

## COMMENTARY

ANALYZING DERIVED STIMULUS RELATIONS  
REQUIRES MORE THAN THE CONCEPT OF  
STIMULUS CLASS

STEVEN C. HAYES AND DERMOT BARNES

UNIVERSITY OF NEVADA AND  
UNIVERSITY COLLEGE CORK, IRELAND

The study of derived stimulus relations has led to new and exciting perspectives on the nature of human verbal behavior (Barnes, 1994; Horne & Lowe, 1996; Sidman, 1994) that go beyond traditional perspectives on the topic within behavior analysis. There is a great deal of conceptual and empirical work to be done, however, to compare and contrast the different approaches being taken by behavior analysts to analyze derived stimulus relations. Our purpose in the present paper is to contrast the emphasis of relational frame theory on the concept of stimulus relation (Barnes & Holmes, 1991; Hayes, 1991, 1994; Hayes & Hayes, 1989, 1992; Hayes & Wilson, 1996) with what we see to be a basic commonality between Sidman and Horne and Lowe: the centrality of the concept of stimulus class.

The concept of stimulus class has played a successful role in behavior-analytic interpretations of many phenomena, so it is not the utility of this concept in general that we wish to question. Rather, we wish to argue that (a) there are important differences between stimulus classes and stimulus relations, (b) popular research methods do not encourage clarity about this distinction, and (c) the attempt to interpret all derived stimulus relations in terms of the stimulus classes that may result is unnecessarily narrow and limits behavioral

approaches to the analysis of language and thinking. Finally, we will suggest methodological alternatives that properly focus attention on the need for relational concepts in the analysis of derived stimulus relations.

*The Dominance of Stimulus Class*

Stimulus classes control common sets of responses because of either physical or functional similarity among a range of stimuli (Donahoe & Palmer, 1994). Class formation can be both a product and a process. For example, stimulus generalization—a term for stimulus class formation based on physical similarity—is universally recognized as a basic behavioral process. Similarly, operant contingencies give rise to functional stimulus classes as products.

This tendency to use class concepts both as products and processes has confused the analysis of derived stimulus relations. By definition, equivalence relations always give rise to stimulus classes as products. The various forms of mutual substitutability of stimuli in equivalence relations (reflexivity, symmetry, and transitivity) are considered by most to be the defining features of the particular class concept called *equivalence classes*. If class formation is also taken to be a process, however, then equivalence classes seemingly require no further explanation, and usually none has been provided. To Horne and Lowe's (1996) credit, they have attempted to go beyond this tautology, but their focus on stimulus classes as the issue is still unequivocal.

The emphasis on class concepts is partly methodological. The matching-to-sample procedure, which has dominated the study of derived stimulus relations, has characteristics that make class-based analyses likely. The re-

---

During the preparation of this article, Dermot Barnes was supported by a grant from the Arts Faculty Research and Development Fund (R250/97), University College Cork.

Requests for reprints may be sent to Steven C. Hayes, Department of Psychology, University of Nevada, Reno, Nevada 89557-0062, or Dermot Barnes, Department of Applied Psychology, University College Cork, Cork, Ireland.

sponse is picking or pointing to a stimulus given another stimulus. If this very performance is taken to be conclusive evidence that the two stimuli involved are in a class, then it is not possible to have consistent performances in a matching-to-sample procedure and *not* have stimulus classes. Because of this methodological characteristic, when matching-to-sample performances are unusual or complex, class-based analyses can always be maintained merely by supposing that there are multiple classes under contextual control. As we will show, this solution seems parsimonious until multiple types of stimulus relations enter the picture. Both Horne and Lowe and Sidman have so far not examined how unwieldy class-based theorizing must become under these circumstances.

*The Challenge of Multiple Stimulus Relations and the Transformation of Functions*

Several researchers have now demonstrated contextually controlled, arbitrary matching-to-sample responding in accordance with multiple stimulus relations such as same, different, opposite, or more than/less than (e.g., Dymond & Barnes, 1995; Steele & Hayes, 1991). The usual procedure involves pretraining with nonarbitrary stimulus sets to establish contextual cues for specific types of stimulus relations, followed by arbitrary matching to sample using these cues. Extraordinarily complex patterns of performances can result from this simple preparation. For example, Steele and Hayes showed that when subjects are trained to pick B3 given A1 in the presence of an OPPOSITE cue (for clarity, in this paper relational cues will be capitalized; in the actual study these cues were nonsense graphical forms), and C3 given A1 and OPPOSITE, they *avoid* C3 given B3 and OPPOSITE, but pick C3 given B3 and SAME. If subjects are also taught to pick D1 given C3 and OPPOSITE, they now pick D1 given C3 and OPPOSITE, but avoid it given SAME. How are such results handled, especially as more and more complex and nonsymmetrical relations (e.g., more than/less than; before/after) are added to the mix?

*Sidman's Answer: Contextually Controlled Equivalence Classes*

It is possible to maintain a class-based account of such results. The only place Sidman

directly addressed the issue of multiple stimulus relations shows the steps that need to be taken. According to Sidman, "the fact that a stimulus pair can be brought via contextual control into such differing relations as same, opposite, different, and so forth, can be handled by any formulation of equivalence that recognizes the role of context" (1994, p. 561).

This statement shows what needs to be done to accommodate the concept of stimulus class to the data on multiple stimulus relations. First, we must be willing to use the term *class* to indicate the reliable selection of one stimulus given another stimulus, because the classes that result need have none of the defining features of equivalence classes (e.g., transitivity). With this use of the term any reliable matching-to-sample result is indicative of a class by definition: The overall patterns of stimulus relations are thereby merely equivalence relations under contextual control. Thus, if the subject selects C3 given A1 and OPPOSITE, the subject is not relating the two as opposite (to agree to that is to overthrow the primacy of classes in the analysis of stimulus relations). Rather, the subject is putting C3 and A1 in a class under the control of an OPPOSITE contextual cue.

Second, the *pattern* of contextual control need not itself be explained. For example, the results of Steele and Hayes (1991) show that OPPOSITE cues functioned in the following way: Stimulus pairs selected in the presence of OPPOSITE that were an odd number of nodes away are now selected (are in a class) given SAME, and stimulus pairs selected in the presence of OPPOSITE that were an even number of nodes away are selected (are in a class) given OPPOSITE. How can OPPOSITE acquire this ability to organize classes in this way merely because subjects are pretrained to pick formally opposite stimuli in the presence of it? Sidman does not provide an account, but if contextual control over an innate equivalence process is learned (Sidman, 1994; see Barnes, 1994, p. 94), presumably he would have to appeal to a history with such contextual cues and with the pattern of results they establish. For example, individuals would have to learn the relations whereby if Stimulus A is more than Stimulus B, then B is less than A. But how would this be learned? First, it would seemingly have to

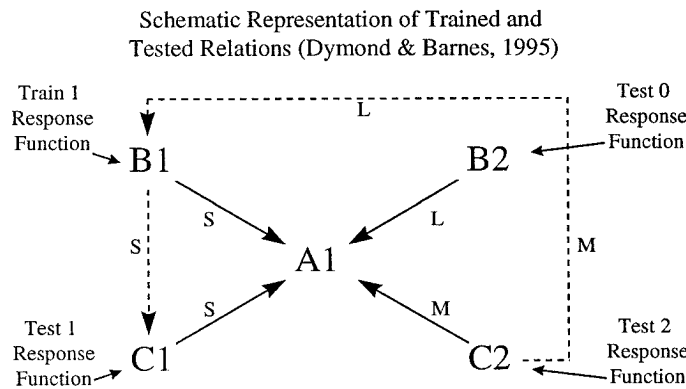


Fig. 1. Schematic representations of the most important of the trained (solid lines) and tested (dashed lines) relations in the Dymond and Barnes (1995) study. S, M, and L indicate the arbitrarily applicable relations of sameness, more than, and less than. The relational network from the Dymond and Barnes study also shows that a one-response function was trained using the B1 stimulus, and tests examined the transformation of the trained self-discrimination response function in accordance with the relations of sameness (C1, one response), more than (C2, two responses), and less than (B2, no response) (adapted from Dymond & Barnes, 1995; copyright the Society for the Experimental Analysis of Behavior, Inc., reprinted by permission).

be learned through the same type of history that, according to relational frame theory, is necessary for establishing multiple stimulus relations. Second, this learning history would have to establish cues that would have different functions in different parts of a stimulus network (e.g., more than/A  $\rightarrow$  B is correct, but more than/B  $\rightarrow$  A is incorrect). These functions are themselves difficult to interpret in class terms, and thus the overall pattern of contextual control still must be accounted for. Merely moving the issue of multiple stimulus relations to a vague appeal to contextual control is not an advance.

The limits of the concept of stimulus class in explaining the results of multiple stimulus relations are revealed even more clearly if we take the additional step of examining how stimulus functions transfer through these multiple stimulus relations. This was done in a study by Dymond and Barnes (1995), who used procedures similar to those of Steele and Hayes (1991) to first train control by the relations of same, more than, and less than with nonarbitrary stimulus sets (e.g., subjects were trained to choose a two-star comparison in the presence of a three-star sample given the LESS THAN cue). When the subjects had successfully completed the pretraining, they were then trained in six arbitrarily applicable relations using the three contextual cues. The four critical relations were SAME/A1-B1, SAME/A1-C1, LESS THAN/A1-B2, and

MORE THAN/A1-C2. The subjects were then tested for seven derived relations, the following three relations being the most important: SAME/B1-C1, MORE THAN/B1-C2, and LESS THAN/B1-B2 (see Figure 1, upper section).

Three response patterns were then shaped via a complex schedule of reinforcement: (a) no response, (b) one response only, and (c) two responses only. The subjects were also trained to pick different stimuli given the response patterns they had been producing, rather than given particular stimuli as samples. Dymond and Barnes (1995) predicted that if choosing Stimulus B1 after making one response was reinforced, a subject, without further training, may then choose (a) C1 following one response (i.e., C1 acquires the same function as B1), (b) B2 following no response (i.e., B2 acquires a response function that is less than the B1 function), and (c) C2 following two responses (i.e., C2 acquires a response function that is more than the B1 function; see Figure 1, upper section). In fact, all 4 subjects demonstrated this predicted transformation of self-discrimination functions (see Dymond & Barnes, 1996, and Roche & Barnes, 1996, for related empirical research). The term *transformation*, rather than *transfer*, is necessary to describe the Dymond and Barnes data, because the explicitly trained one-response function of B1 transferred to C1, which was in an equivalence re-

lation with B1, but not to B2 and C2, which were not. B2 and C2 did not acquire one-response functions. Instead, the one-response function of B1 was transformed in accordance with more than and less than relations between the stimuli, such that B2 and C2 acquired zero- and two-response functions, respectively.

It is difficult to predict, or even to describe, the test performances reported by Dymond and Barnes (1995) in terms of equivalence or other stimulus classes. Different functions emerged for C1, B2, and C2, and the functions seen were in accord with the derived relations between these stimuli and B1. The account in terms of relational frame theory is straightforward, but an account based on stimulus class seems to require that we invoke three separate classes (i.e., one for each function). These three cannot all be equivalence classes, however, because the functions did not change if the underlying relation was an equivalence relation but did change in the other two cases. Moreover, simply invoking three different classes would not allow us to predict the specific transformation of functions observed in the study (i.e., even if B1, B2, and C2 were members of three distinct classes, establishing a one-response function for B1 leaves the untrained functions of B2 and C2 unspecified). Perhaps the stimuli could all be in an equivalence class in which the function transformation was controlled, to some extent, by the nodal distances between the stimuli participating in the class (see Fields, Adams, & Verhave, 1993), but that would not explain the derived stimulus relations seen in matching to sample or the direction of the changed self-discrimination functions, because B2 and C2 were both removed by one node (i.e., A1) from the B1 stimulus (see Figure 1, upper section). Dymond and Barnes examined two further class-based accounts of their data, separable stimulus compounds and ordinal classes, and also found these to be inadequate (Dymond & Barnes, 1995, pp. 182–183).

The core of the problem that multiple stimulus relations present to a class-based account is this: If any relation between a stimulus pair is defined in terms of a contextually controlled equivalence class, the mathematical definition of equivalence in terms of reflexivity, symmetry, and transitivity must col-

lapse, and with it the need for equivalence as a concept (see Barnes & Roche, 1996). Consider, for example, a subject who is trained to choose B in the presence of A given a MORE THAN context. Presumably this subject should not then choose, in the same context, A given B (i.e., A is in fact less than B) or A given A or B given B (a stimulus cannot be more than itself). In this particular example, therefore, we are left with neither symmetry nor reflexivity as defining properties of the subject's pattern of responding to the stimulus pair. If any two stimuli that go together, or are partitioned into a set, are members of a contextually controlled equivalence class, we are thereby left with no additional specific pattern of behavior that can be isolated as characteristic of an equivalence class. In an effort to retain *class* as the core outcome concept, and equivalence classes as its manifestation in the case of derived stimulus relations, the distinction between equivalence and other classes has to be undermined. We are left with the term *equivalence*, but the concept itself has broken down.

Evidence for this breakdown is accumulating. Equivalence researchers are suggesting new forms of equivalence classes (e.g., ordinal/sequential; see Green, Stromer, & Mackay, 1993). Sidman has backed away from such key points as his four- or five-term contingency analysis (Sidman, 1994, pp. 378–379) and stimulus selection as the basis for equivalence class formation (Sidman, 1994, p. 399). As we have tried to show above, Sidman's view of nonequivalence relations as contextually controlled equivalence classes (Sidman, 1994, p. 561) undermines his own set theory definition of equivalence in terms of reflexivity, symmetry, and transitivity. That leaves only the concept of partition or class standing intact in the traditional approach to equivalence, and that was a concept we had before equivalence arrived. We believe that this inconsistency and process of decomposition show a fundamental error: The refusal to go beyond "stimulus class" has hamstrung the analysis of derived stimulus relations.

#### *Horne and Lowe's Answer: Naming and Stimulus Classes*

In contrast to Sidman, Horne and Lowe (1996) present a much more elaborate view of stimulus equivalence, although it is one

with considerably less empirical support. In agreement with relational frame theory, they see stimulus class formation as the result of operant activity, but in a major point of disagreement, the end result is still simply a stimulus class.

Lowe and Horne suggest that naming emerges, as a higher order behavioral relation, when listener and speaker behavior combine (Lowe & Horne, 1996, p. 315). In their article, they propose in great detail how a young child's interaction with the verbal community might establish and maintain generalized or higher order classes of listener and echoic behavior. These behavioral repertoires then combine, according to Horne and Lowe, to produce the higher order name relation. In their own words,

When listener and echoic relations combine in the presence of particular objects or events, this creates the conditions for the emergence of a new response class of speaker behavior that is directly evoked by these objects and events. Thus, objects now give rise to speaking and then to listening, that is, reorienting to the objects, which in turn reevokes speaker behavior and so on. This closes the circle and establishes a functional unity of these three generalized classes of behavior. At this point the higher order name relation is established. . . . the first instances of this new unit are explicitly reinforced by caregivers. What is now reinforced, of course, is the behavior class as a whole. With each reinforced repetition of the name relation, perhaps as new object class members are encountered (e.g., a new dog, a new chair), naming as a functional higher order class is further strengthened. Thereafter, explicit reinforcement by caregivers for new name relations becomes less important as the automatic reinforcing consequences of naming things become the more potent source of control. (Lowe & Horne, 1996, pp. 317–318)

Having outlined how naming is established in the behavior of a young child, Horne and Lowe (1996) then attempt to use the concept of naming to explain the formation of stimulus classes. In particular, they suggest that the name relation helps to explain the formation of functional stimulus classes (pp. 204–206) and stimulus equivalence classes (e.g., p. 207). With regard to equivalence classes, Horne and Lowe suggest that naming may produce an equivalence class via common naming (pp. 215–218), via intraverbal

naming (pp. 218–221), or via some other verbal behavior (pp. 221–222). The details of these explanations for the formation of equivalence classes are not important here. What matters is that in focusing on naming explanations for the formation of stimulus classes per se, Horne and Lowe have yet to provide an explanation for behavior that is not easily described in terms of stimulus classes alone. It is not clear, for example, why a named stimulus relation might operate as a stimulus relation merely because it is named.

Horne and Lowe were challenged on this very issue when they were asked to “explain the Steele and Hayes data using their approach” (Hayes, 1996, p. 311). They did not rise to the challenge other than to say that subjects “will have used verbal behavior (i.e., names and rules) to solve the problems posed” (Lowe & Horne, 1996, p. 333). Horne and Lowe apparently believe that the derivation of multiple stimulus relations requires not just naming but also rule governance, a level of verbal ability that extends well beyond the name relation (Horne & Lowe, 1996, pp. 212–213).

This claim by Horne and Lowe is hard to evaluate, especially because rule governance is even less well understood, both empirically and conceptually, than derived stimulus relations. But some form of empirical evaluation may be possible over the short term, because Horne and Lowe seem to be invoking forms of verbal control that are fairly elaborate. If complex multiple stimulus relations can be readily established in very young children, a simpler explanation may be warranted.<sup>1</sup>

If Horne and Lowe's account is to be parsimonious, they must explain the derivation of multiple stimulus relations using naming only, and that they have not done. It is not enough merely to explain derived stimulus relations by a reference to verbal behavior. The precise verbal performances needed, and their source, must be specified. If the performance is explained by rules, Horne and Lowe must specify exactly how naming gives rise to rule governance, and then how rule governance gives rise to the derivation

<sup>1</sup> Vaughan, M., & Barnes, D. (1994, November). *Children's perception of pain and bullying: A relational-frame interpretation*. Paper presented at the Annual Conference of the Psychological Society of Ireland, Kilarney, Ireland.

of multiple stimulus relations. If this explanation is not cast in terms applicable to very young children, and multiple stimulus relations are then shown in that population (see footnote 1), Horne and Lowe's account is disconfirmed.

We cannot see how naming alone—as described in Horne and Lowe's (1996) account—can produce multiple stimulus relations, in part because the end product of naming is merely a stimulus class. All of our objections in the previous section about the limitations of that concept as the explanation for multiple stimulus relations apply with equal force to classes formed via naming. We do not deny, of course, that subjects may name the relations involved, but in terms of relational frame theory this shows only that stimulus relations themselves can enter into a “frame of coordination” with a name. It does not show why named relations relate. In the same way that Sidman relies on contextual control to address the issue, Horne and Lowe rely on naming, but both fail to deal with relational operants directly and in detail, or to show why multiple relational operants cannot be learned the way most operants are learned.

Relational frame theory accounts for all of the existing naming data quite readily and without modification. It does so by allowing for the acquisition of a variety of relational operants, not just one (in the case of Horne and Lowe) or none (in the case of Sidman). When stimulus relation is the core concept, equivalence classes can sometimes result (e.g., through frames of coordination), but sometimes they do not.

#### *Developing Methods for Examining Stimulus Relations*

The matching-to-sample procedure emphasizes class formation as the product because the results can always be analyzed as a matter of partitioning stimuli. The concept of stimulus relation is built into basic behavior analysis (in concepts like *contingency*, for instance) but the methods that exist for the assessment of nonarbitrary stimulus relations have not been adequately applied to those that are abstracted and brought under contextual control.

Some methodological advances might be made by examining these methods. For ex-

ample, classical conditioning is a kind of stimulus relation, but respondent-type training procedures have only been applied to equivalence classes quite recently (Barnes, Smeets, & Leader, 1996; Leader, Barnes, & Smeets, 1996; Smeets, Leader, & Barnes, 1997). The basic procedure involves presenting an arbitrary Stimulus A that reliably predicts the appearance of a second arbitrary Stimulus B (i.e.,  $A \rightarrow B$ ; note that A and B are never presented simultaneously). At least two trial types may be trained initially (e.g.,  $A1 \rightarrow B1$  and  $A2 \rightarrow B2$ ). Following sufficient exposure to this respondent training procedure, subjects (both adults and children as young as 5 years old) presented with these stimuli in a matching-to-sample procedure normally pick Stimulus A1 (as a comparison) in the presence of Stimulus B1 (as a sample); similarly, they pick Stimulus B2 given A2. In effect, having been exposed to  $A \rightarrow B$  respondent training, adults and children respond in accordance with the B-A symmetry relation. Furthermore, when the same subjects are exposed to  $A \rightarrow B \rightarrow C$  respondent training, they normally respond in accordance with C-A equivalence relations on a matching-to-sample test.

These findings show that “picking” is not necessary when training for equivalence class formation, but they are only a first step, because they are designed to examine only equivalence classes; furthermore, a matching-to-sample (equivalence) test was used to examine the effects of the training. But this same approach can be expanded into a method designed to assess multiple stimulus relations. We believe this is worth an extended discussion, because the failure to go beyond the class concept seems so much determined by the methodological narrowness of the entire area.

#### *The Relational Evaluation Procedure and the Derived Stimulus Relations of Before and After*

The new method described here is called the *relational evaluation procedure* because it was designed to allow subjects to report on, or evaluate, the stimulus relation or relations that are presented to them on a given task. This present example is focused on the derived stimulus relations of before and after, but a similar approach can be used for vir-

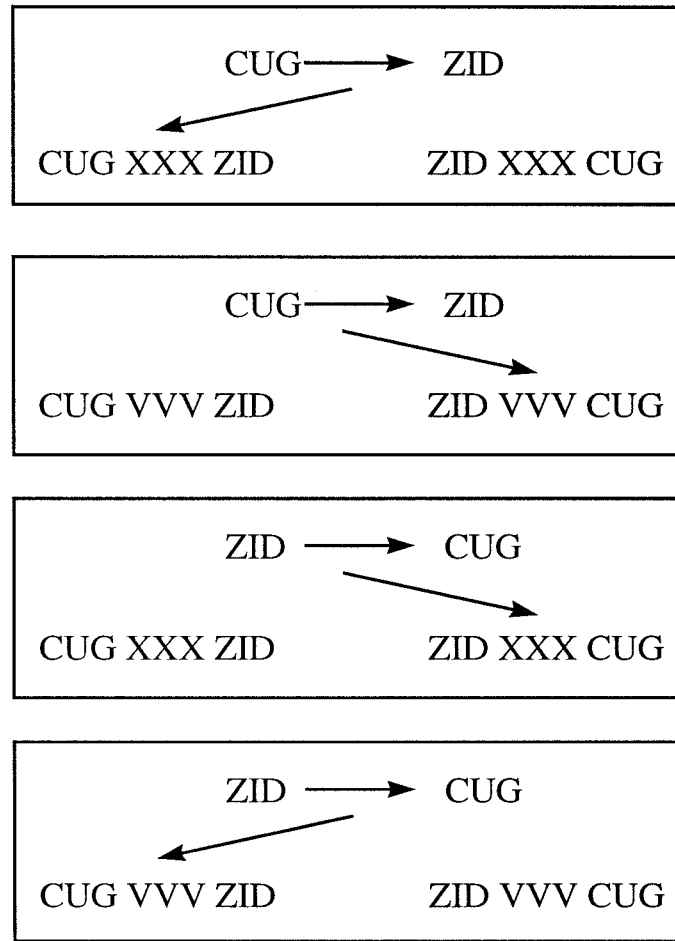


Fig. 2. Schematic representation of the four tasks used to establish contextual control by the before and after cues (see text for details).

tually any relation. The core of the method is this: Establish a procedure in which subjects may confirm or deny the applicability of particular stimulus relations to sets of stimuli. In so doing, the focus shifts from stimulus partitioning and picking (with its class connotations) to relational specification and evaluation.

*Before/after training.* On each trial, a nonsense syllable (e.g., CUG) appears in the center of a computer screen for 1 s, disappears for 1 s, and is followed by a second syllable (e.g., ZID) for 1 s. The presentation of these two stimuli, one after the other, functions as a sample stimulus, and following a further 1-s delay, two three-element comparison stimuli appear on the screen, one in the lower left corner and the other in the lower right cor-

ner. Reading from left to right, both comparisons (which we will term here *statements*) contain the nonsense syllable just displayed (e.g., CUG), an arbitrary relational contextual cue (e.g., XXX or VVV), and the other nonsense syllable just displayed (e.g., ZID). Subjects must select one of the two comparison statements, and are then given feedback.

Examples of four such tasks are shown in Figure 2. On one of the trials in which CUG was shown before ZID in the first part of the trial, for instance, "CUG XXX ZID" is correct but "ZID XXX CUG" is not, and "ZID VVV CUG" is correct but "CUG VVV ZID" is not. In this fashion, XXX is treated as functionally equivalent to "before" and VVV to "after." What is critical for our present purposes is that, like statements in natural language, the

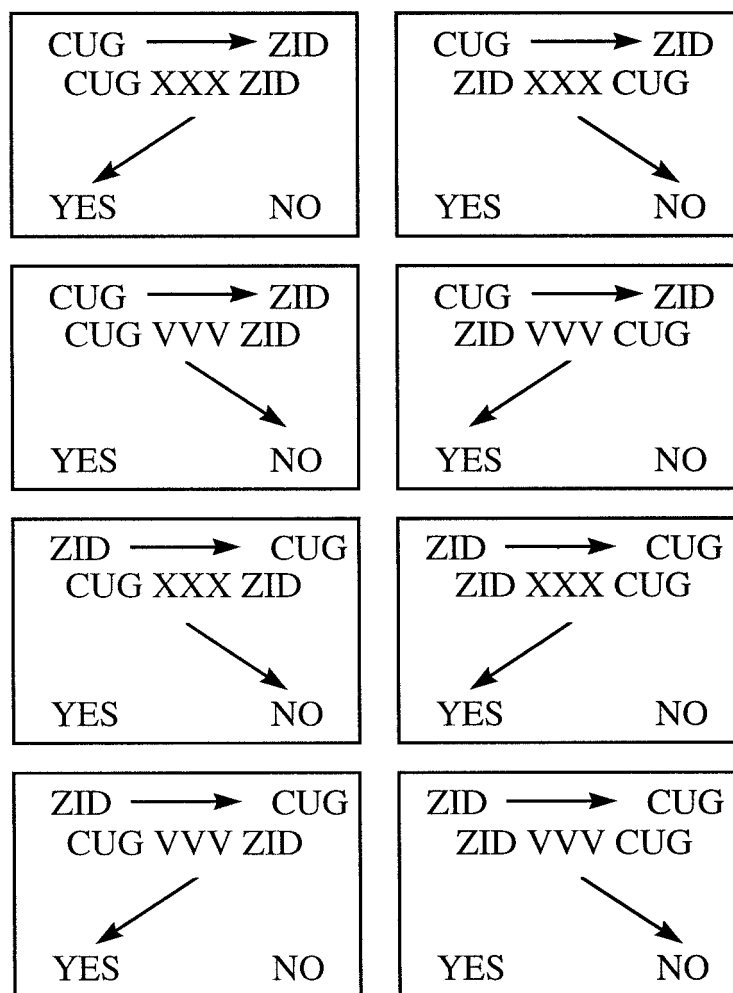


Fig. 3. Schematic representation of the eight tasks used to train "yes" and "no" functions in combination with the before and after cues (see text for details). During a subsequent test phase (no feedback) the same eight tasks were used, but the CUG and ZID stimuli were replaced with novel nonsense syllables. Note that in the actual procedure "yes" and "no" are arbitrary stimuli, not known words.

correctness of the items cannot be distinguished on the basis of the two nonsense syllables or on the basis of the relational contextual cue, but only on the relation among all of these to some state of affairs.

*Evaluation of statements.* Once the meaning of the relational contextual cue is established, these statements can be affirmed or denied, rather than picked from a set of comparison statements (see Figure 3). For example, if ZID is shown before CUG, the statement "CUG XXX ZID" is incorrect, and reinforcers are contingent upon selection of one of two novel nonsense syllables (that is thus functionally equivalent to "no"; on other tri-

als the "yes" stimulus is selected, as is shown in Figure 3).

Once such a procedure is established, subjects can be trained and tested on entirely new sets of stimulus relations using the "yes" and "no" stimuli. As in natural language, stimulus relations can even be trained without any explicit overt responding at all (e.g., "YES" TUH VVV FRE might establish that FRE came first). Any relational stimulus can be similarly trained, provided only that pre-training exemplars can be provided. Note also that this procedure does not require any instructions that would preclude its use with nonhumans.

We have only recently begun working with the relational evaluation procedure, but our preliminary research has shown that subjects can produce such performances with ease.<sup>2</sup> The flexibility of the procedure can lead quickly to remarkably complex behavior. On one typical trial, for example, the sample stimulus consists of the following string of characters, "CUG XXX VEK XXX ZID XXX YIM XXX DAX," which appears above the string "CUG XXX YIM VVV VEK" (XXX and VV have previously been established as before and after). The subject is required to choose between the "yes" and "no" nonsense syllables to indicate whether the lower string may be derived from the top string. In this case, the correct choice is "yes."

Our point in reviewing the relational evaluation procedure is to show how much our current methods emphasize stimulus classes over stimulus relations, and to give an example of what methods focused on stimulus relations might look like. It is extremely difficult to think in terms of stimulus classes with the resulting relational performances. Consider, for example, the following test performance:  $C \rightarrow D/D$  XXX C, pick "no." When the subject chooses the "no" nonsense syllable, should we define it as participating in an equivalence class with D before C? If so, then its participation in this class must be under a complex form of contextual control, because on other tasks D before C controls picking "yes," and on yet other tasks D after C and C after D also control picking "no."

We are not suggesting here that a class-based account could not be constructed for these data, but we question the functional utility of doing so, particularly as more complex tasks with multiple relations are used (as indeed, we would argue, is occurring this moment as readers read this material: We would argue that the reader is deriving stimulus relations among verbal classes, not merely partitioning stimuli into classes). In contrast to a class-based account, however, consider the relative ease with which the same data may be interpreted in terms of multiple stimulus relations. From this perspective, the subject is

presented, on any given trial, with a relational network (e.g., the statement) that is to be compared to what is known about the network elsewhere. Thus, when a subject chooses "no" when presented with  $C \rightarrow D/C$  VVV D, the response is controlled by the nonarbitrary temporal relation C before D, the arbitrarily applicable relation VVV being applied to C and D, and the relation of distinction that obtains between the two (i.e., the nonarbitrary relation C before D is different than the arbitrarily specified C after D). This simple formulation in terms of a relational network may be applied with relative ease to any of the tasks outlined above, or to their more complex forms. It is hard to think of performances on the relational evaluation procedure in simple partitioning or class terms, because what distinguishes correct and incorrect responses is the applicability of stimulus relations, not mere stimulus partitions.

#### Conclusion

Relational operants are classes of operant behavior, but their results are abstracted and arbitrarily applicable stimulus relations, not necessarily stimulus classes. The concept of stimulus class, which has an important place in behavior analysis, needs to be added to, not thrown over. It seems better to make the needed conceptual and methodological modifications than to deal with the narrowness, decomposition, and distortion that is resulting from an attempt to hold to the concept of stimulus class as the single organizing principle and result in the area of derived stimulus relations.

#### REFERENCES

- Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44*, 91-124.
- Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record, 41*, 19-31.
- Barnes, D., & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders. *The Psychological Record, 46*, 489-507.
- Barnes, D., Smeets, P. M., & Leader, G. (1996). Procedures for generating emergent matching performances: Implications for stimulus equivalence. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals: Advances in psychology* (pp. 153-171). Amsterdam: Elsevier.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and*

<sup>2</sup>Hegarty, N., & Barnes, D. (1996). *Responding in accordance with the relational frame of before and after: An experimental analysis*. Paper presented at the annual conference of the Experimental Analysis of Behaviour Group, London.

- complex behavior*. Needham Heights, MA: Allyn & Bacon.
- Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior*, *64*, 163–184.
- Dymond, S., & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record*, *46*, 271–300.
- Fields, L., Adams, B. J., & Verhave, T. (1993). The effects of equivalence class structure on test performances. *The Psychological Record*, *43*, 697–712.
- Green, G., Stromer, R., & Mackay, H. A. (1993). Relational learning in stimulus sequences. *The Psychological Record*, *43*, 599–615.
- Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior: The first international institute on verbal relations* (pp. 19–40). Reno, NV: Context Press.
- Hayes, S. C. (1994). Relational frame theory: A functional approach to verbal events. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 9–30). Reno, NV: Context Press.
- Hayes, S. C. (1996). Developing a theory of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, *65*, 309–311.
- Hayes, S. C., & Hayes, L. J. (1989). The verbal action of the listener as a basis for rule-governance. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 153–190). New York: Plenum.
- Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist*, *47*, 1383–1395.
- Hayes, S. C., & Wilson, K. (1996). Criticisms of relational frame theory: Implications for derived stimulus relations. *The Psychological Record*, *46*, 221–236.
- 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.
- Leader, G., Barnes, D., & Smeets, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record*, *46*, 685–706.
- Lowe, C. F., & Horne, P. J. (1996). Reflections on naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*, 315–340.
- Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and human sexual categorization: A critical test of the derived difference relation. *The Psychological Record*, *46*, 451–475.
- Sidman, M. (1994). *Stimulus equivalence: A research story*. Boston: Authors Cooperative.
- Smeets, P. M., Leader, G., & Barnes, D. (1997). Establishing stimulus classes with adults and children using a respondent training procedure: A follow-up study. *The Psychological Record*, *47*, 285–308.
- Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, *56*, 519–555.

---

### THE ROLE OF JOINT CONTROL IN THE DEVELOPMENT OF NAMING

BARRY LOWENKRON

CALIFORNIA STATE UNIVERSITY, LOS ANGELES

---

In my earlier comments (Lowenkron, 1996), I pointed out that Horne and Lowe's (1996) account of the naming relation seems to be deficient in explaining how novel stim-

uli come to be selected in response to their names after the names are learned as responses to the stimuli. I also suggested that this deficiency could be remedied, and several strengths could be gained, by appreciating the role *joint control* plays within the naming relation. Lowe and Horne (1996, p. 318), however, assert that applying the joint control account to the naming relation involves two problems: first, that it engenders an anachronism with respect to the order in which the

---

Preparation of this article was supported by Grant HD 29232 from the National Institute of Child Health and Human Development. I thank Michael Hixon for his clarifying comments on some issues that appear here.

Address correspondence to Barry Lowenkron, Department of Psychology, California State University, Los Angeles, Los Angeles, California 90032.

component responses develop, and second, that the notion of joint control is merely redundant with the notion of the naming relation. I argue here that neither assertion is correct. Interpreting the development of generalized stimulus selection in terms of the role of joint control does not involve an anachronism, nor is the account redundant with the naming account. Rather, appreciating the role of joint control allows for an account that is significantly more explicit and far more general than the naming account.

*Does the Joint Control Account  
Stumble on an Anachronism?*

As described earlier (Lowenkron, 1996, p. 253), when it is under joint control, the selection response is actually an *autoclitic* that reports which comparison stimulus brings the currently rehearsed sample topography under joint control. Thus, as illustrated there (Lowenkron, 1996, Figure 1, Panel A), when a person is vocally instructed to find the printed number sequence 135476 from a list of six-digit sequences, he or she will rehearse that topography (“135476”) as a self-echoic while perusing the list. When the correct sequence is encountered, any further rehearsals of the topography occur under joint control, for now, in addition to self-echoic control, the rehearsed topography is also (i.e., jointly) emitted under the *tact* control produced by the printed sequence on the page. By pointing to the printed number sequence that initiated this change from self-echoic to joint self-echoic/tact control, the person selects the sequence specified by the original spoken numbers. The pointing response is thus an autoclitic, because it reports to the observer about an event (the onset of joint control) that affected the verbal behavior (the rehearsed topography) of the speaker (Lowenkron, 1991).

Now Horne and Lowe (1996) document the fact that children learn to select objects in response to spoken words before they learn other verbal responses. Because the joint control account has the selection response as an autoclitic, Lowe and Horne (1996) observe that the joint control account seems to require that children learn autoclitic responses before they learn other responses. This, they argue, would involve an anachronism because autoclitics report

about stimulus control over other responses. Therefore, these other responses must develop first.

Their observation is not unreasonable given the brevity of my earlier account, but neither is it correct. The actual course of development of joint control, and its relation to the naming relation, may be seen most clearly by tracing through Horne and Lowe’s (1996) own account. Thus, the first panel in Figure 1 illustrates the three primary verbal relations and their order of occurrence as documented by Horne and Lowe. The first repertoire infants acquire consists of the *unmediated selection* of stimuli in response to their names. Later (up to several months), children acquire vocal responses in the form of echoic and self-echoic behavior (the second repertoire) and tacts (the third repertoire).

Now we come to the crucial question. What happens next, so as to move the child from these simple operants to the generalized semantic processes attendant to language usage? Horne and Lowe (1996, p. 200) provide one account by postulating the emergence of the higher order naming relation. Another possibility is illustrated in Figure 1 (Panel II): In an environment more complex than the one in which the unmediated selection response was originally trained (i.e., with more alternatives available, or with delays until named objects appear), the sample name is given. To maintain the sample in this complex environment, the child rehearses it as a self-echoic (Panel IIA) while searching for the named object (Michael, 1996). When the sought-after object is encountered, the onset of joint control occurs as the child emits a self-echoic that is now also a tact for that object (Panel IIB).

Here is where the anachronism is resolved. As a result of the prior training in unmediated selections of this and other objects in response to their names (e.g., Figure 1, Panel I), the child now selects the named object here, and doing so is reinforced (Panel IIC). As a result, the selection of an object that evokes a tact that enters into joint control with the currently rehearsed self-echoic is adventitiously reinforced (Panel IID). If this same process were to continue with many other objects, whereby the subject rehearses their names while seeking them out, emits the name under joint tact/self-echoic control

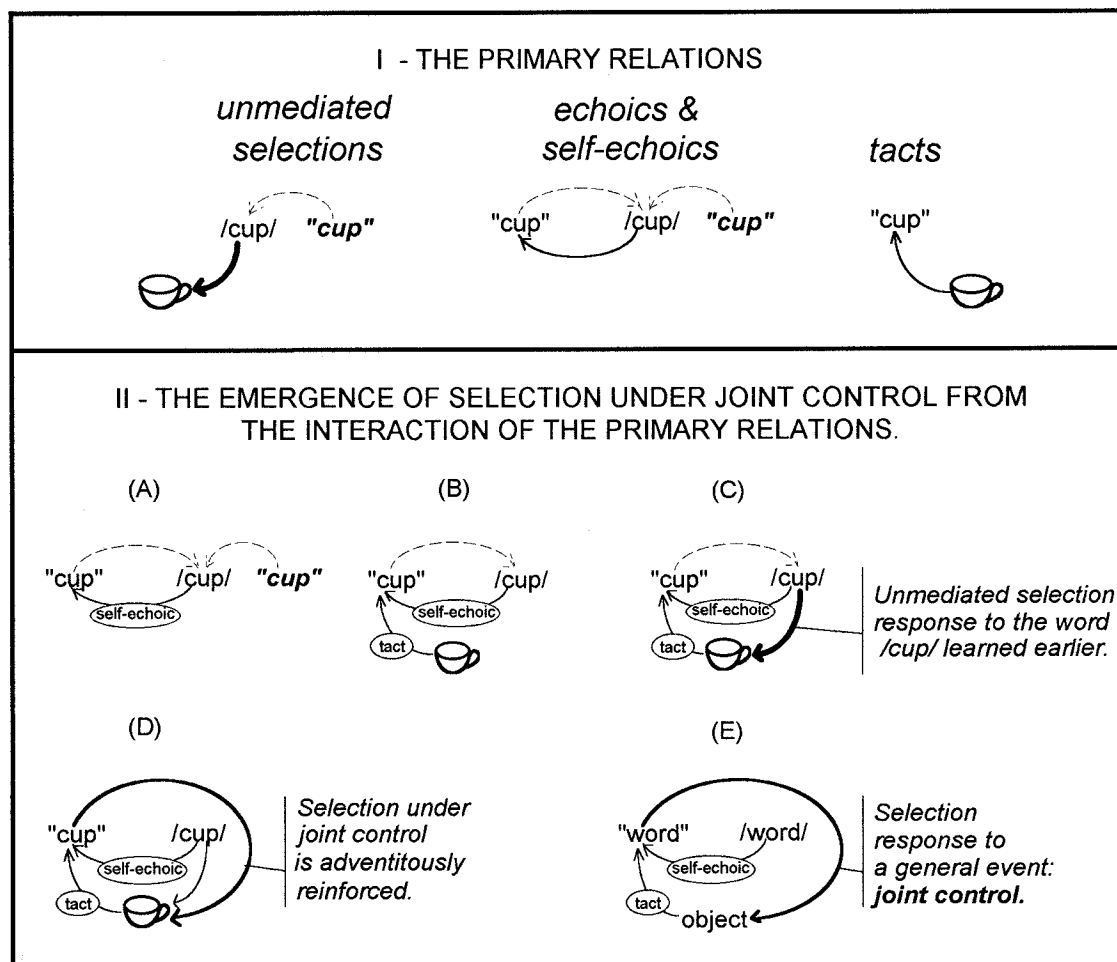


Fig. 1. The development of selection under joint control. Words spoken are denoted as "cup," and words heard by the subject are denoted as /cup/. Solid lines indicate responses, and dashed lines indicate auditory productions. Panel I: the three primary verbal repertoires. Panel II: the interaction of the three primary verbal repertoires to produce selection under joint control. (A) In a selection task, upon hearing /cup/, the child rehearses it as a self-echoic. (B) When a cup is encountered (for which the child previously acquired the word *cup* as a tact), the child's rehearsal now occurs under joint tact/self-echoic control. (C) As a result of the prior training in the unmediated selection of the cup in response to the heard word /cup/, the child now selects the cup. (D) Selecting the cup results in adventitious reinforcement for selecting a stimulus that evokes a tact that enters into joint control with the topography of the current self-echoic. (E) After experience selecting several stimuli under joint control, stimulus control of the selection response transfers from the unmediated selections illustrated in Panel I to the generic joint control event, so that any object evoking a tact that enters into joint control with the current self-echoic will be selected by the now autoclitic selection response, thereby producing generalized selection (the name-object relation) to any stimulus that evokes a consistent tact (the object-name relation).

when the object is encountered, and selects the object because of the earlier, unmediated, selection training, one could reasonably expect that stimulus control of the selection response would transfer from the individual objects to the generic, joint control event itself (Panel IIE). There is thus no anachronism:

Before the requisite tacts and echoics have been acquired, selection responses are unmediated: They are evoked directly by the named objects. Selection under joint control, by an autoclitic response, only occurs after the tacts and self-echoics have been acquired and have come to interact.

*Is the Joint Control Account  
Redundant with Naming?*

The second point Lowe and Horne (1996) raise, regarding the redundancy of the joint-control account and their own account, may be answered by noting that the former is both more parsimonious and more general. Thus, unlike naming, the joint control account finds the varieties of speaker behavior described by Skinner (1957) entirely sufficient to describe symbolic behavior in the listener. In doing so, it treats the development of semantic language competence strictly in terms of the acquisition of specifiable operants and changes in stimulus control over these operants. Rather than appealing to generalized response classes and higher order relations, under joint control, generalization is the result of the transfer of stimulus control to a generic event: the onset of joint control.

Aside from its parsimony, the mechanism of joint control is both more efficient and more plausible. On the naming account, each time the child learns a new tact, both listener and speaker behavior for the name-object pair must be practiced by implicit rehearsal as some higher order naming relation is activated (Lowe & Horne, 1996, pp. 317–318). Thus, all novel behavior is considered to be the product of prior covert rehearsal. The joint control account, on the other hand, makes no such demands, either mechanical or conceptual. On this account, only tacts need be acquired. Generalized selection occurs de novo whenever one of these novel tacts enter into joint control with the child's self-echoic rehearsal of the name of the stimulus currently sought. Efficiency is thus gained in this account because it assumes less practice on the part of the subject to produce the same behavior as the naming relation, and because truly novel behavior can emerge here untrained and unrehearsed. Plausibility is gained by not positing performances (covert rehearsal) that we do not observe in our own behavior as we learn new words.

Beyond its greater efficiency, joint control allows a more general account than does naming. Thus, as illustrated earlier (Lowenkron, 1996, Figure 1, Panel B), joint control easily accounts for responding that involves prepositions describing relations between stimulus elements, something the naming re-

lation seems not to address. Thus, the selection of a dot in a circle in response to that description only requires the subject to select the comparison that evokes the phrase *dot in circle* as a tact, while the child rehearses the description as a self-echoic. Under joint control, selection in response to phrases that contain prepositions only requires that the child be capable of accurately tacting with those prepositions.

There is more: The role of joint control has been documented across a wide range of generalized relations (e.g., identity matching; matching based on order [before/after], size, orientation, and oddity) as well as in generalized goal-oriented behavior (Lowenkron, 1984, 1988, 1989; Lowenkron & Colvin, 1992). Joint control also provides an explicit account of many sorts of symbolic and conceptual behavior that have heretofore resisted a rigorous behavioral account. All of this suggests that replacing the notion of higher order classes with joint control as the mechanism of semantic relation provides a significant improvement in the parsimony, the efficiency, the plausibility, and the generality of the account.

## REFERENCES

- 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.
- Lowe, C. F., & Horne, P. J. (1996). Reflections on naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*, 315–340.
- 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., & Colvin, V. (1992). Joint control and generalized nonidentity matching: Saying when something is not. *The Analysis of Verbal Behavior*, *10*, 1–10.
- Michael, J. (1996). Separate repertoires or naming? *Journal of the Experimental Analysis of Behavior*, *65*, 296–298.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

## AN ARCHEOLOGY OF MEANING

ERNST L. MOERK

CALIFORNIA STATE UNIVERSITY, FRESNO

A potentially creative tension, a critical commentator might call it a fault line, runs through the project of Horne and Lowe (1996): The matching-to-sample paradigm, a “highly artificial . . . odd experimental paradigm,” as the authors themselves call it (p. 238), confronts an immense problem area of naturalistic meaning and reference. Such a discrepancy in breadth is not necessarily a disadvantage in the first stages of an analytical investigation, yet the chasm separating this study from most discussions of semantics will have to be bridged, because the authors themselves endeavor to “foster productive interaction with scientists from other traditions” (p. 186). For such bridging, the existence of scholars from other fields and the potential relevance of their work will have to be acknowledged. The “archeology” aspect indicated in the above title, referring to older contributions from a variety of fields, is intended to contribute to such interaction. Three aspects are emphasized: (a) a bridge to naturalistic semiotic performance in animals, (b) additional research on children’s semiotic development, and (c) a bridge to philosophical and psycholinguistic approaches to meaning.

*Stimulus Equivalences and Semiotic Performance in Animals*

I was astonished to read in Horne and Lowe (1996) that stimulus equivalences could not be established in animals. It would have seemed to me that no higher animal (whether predator or prey) could survive without being a quick learner of stimulus equivalences (e.g., that the smell, sound, sight, tracks, excreta of a predator mean danger). Admittedly, these are *functional equivalences* (p. 190) only and do not conform to the operational definition of *stimulus equivalence*, in that they do not entail all three features of the math-

ematical concept of equivalence. Yet, when a prey animal learns the equivalence between predators and the sounds they produce as well as the equivalence of predators and their scents, it will sniff when hearing this sound and search visually for the predator’s location (cf. Rubenstein & Wrangham, 1986, *passim* for discussions of “predator vigilance”), similar to a child engaging in listener behavior after being asked “Where is the teddy bear?” Such gazelle behavior appears to be similar to the emergent equivalences BC and CB in Figure 2 of Horne and Lowe. Following the authors’ reasoning (p. 187), I would therefore suggest that the gazelle has formed a stimulus class that seems to involve symmetry and transitivity, even if this has not been observed in the laboratory.

The above pertains obviously only to *perceptual* functioning, and one of the main aspects of symbolic behavior is the production of a member of the equivalence class. Although animals can quite easily be trained to react appropriately to commands, their symbolic productivity is often doubted. This should be examined. There exists, of course, a broad literature on “language” learning of apes (Harnad, Steklis, & Lancaster, 1976; Miles, 1990; Savage-Rumbaugh, 1994). But even without training, *mand naming* appears to be common. A real-life example might serve as a starting point: When my kitty wanted to be let out, he jumped on the arm of the couch, hit the safety-chain that dangled within reach of his paw, and looked back at me. If I did not respond expeditiously, he repeated this sequence. He was never explicitly trained in this, yet quite complex relationships exist between multiple stimuli: access to the outdoors as a reinforcing consequence, the door as an obstacle, the chain that can be manipulated, and the person whose behavior is the means to produce the consequence. Originally, the jingling chain and the opened door preceded the getting out. Later, getting out was the previously experienced consequence and jin-

---

Address correspondence to the author at the Department of Psychology, California State University, Fresno, Fresno, California 93740 (E-mail: ernstm@csufresno.edu).

gling was employed as a mand name; this performance suggests symmetry, because the cat had not previously produced the jingling. What is the behavioral difference between this performance and our ringing the doorbell when visiting?

Stories of dogs bringing the leash to the master when they want to go on a walk are commonplace. There are reports (Cheney & Seyfarth, 1980) that wild vervet monkeys produce different warning signals for different predators. This performance would be another precursor to naming. Prevarication, a unique language feature strongly emphasized by Hockett (1960), was found in some of Köhler's (1951) apes. These apes also took a person's hand and guided him or her to perform a specific desired activity. Specialists in animal ecology could probably report much more intentional communication. As a developmental psychologist, I am therefore skeptical whether such a sharp distinction should be drawn between animals and the name-giving homo sapiens in their capacity to establish various degrees of equivalence relations. Vauclair (1990) discusses such continuities from a Piagetian perspective.

*Additional Research on Children's Semiotic Development*

Proceeding from animals to the neonate and very young infant, stimulus equivalences are also early established: The mother's voice, sight, and even her body odor are responded to as interchangeable in many respects, even if this might again be only functional equivalence. Piaget (1952) has described in detail how even slight noises in the room induced the infant to anticipate the appearance of the mother. Piaget has derived later symbolic behavior from these early equivalences. Multiple symmetrical relationships between sights and sounds in relation to feeding and other important events have been documented in careful observations.

Proceeding to message production, it can be shown how the infant's instinctual crying changes within a few weeks after birth to "calling," as indicated by changes in the sound spectra. Piaget (1952) has reported that preverbal infants have also been observed to take the hand of a parent to guide him or her to a thing that needs to be handled. This comes very close to demanding an

action, and it is acquired without any explicit training. More generally, social signaling needs to be considered here. Pointing serves both as a discriminative stimulus and as an operant response for children with specific expressive language delay. It probably derives from instinctive reaching that precedes verbal reference (Bates, Bretherton, & Snyder, 1988) but later becomes an act of communication through social consequences. Words can then readily replace the pointing (Whitehurst & Fischel, 1994), showing that the equivalence principle has been operative pre-linguistically.

Closer in specific semantic emphasis is Nelson's research program on the acquisition of meaning, culminating in her book, *Making Sense: The Acquisition of Shared Meaning* (1985). Moerk (1986), especially in the section on "The Environmental/Instructional Impact," has provided a more behavioral evaluation of her work. Clark, who spent her entire professional life studying the development of meaning, has summarized her conclusions in *The Lexicon in Acquisition* (1993). Her name, too, is missing from Horne and Lowe's reference list. During 1996, informative discussions on early vocabulary have proceeded on the CHILDES network, which could not be included in Horne and Lowe's article, but these could be an important source of information for the reconceptualization of the argument. The 1995 *Handbook of Child Language* also contains a relevant chapter by Barret on "Early Lexical Development."

From an evolutionary and developmental perspective, the question should be raised as to how Sidman's stimulus equivalence as a biologically given function (p. 189) relates to the multimodal perception that is centered in the superior colliculus in the brain (Pöppel, Held, & Dowling, 1977) and is most functional during early development. Considering what was indicated above about the survival value of multimodal exploration, such phylogenetically and ontogenetically early roots would be expected. Moerk (1984) has related this evidence to symbolic development. Because multimodal perception seems to imply considerable symmetry and perhaps even transitivity in the relationships between stimuli, stimulus equivalence could have relatively

primitive neurological roots. Werner (1940) has explored related topics as “synesthesia.”

In past developmental discussions, a trichotomy of *index*, *icon*, and *symbol* has been generally employed. The index could probably be equated with the discriminative stimulus; responses to iconic aspects are already found in fearful reactions of apes to unknown percepts, as reported by Langer (1942). Signs are verbal labels. Why did Horne and Lowe neglect this widely accepted developmental trichotomy?

Two generalizations can be drawn from the observations presented so far. First, mand naming (in a broad sense) is functional. In accordance with this principle, many early vocal-verbal communications, such as *up*, *down*, *want*, are employed as operants. And certainly animals induce their masters, sometimes through manipulating acquired symbols, to provide particular reinforcers. Second, a large repertoire of “names,” that is, of specific signals related to ends (particular consequences), provides the potential for effective messages in a variety of situations. The establishment of a large number of words could first be a generalized operant to control the social environment and could thereafter become a tool of cognitive mastery through the classification of experiences, as Langer (1942) argued and Horne and Lowe (1996) implied.

*An Extension to Philosophical, Psycholinguistic, and Cognitive Approaches to “the Meaning of Meaning”*

The title chosen for the present comment was adapted from the famous book of Foucault (1970), *The Order of Things: An Archeology of the Human Sciences*. The book’s original title was *Les mots et les choses (Words and Things)*, and it surveys conceptions of meaning from the Renaissance to the present. My title suggests that pre-Skinnerian writings, such as *The Meaning of Meaning* (Ogden & Richards, 1923) might be worth digging up when dealing with this complex topic. A central omission is the work of Frege (1848–1925) and his pivotal distinction between reference and meaning (or sense), as succinctly summarized in Ogden and Richards (1923). A more extensive source for Frege’s conceptualizations is the translation provided by Geach and Black (1952). This distinction has been ac-

cepted in most discussions on semantics of the last hundred years (cf. Whitehead & Russell, 1910–1913). Frege, a mathematician, employed a mathematical perspective in his semantic analyses (Geach & Black, 1952). His concept of *identity* is close to the equivalence relation. A system of stimulus equivalences, such as that presented by Sidman and Tailby (1982), relying on the mathematical principle of equivalence, would therefore seem to have special affinity to Frege’s ideas. The argument against “sameness” made by Horne and Lowe (p. 234ff) has been discussed since the Renaissance under the notion of “resemblance” (Foucault, 1970, p. 67ff). What the authors refer to on page 189 as the relationships between speaker, listener, and object in Skinner’s (1957) theory was analyzed much more extensively, and profoundly, by Karl Bühler (1934/1965) in his “organon model” of language. A triadic model is common in recent history as it was in the writings of Peirce (1839–1914; as surveyed in Murphey, 1961), in those of Frege (1848–1925), in Ogden and Richards (1923), and in many subsequent discussions. It originated with the Stoics. Pre-Skinnerian work with special affinities to behaviorism might have been brought to bear, notably the extensive work by Charles Morris (e.g., 1946) on semantics and Mowrer’s (1960) theories of language origins. I am surprised that these are reflected neither in Horne and Lowe’s lead article nor in the commentaries.

Without such complex epistemological grounding, a serious category error might have been committed by Horne and Lowe (1996) when they reject the match-to-sample paradigm in favor of a naming explanation. With this shift, the explanandum, that is, the symbolic capacity for naming, has become the explanans under the subterfuge of “echoic behavior” and “listener behavior,” seemingly without the authors noticing it. However, if the symbolic capacity (naming) is based on simple principles, then many species should be trainable in language skills, or should have developed them spontaneously, because naming is so functional in shared endeavors. As this does not seem to be the case, the learnability of naming needs to be explained, not just presumed.

Whereas the above remarks are intended as constructive criticisms, they imply also an

acknowledgment of the courage of Horne and Lowe, who have tackled a topic which the best minds of Western civilization have struggled with at least since Plato's *Kratylos*. Other civilizations have done the same, as briefly discussed in Ogden and Richards (1923). Equally, the wisdom of the editors is to be applauded; they knew how many minds might be needed to contribute building blocks to an empirical and process-based theory of reference and meaning. My contributions are intended as a medley of potential constituents to the challenging endeavor of Horne and Lowe.

## REFERENCES

- Barret, M. (1995). Early lexical development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 362–392). Cambridge, MA: Basil Blackwell.
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar*. Cambridge, MA: Cambridge University Press.
- Bühler, K. (1965). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart: Fischer. (Original work published 1934)
- Cheney, D. L., & Seyfarth, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behavior*, 28, 362–367.
- Clark, E. (1993). *The lexicon in acquisition*. Cambridge, MA: Cambridge University Press.
- Foucault, M. (1970). *The order of things. An archeology of the human sciences*. New York: Pantheon Books. (Originally published as *Les mots et les choses*. Paris: Gallimard, 1966)
- Geach, P., & Black, M. (Eds.). (1952). *Translations from the philosophical writings of Gottlob Frege*. Oxford: Oxford University Press.
- Harnad, S. R., Steklis, H. D., & Lancaster, J. (Eds.). (1976). Origin and evolution of language and speech. *Annals of the New York Academy of Sciences*, 280. New York: New York Academy of Sciences.
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203, 1–10.
- 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.
- Köhler, W. (1951). *The mentality of apes*. New York: The Humanities Press.
- Langer, S. K. (1942). *Philosophy in a new key. A study of the symbolism of reason, rite, and art*. New York: The New American Library.
- Miles, H. K. W. (1990). The cognitive foundations for reference in a signing orangutan. In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes. Comparative developmental perspectives* (pp. 511–539). Cambridge, MA: Cambridge University Press.
- Moerk, E. L. (1984). Object, percept, concept, word. In L. Spillman & B. R. Wooten (Eds.), *Sensory experience, adaptation, and perception* (pp. 143–167). Hillsdale, NJ: Erlbaum.
- Moerk, E. L. (1986). The meaning of meaning: A review of *Making Sense: The Acquisition of Shared Meaning. Developmental Review*, 6, 365–385.
- Morris, C. (1946). *Signs, language, and behavior*. New York: Prentice Hall.
- Mowrer, O. H. (1960). *Learning theory and the symbolic processes*. New York: Wiley.
- Murphey, M. G. (1961). *The development of Peirce's philosophy*. Cambridge, MA: Cambridge University Press.
- Nelson, K. (1985). *Making sense: The acquisition of shared meaning*. Orlando, FL: Academic Press.
- Ogden, C. K., & Richards, I. A. (1923). *The meaning of meaning*. New York: Harcourt, Brace & World.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: Norton.
- Pöppel, E., Held, R., & Dowling, J.E. (1977). *Neural mechanisms in visual perception*. Cambridge, MA: MIT Press.
- Rubenstein, D. I., & Wrangham, R. W. (Eds.). (1986). *Ecological aspects of social evolution*. Princeton, NJ: Princeton University Press.
- Savage-Rumbaugh, E. S. (1994). *Kanzi: The ape at the brink of the human mind*. New York: Wiley.
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37, 5–22.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Vauclair, J. (1990). Primate cognition: From representation to language. In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes. Comparative developmental perspectives* (pp. 312–329). New York: Cambridge University Press.
- Werner, H. (1940). *Comparative psychology of mental development*. New York: International Universities Press.
- Whitehead, A. N., & Russel, B. (1910–1913). *Principia mathematica*. Cambridge, MA: Cambridge University Press.
- Whitehurst, G. J., & Fischel, J. E. (1994). Early developmental language delay: What if anything should the clinician do about it? *Journal of Child Psychology and Psychiatry*, 35, 613–648.

WHAT'S IN A NAME? EQUIVALENCE BY  
ANY OTHER NAME WOULD SMELL AS SWEET

RONALD J. SCHUSTERMAN, DAVID KASTAK,  
AND COLLEEN J. REICHMUTH

INSTITUTE OF MARINE SCIENCES,  
UNIVERSITY OF CALIFORNIA

Horne and Lowe (1996) argue that our results demonstrating California sea lion Rio's success on tests of equivalence relations (Schusterman & Kastak, 1993) were likely due to four procedural artifacts (p. 223). However, four lengthy commentaries directly addressed Horne and Lowe's attempt to devalue and explain away our positive findings, and all of them found the alternative explanations and arguments neither persuasive, compelling, parsimonious, nor even accurate (Fields, 1996, p. 280; McIlvane & Dube, 1996, p. 269; K. Saunders & Spradlin, 1996, p. 306; R. Saunders & Green, 1996; pp. 313–314). It is therefore unnecessary for us to further elaborate on either our procedure or the interpretation of our results. Instead, in this paper, we will concentrate our effort on two points made by Horne and Lowe. First, we will focus on their assumption that thought is dependent on words rather than the other way around. Second, we will address their point that one must be cautious about overinterpreting results from a single study (Horne & Lowe, 1996, p. 330) by reviewing the results of recent work at our laboratory as well as other important studies concerning the cognitive capabilities of nonhuman animals.

We believe that Horne and Lowe's basic premise that naming or verbal reasoning is necessary for the emergence or the formation of equivalence classes in an individual is wrong. In our view, the way some nonhuman animals and preverbal children classify the relations between and within objects and events determines whether and how codes, symbols, or words are acquired and used and not vice versa. Nonlinguistic animals have repeatedly

shown that they are capable of reasoning and conceptualizing about relations dealing with time, space, and objects (for reviews, see Cheney & Seyfarth, 1990; Mackintosh, 1994; Schusterman, Thomas, & Wood, 1986; Vauclair, 1996; Wasserman, 1993). Animals, along with nonlinguistic human babies and adults, have been shown to display many abstract types of thinking (for babies, see Wynn, 1992; for adults, see Schaller, 1991; Shepard & Cooper, 1982).

*Thinking Without Words*

Terrace (1993) would say that Horne and Lowe (1996) have bought into the Cartesian and non-Darwinian belief that the word is the sole sign of thought. However, many investigators believe that there is currently enough empirical data to support the hypothesis that words depend on thinking, or what Pinker (1994) calls "mentalese" and what Terrace (1993) calls nonverbal thinking, and that nonlinguistic cognitive processes may draw upon the same cognitive mechanisms that are used in the processing of language. Indeed, Sidman (1994) thinks that the emergence of equivalence relations does not depend on naming but that some aspects of language like naming do depend upon the emergence of equivalence relations.

The idea that nonlinguistic beings are capable of forming equivalence classes has been our working hypothesis for the past several years (Gisiner & Schusterman, 1992; Kastak & Schusterman, 1994; Schusterman & Gisiner, 1997; Schusterman & Kastak, 1993). In the remainder of this paper we will try to bolster this argument by first describing some additional experiments with California sea lions that support the hypothesis that language and naming are unnecessary for the emergence of equivalence classes. These experiments will show that California sea lion Rio can form equivalence classes in a conditional

The writing of this article was supported by ONR Grant N00014-56-J-0011 to the first author.

Address correspondence to Ronald J. Schusterman, Joseph M. Long Marine Laboratory, 100 Shaffer Road, Santa Cruz, California 95060 (E-mail: rjchust@cats.ucsc.edu).

discrimination and transfer them to a simple discrimination and can do the opposite (i.e., form equivalence in a simple discrimination and transfer the equivalencies to a conditional discrimination). Then we will discuss how the emergence of equivalence is likely to play an important role in the social and communicative interactions of social living animals like California sea lions and vervet monkeys.

#### *California Sea Lions: Additional Data*

As previously noted, studies showed that Rio, after being trained to relate Stimulus Pairs AB and BC (i.e., choosing Comparison B conditionally upon Sample A and choosing Comparison C conditionally upon Sample B), could then demonstrate emergent reflexive (AA, BB, and CC), symmetrical (BA and CB), transitive (AC), and equivalence (CA) relations among those visually presented stimuli (Kastak & Schusterman, 1994; Schusterman & Kastak, 1993). Two weeks after Rio successfully passed the equivalence relations tests, we switched from a conditional discrimination procedure to a two-choice simple discrimination. The simple discrimination experiment was conducted in order to test the transfer of equivalence relations formed in one context (matching to sample) to a second, novel context (simple discrimination). In terms of the simple discrimination procedure, a stimulus pairing was defined as members from two potential classes pitted against each other as alternate choices. No sample object was presented on these trials. The stimulus configurations for the simple discrimination were formed by pairing the A, B, and C members from each of two randomly chosen three-member equivalence classes that had been learned by Rio in the earlier experiment (Schusterman & Kastak, 1993). Rio's 30 equivalence classes were divided into a total of 15 class pairings, each consisting of three simple discriminations pitting A versus A, B versus B, and C versus C. On the first discrimination (training phase), one of the three possible member pairings was presented to the sea lion, and her choice on Trial 1 was reinforced, regardless of class membership. Following criterional learning of this discrimination (90% correct in a block of 10 consecutive trials), one of the two remaining stimulus pairings from the same two classes was tested (Transfer Test 1). The third pair-

ing from the same classes (Transfer Test 2) was introduced following 90% performance on the second pairing. This procedure was repeated for the remaining 14 class pairings. Our measure of transfer was Rio's performance on Trial 1 of each novel stimulus pairing, giving a total of 30 completely novel trials (15 trials were used in training to establish the class to be denoted as correct for each pairing).

As an example, a potential class pairing might be between Equivalence Classes 1 and 2. In the training phase, the B member from Class 1 (B1) might be pitted against the B member from Class 2 (B2). Rio's choice on Trial 1 then established the reinforcement contingencies for the remainder of the training phase. For example, a choice of B1 on Trial 1 established B1 as the positive discriminative stimulus and B2 as the negative discriminative stimulus. Subsequently, responses to A1 and not A2 would be reinforced on Transfer Test 1 and responses to C1 and not C2 would be reinforced on Transfer Test 2. Stimulus pairings in Transfer Tests 1 and 2 would be completely novel on their first presentation to Rio, in that these class members had previously been paired only in conditional discriminations and never in a simple discrimination. Performance on the first trial of each transfer test constituted the critical measure of the equivalence responding in this experiment.

On the first trial of each of the 30 transfer tests, Rio made 28 correct (i.e., class consistent) responses; such performance is significantly better than that expected by chance (two-tailed binomial test,  $p < .01$ ). These results indicate the sea lion's ability to transfer equivalence relations learned in matching to sample to simple discriminations, a completely novel context.

In the next series of experiments, we attempted to determine whether a modification of the discrimination reversal technique first used successfully to generate two functionally equivalent classes in pigeons (Vaughan, 1988) could do the same for California sea lions. Moreover, we went on to test whether Rio would show conditional discriminations emerging between members of a functional class. Finally, we tested whether the sea lion, after being taught to relate novel stimuli to existing class members, could then

match other class members to the novel stimuli, thus demonstrating stimulus equivalence. This series of experiments was modeled after those designed by Sidman, Wynne, Maguire, and Barnes (1989) to examine equivalence class formation by human subjects.

The subjects of the first stage of testing were 2 female California sea lions, Rio and Rocky. Like Rio, Rocky had extensive experience with other learning tasks, including conditional discrimination learning. For this task, we used 20 different visual stimuli that were divided into two sets of 10. All of the stimuli in the sets were discriminably different, and for our own convenience, we identified each object as either a letter or a number.

The general procedure was a two-choice simple discrimination. In this task, the subject was presented with two different stimuli, one from each set (one letter and one number), and selection of the object designated as correct by the experimenter was reinforced. The letter-number pairings were shuffled so that any member of the letter class and any member of the number class could appear together on a given trial. The sequence of trials in a session consisted of four sets of 10 trials. Each set included one presentation of each class member in a randomized order. The reversal procedure consisted of training a series of simple discriminations in which selection of members of a particular set was reinforced until the subject met a preset criterion (90% correct responses in one session). Subsequently, the contingencies were reversed; choices of members of the previously negative set were reinforced until the subject met the same criterion. This series of reversals from the letter set to the number set continued throughout the experiment.

Like Vaughan (1988), we defined functional classes as groups of stimuli that become interrelated through common behavioral functions, such that when responses to some members of a class change, responses to the remaining members of that class change accordingly. Based on this definition, performance on specific trials following a reversal in reinforcement contingencies may be influenced by the outcome of previous trials with other class members. Both of our sea lions eventually showed evidence of functional class formation in this task, reversing their

choices from members of the letter class to the number class (and vice versa) after several trials with specific class members. Following a reversal, performance on the first exposure of the 10 stimuli from each class should be close to zero, because the last time the sea lions encountered each stimulus, it appeared as an incorrect alternative. However, Rio and Rocky scored an average of 3.4 and 3.1 correct responses out of 10, respectively, on the first presentation of each stimulus following a reversal. This effect was enhanced by the introduction of differential outcomes to each class, such that correct responses to members of the letter class and correct responses to members of the number class were differentially reinforced with specific types of fish. Under this condition, Rio and Rocky scored an average of 6.5 and 6.0 correct, respectively, on the first presentation of each of the 10 stimuli in a class following a reversal in reinforcement contingencies. Further, performance on Trial 1 for the last five class members following a reversal was significantly better than chance, averaging 4.0 of 5 for Rio, and 3.5 of 5 for Rocky. This indicates rapid reversal of choices following exposure to a few exemplars from each class.

Following the formation of functional classes in the first phase of testing, we used a two-choice conditional discrimination procedure to assess whether members of a particular class would be related in a novel context. Specifically, the task was designed to assess whether a sea lion trained to sort stimuli into classes based on contingencies of reinforcement could match stimuli from these functional classes to one another in conditional discriminations on the basis of their class memberships. Rio was the subject in this experiment.

A matching-to-sample procedure was used to test the transfer of functional classes to conditional discriminations. The experiment consisted of presenting one stimulus as a sample object and then two comparison objects, one from each set, as choice stimuli. Correct responses consisted of matching the sample stimulus to the choice object that belonged to the same class. If a number was the sample, then picking a number as the correct match was reinforced, and conversely, if a letter was the sample, choosing a letter as the correct match was reinforced. Correct responses were differentially reinforced, as in the latter

half of the previous test. Performance on each novel stimulus pairing from the two classes was measured to assess transfer. This testing procedure generated a total of 180 novel trial combinations. The presentation sequence of these trials during testing was randomized. Rio scored 154 out of 180 possible correct responses on novel trials in this transfer test. Her performance was significantly better than expected by chance (two-tailed binomial test,  $p < .001$ ).

The final testing phase in this experiment was conducted to determine whether functional classes could also form equivalence classes. Specifically, we tested whether Rio, after being taught to relate novel stimuli to existing class members, could then match other class members to the novel stimuli. The experiment consisted of a training phase and a testing phase, both utilizing a matching-to-sample procedure. The training phase consisted of training four new conditional discriminations in which two new stimuli (K and 11) were mapped onto two existing class members (J and 10). The four conditional discriminations trained were JK, KJ, 10-11, and 11-10. Following training of the new conditional discriminations, the new stimuli, K and 11, were tested against the remaining members of each functional class to determine whether they had become related through equivalence relations. If K appeared as the sample, Rio's selection of another letter (not a number) was reinforced. Conversely, if a letter appeared as the sample stimulus, Rio's choice of K (not 11) was reinforced. The trials testing the transfer of 11 into the number class were set up the same way, and all test trials appeared in a randomized order. Correct matches of stimuli in either class were differentially reinforced, as in the earlier two phases of this experiment. The procedure generated a total of 36 novel trial combinations, and performance on these relations was measured by performance on Trial 1. Rio immediately and accurately matched the new stimuli to each of the remaining functional class members, scoring 36 out of 36 possible correct responses on Trial 1. Her performance was significantly better than would be expected by chance (two-tailed binomial test,  $p < .001$ ).

In Sidman and Tailby's (1982) original formulation of stimulus equivalence, the learned

equivalence of dissimilar stimuli was based on the notion that equivalence relations or conditionally related stimuli (which have become interchangeable in their control over behavior) have characteristics similar to those delineated in mathematics: reflexivity, symmetry, and transitivity. Currently, Sidman (1994) states that equivalence can also be the product of simple discriminations (three-term contingencies) and that responses and reinforcers, as well as discriminative and conditional stimuli, can be members of equivalence classes. The results from the present two experiments, showing that equivalence relations formed by a sea lion under one set of conditions can be transferred to a novel set of conditions, are supportive of the expanded view of stimulus equivalence taken by Sidman. We believe that our results support the notion that behaviorally as well as mathematically, functional classes imply equivalence relations and vice versa (Sidman, 1994, pp. 418–419). All of our work with California sea lions suggests that some members of the species are capable of forming equivalence relations and can do so under a variety of contexts in the absence of any symbolic or "naming" behavior.

#### *Individual Recognition and Cross-Modal Transfer*

*Pinnipeds.* Observations on the natural behavior of California sea lions indicate that equivalence classes may play a role in their social and communicative interactions. We suggest that the evolutionary learning program that promotes imprinting by California sea lion pups on their mother's voice (see Schusterman, Hanggi, & Gisiner, 1992; Trillmich, 1981), her smell, and various aspects of her visual appearance, such as her gait, posture, and facial expression, is likely to facilitate the pup's ability to interrelate these cues with one another and with the powerful reinforcers of milk, suckling, protection, and warmth that it receives from its mother. Thus, individual recognition beginning with mother-pup recognition via a variety of sensory modalities may be a critical developmental requirement in group-living animals such as California sea lions (Schusterman, Kastak, & Reichmuth, 1995). Perhaps the same skills that enable the pup to react differentially to its mother's odor and visual appearance when

it has only heard her voice may later enable the mature sea lion to recognize its sisters through their relationship with the mother and each other. There is evidence that like several other group-living animals, California sea lions discriminate particular kin on the basis of any of several sensory modalities (Hanggi & Schusterman, 1990). Therefore, it is reasonable to suggest that such perceptual learning abilities are related to their skill at successfully passing tests of stimulus equivalence. In contrast, harbor seals are notoriously poor at performing in matching-to-sample tasks (Constantine, 1981; Hanggi & Schusterman, 1995), and harbor seal mothers, during their brief (1-month) period of attending to the pup, do not emit pup attraction calls as do California sea lion females, who attend to their pups for 6 months or longer.

*Vervet monkeys.* Vervet monkeys have been studied by Cheney and Seyfarth (1990) in Amboseli National Park. They live in stable social groups consisting of a number of adult males, adult females, and their juvenile and infant offspring. These field investigators have pointed out that these animals, like several other social mammals and birds, recognize one another individually, have extended families, and form alliances against one another, like the Montagues and the Capulets. By doing an extensive series of playback studies, Cheney and Seyfarth have demonstrated that vervets place physically dissimilar stimuli consisting of other group members, predators, acoustic signals, responses, and a host of social reinforcers into equivalence classes. For example, when for the first time free-ranging adult female vervet monkeys are played the scream of an absent juvenile from a concealed loudspeaker, the adult females frequently respond to the playback scream by looking at the juvenile's mother. This occurred before the mother looked toward or approached the speaker herself. One could say that adult females, through previous experience, related the sound of the juvenile's scream (A) with the juvenile itself (B) and related the juvenile (B) with its mother (C) and therefore, the first time they heard the scream of the absent juvenile vervet monkey, without any additional experience they immediately oriented to its mother. Thus, responsiveness to the mother by the other adult females in the group emerged in this novel

context, even though these female monkeys may have never explicitly had the experience before.

An equivalence learning model might argue that the existing relation between the scream (A), the juvenile itself (B), and the frequent relations between the infant (B) and its mother (C) resulted in a three-member equivalence class consisting of the scream, the juvenile, and the mother (A, B, and C). That is, under certain circumstances, these events and individuals may all be classified as the same.

Learned equivalence of stimuli can be readily applied to affiliative and aggressive behavior of group-living individuals (e.g., Hanggi & Schusterman, 1990). Following an aggressive interaction between 2 vervet monkeys from two different genetic lineages (A1 and A2), other individuals from those lineages (B1 and B2), having observed A1 and A2 fight, are more likely to fight even though they were not involved in the original conflict (Cheney & Seyfarth, 1990). The aggressive events in this description are interchangeable with one another not because the individuals necessarily resemble one another but because the individuals in each of the lineages have previously shared a common functional association in terms of temporal or spatial proximity and positive and negative reinforcers. In general, kinship- and friendship-based coalitions seem to depend on a history of common functional relations that establish the equivalence of stimuli.

In referential communication studies with vervet monkeys, Cheney and Seyfarth (1990) have used playback-habituation experiments to demonstrate that the subjects classify the following calls as functionally equivalent despite their different acoustic properties: (a) vervet eagle alarm calls and starling raptor alarm calls and (b) calls labeled "wrrs" and "chutters" used to coalesce group members to the approach of another group. Thompson (1995) has recently pointed out that results like these, in which vervet monkeys classify calls on the basis of their common referent, are similar to the way pigeons use a prior association with a common response to produce a class of functionally equivalent but physically different visual stimuli (Wasserman & DeVolder, 1993).

In addition, it has been shown in matching-

to-sample tests that Java monkeys classify individuals into affiliative pairs (Dasser, 1988). Finally, it has been pointed out that nonhuman primates appear to switch from one behavioral social reinforcer to another (e.g., exchanging a mount for tolerance at a food source or grooming for later support in an alliance; Cheney, Seyfarth, & Smuts, 1986). These observations strongly suggest that some species of nonhuman primates, and perhaps other animals living in social groups, are capable of forming an equivalence class consisting of social reinforcers.

### Conclusion

In conclusion, we believe that field studies of social nonlinguistic animals provide strong evidence for the formation of equivalence classes without the prerequisite linguistic ability suggested as necessary by Horne and Lowe (1996). Further, we have expanded upon previous experimental evidence for the formation of equivalence relations in California sea lions, and believe that, although the requisite tests for reflexivity, symmetry, and transitivity have not been conducted successfully on other nonlinguistic animals, many such animals are more than likely capable of forming equivalence classes. The most parsimonious explanation for the appearance of equivalence in both humans and other animals is that the ability evolved in a social or ecological context, rather than as a result of linguistic competence. Thus, we believe that naming is not a prerequisite for equivalence, but equivalence is likely to be essential for naming.

### REFERENCES

- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.
- Cheney, D. L., Seyfarth, R. M., & Smuts, B. B. (1986). Social relationships and social cognition in nonhuman primates. *Science*, *234*, 1361–1366.
- Constantine, B. J. (1981). *An experimental analysis of stimulus control in simple conditional discriminations: A methodological study*. Unpublished doctoral dissertation, Northeastern University, Boston.
- Dasser, V. (1988). A social concept in Java monkeys. *Animal Behaviour*, *36*, 225–230.
- Fields, L. (1996). The evidence for naming as a cause or facilitator of equivalence class formation. *Journal of the Experimental Analysis of Behavior*, *65*, 279–282.
- Gisiner, R., & Schusterman, R. J. (1992). Sequence, syntax and semantics: Responses of a language-trained sea lion (*Zalophus californianus*) to novel sign combinations. *Journal of Comparative Psychology*, *106*, 78–91.
- Hanggi, E. B., & Schusterman, R. J. (1990). Kin recognition in captive California sea lions (*Zalophus californianus*). *Journal of Comparative Psychology*, *104*, 368–372.
- Hanggi, E. B., & Schusterman, R. J. (1995). Conditional discrimination learning in a male harbor seal (*Phoca vitulina*). In R. A. Kastelein, J. A. Thomas, & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 543–559). Woerden, The Netherlands: De Spil Publishers.
- 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.
- Kastak, D., & Schusterman, R. J. (1994). Transfer of visual identity matching-to-sample in two California sea lions (*Zalophus californianus*). *Animal Learning & Behavior*, *22*, 427–435.
- Mackintosh, N. J. (1994). *Animal learning and cognition*. San Diego: Academic Press.
- McIlvane, W. J., & Dube, W. V. (1996). Naming as a facilitator of discrimination. *Journal of the Experimental Analysis of Behavior*, *65*, 267–272.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow.
- Saunders, K. J., & Spradlin, J. E. (1996). Naming and equivalence relations. *Journal of the Experimental Analysis of Behavior*, *65*, 304–308.
- Saunders, R. R., & Green, G. (1996). Naming is not (necessary for) stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, *65*, 312–314.
- Schaller, S. (1991). *A man without words*. New York: Summit Books.
- Schusterman, R. J., & Gisiner, R. (1997). Pinnipeds, porpoises and parsimony: Animal language research viewed from a bottom-up perspective. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), *Anthropomorphism, anecdotes, and animals* (pp. 370–382). Albany: SUNY Press.
- Schusterman, R. J., Hanggi, E. B., & Gisiner, R. (1992). Acoustic signaling in mother-pup reunions, interspecies bonding, and affiliation by kinship in California sea lions (*Zalophus californianus*). In J. A. Thomas, R. A. Kastelein, & A. Y. Supin (Eds.), *Marine mammal sensory systems* (pp. 533–551). New York: Plenum Press.
- Schusterman, R. J., & Kastak, D. A. (1993). A California sea lion (*Zalophus californianus*) is capable of forming equivalence relations. *The Psychological Record*, *43*, 823–839.
- Schusterman, R. J., Kastak, D. A., & Reichmuth, C. (1995). Equivalence class formation and cross-modal transfer: Testing marine mammals. In R. A. Kastelein, J. A. Thomas, & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 579–584). Woerden, The Netherlands: De Spil Publishers.
- Schusterman, R. J., Thomas, J. A., & Wood, F. G. (Eds.). (1986). *Dolphin cognition and behavior: A comparative approach*. Hillsdale, NJ: Erlbaum.
- Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- Sidman, M., & Tailby, W. (1982). Conditional discrimi-

- nation vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37, 5–22.
- Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior*, 52, 261–274.
- Terrace, H. S. (1993). The phylogeny and ontogeny of serial memory—list learning by pigeons and monkeys. *Psychological Science*, 4, 162–169.
- Thompson, R. K. R. (1995). Natural and relational concepts in animals. In H. Roitblat & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 175–224). Cambridge, MA: MIT Press.
- Trillmich, F. (1981). Mutual mother-pup recognition in Galapagos fur seals and sea lions: Cues used and functional significance. *Behavior*, 78, 21–42.
- Vauclair, J. (1996). *Animal cognition: An introduction to modern comparative psychology*. Cambridge, MA: Harvard University Press.
- Vaughan, W., Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 36–42.
- Wasserman, E. A. (1993). Comparative cognition: Beginning the second century of the study of animal intelligence. *Psychological Bulletin* 113, 211–228.
- Wasserman, E. A., & DeVolder, C. L. (1993). Similarity and nonsimilarity-based conceptualization in children and pigeons. *The Psychological Record*, 43, 779–794.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749–750.

---

### EQUIVALENCE RELATIONS

MURRAY SIDMAN

NEW ENGLAND CENTER FOR CHILDREN

---

When I was asked to reply to Horne and Lowe's (1996) criticisms of my position on equivalence relations, I replied that I did not hold to a *position*. In my recent book, *Equivalence Relations and Behavior: A Research Story* (Sidman, 1994; in the present article, subsequent references to chapters or to pages are citations of that book), I made many suggestions about how to view phenomena that are subsumed under *equivalence relations*. In each instance, I first detailed my reasons for making the suggestion. Then, I outlined experiments that might either support or fail to support the suggestion. If I have any position, it is that data rather than debate will show the way.

With one exception (Rumbaugh, 1995), most of the more important proposals in the book have received little theoretical and no empirical commentary, either from critics or from those who might be favorably

inclined. I suggested, therefore, that I might simply reproduce selected paragraphs from my book. Somewhat to my surprise, this suggestion was received favorably. So here are some isolated paragraphs, repeated. Abstracting them from the general clutter will perhaps make them stand out more effectively.

Still, the surrounding material, although not included here, performs important functions, describing both the origins and possible consequences of each suggestion. Some proposals will not stand the test of data; for those, the originating problems will still remain. And so, I hope that anyone who really wishes to evaluate the following paragraphs will also attend to the context.

### EQUIVALENCE AND THE REINFORCEMENT CONTINGENCY

Page 325

The study of equivalence relations has contributed some new data to behavior analysis and perhaps some new principles, but none

---

Correspondence concerning this article should be addressed to Murray Sidman, 242 Beacon Street, Boston, Massachusetts 02116-1232 (E-mail: msidman@aol.com).

of these requires the abandonment of data and principles that have already proven their worth. The enlargement of the analytic unit that is outlined in the *Emergent Verbal Classes* paper [Sidman, 1986] is just that—an *enlargement*. It encompasses equivalence relations and their contextual control but requires no fundamental change in the methods of analysis or in the underlying empirical and theoretical structure. In fact, I am convinced that the *Emergent Verbal Classes* paper provides a useful framework within which to organize the existing data and principles of behavior analysis.

*Pages 324–325*

It is true that laboratory research on equivalence relations was examining phenomena that behavior analysis had not previously considered. Our formulation of those phenomena had introduced a new set of terms to the behavior analytic vocabulary. To some, our introduction of new terms and concepts seemed to suggest that that we were discarding the old ones. In spite of the novelty of topic and terminology, however, I never felt that we were abandoning the system of behavioral analysis that was founded on the experimental, theoretical, and philosophical contributions of B. F. Skinner. Rather, I viewed the work on equivalence relations as a natural extension of that tradition. A major aim of the *Emergent Verbal Classes* paper was to show that the equivalence relation, while perhaps a new behavioral concept as we had defined it, was an outgrowth of the same kind of contingency analysis that had given rise to basic relational concepts like stimulus, response, reinforcement, discrimination, conditioned and generalized reinforcement, and conditional discrimination.

*Page 367*

In [the *Emergent Verbal Classes* paper], I described the equivalence relation as emerging at the level of the four-term contingency. I now believe that this restriction of equivalence to the four-term unit placed too strong a constraint on the relation between equivalence and the units of behavioral analysis. Also, in the *Where Does Equivalence Come From?* paper [Sidman, 1990], I suggested that we have to consider seriously the possibility that equivalence is a basic stimulus function, not

derivable from more fundamental processes. I now believe more strongly in this possibility. These developments in my conception of equivalence—the weakening of one belief and the strengthening of the other—are related, but definitive evidence is not yet in.

*Pages 387–388*

*A terminological note.* It must be recognized that to speak of the “establishment” of equivalence relations is a circumlocution . . . it saves words and eliminates awkward sentence constructions. *Equivalence relation* refers neither to a theoretical entity nor . . . to processes or entities that are beyond observation, but rather, summarizes a set of observed regularities. Strictly speaking, reinforcement contingencies do not *create* equivalence relations; rather, they create prerequisites, or the potential, for demonstrating the properties that define an equivalence relation. Additional factors, like the test conditions, contextual control, and a subject’s behavioral history will help determine whether and how that potential is realized. . . .

An equivalence relation, therefore, has no existence as a *thing*; it is not actually *established, formed, or created*. It does not *exist*, either in theory or in reality. It is defined by the emergence of new—and predictable—analytic units of behavior from previously demonstrated units. . . .

The equivalence relation is not itself a unit of behavior from which more complex units are built. Nor is the equivalence relation a structure that is composed of more basic units. Although the diagrams that are typically used to depict experimental procedures may give a misleading impression of sequential or mediated learning processes, the definition of an equivalence relation does not require the component pairs to possess any temporal or structural property that might define a *mediating* event, a temporal or spatial *sequence*, an *association*, a *link*, an *associative link*, a *distance*, a *chain*, a *network*, a *conditioned stimulus*, a *conditioned response*, or any other kind of presumed basic structure or unit of behavior.

*Page 415*

In chapter 10, I pointed out that equivalence relations have been shown to include all possible ordered pairs of the stimuli in a

four-term analytic unit (conditional discrimination)—the conditional, discriminative, and reinforcing stimuli. Then, theoretical considerations along with some obtained and some anticipated empirical findings were advanced to support the inclusion of the unit's defined responses, too, among the components of the ordered pairs that make up an equivalence relation. These findings, actual and speculative, gave rise to the proposal that the contingency responsible for establishing the analytic unit is also responsible for the equivalence relation. . . .

Both of these proposals, (a) that defined responses be included as components of the equivalence relation's event pairs and (b) that the reinforcement contingency creates the equivalence relation, were said to be supported by findings that three-term contingencies . . . and perhaps even more restricted contingencies (response-reinforcer and stimulus-reinforcer) could also establish equivalence relations. If the four-term units that are needed for direct documentation of the . . . properties that define an equivalence relation can emerge from three-term (and perhaps two-term) contingencies that specify different reinforcers or defined responses, then, it was argued, reinforcers *and* defined responses would have to be included among the components of the relation. Only then could the emergence of equivalence from three- and perhaps two-term units be accounted for.

*Pages 378–379*

I am not calling for the inclusion of unobservable or invented responses in the equivalence relation. This is not just a repetition of mediation theory's practice of postulating the occurrence of responses in order to satisfy theoretical needs. Defined responses are neither invented nor inferred. They are . . . specified components of the reinforcement contingency. The contingency decrees that reinforcement be withheld unless the subject . . . emits the defined responses. . . .

*The three-term unit and the definition of equivalence.* The inclusion of defined responses as elements of event pairs that make up the equivalence relation turns out to be more than just an arbitrary assignment of labels. . . . One theoretically significant feature arises from the demonstration that discriminative stimuli can become related by equivalence

even when they are involved only in three-term contingencies. This demonstration [Sidman, Wynne, Maguire, & Barnes, 1989] . . . calls into question our original behavioral definition of the equivalence relation. . . . In the face of demonstrations that three-term contingencies can generate equivalence . . . I was compelled to drop the notion that equivalence emerges only at the level of the four-term unit.

Because the direct evaluation of reflexivity, symmetry, and transitivity requires four-term units, abandonment of the four-term unit as the necessary origin of equivalence might cause one also to abandon the set-theory definition of the equivalence relation. This would be unfortunate. . . .

It turns out that the inclusion of responses in the equivalence relation not only permits but forces us to maintain our set-theory definition. This is because even though four-term units need not be involved in generating equivalence relations, it is still necessary to take the defining properties into account if we are to understand in principle how an inferred equivalence relation could have arisen from three-term units.

*Page 380*

In addition to making emergent conditional discriminations predictable [from smaller units], the inclusion of differential responses in the equivalence relation . . . permits us to escape the theoretical intricacies in which we involve ourselves when we hypothesize response mediation as the process responsible for emergent stimulus-stimulus relations. A major complexity of mediation theory is the requirement that the mediating responses must occur, although perhaps in a reduced form, whenever a subject demonstrates either baseline or emergent stimulus-stimulus relations. To maintain the necessary linear chain of stimulus-response-stimulus and so on, mediation theorists had to assume the occurrence of unobserved responses between each related pair of stimuli. This assumption leads to greater and greater awkwardness in the necessary explanatory constructions as derived relations come to involve more and more baseline nodes. The very inelegance of mediation theories of stimulus equivalence ought to occasion some skepticism as to their explanatory utility.

*Page 381*

If . . . we simply include defined responses as elements of event pairs that constitute an equivalence relation, we need postulate no linear mediating process in baseline or emergent relations, no unobserved responses, and no backward conditioning to account for symmetry. . . . By definition, the equivalence relation will include all of the stimulus–stimulus, stimulus–response, response–stimulus, and response–response pairs that are directly taught and all of the pairs that emerge in the tests. We need nothing more than our behavioral definition of equivalence to predict the emergent relations.

*Pages 384–385*

What does the inclusion of responses in the equivalence relation have to say about the distinction between stimulus and response? It says that *with respect to the equivalence relation*, such a distinction is unnecessary [emphasis added]. An equivalence relation is made up of pairs of events, with no restriction on the nature of the events that make up the pairs. The locus of those events, whether it be the living organism or the organism's living or nonliving environment, is irrelevant.

This does not mean there are no differences between stimuli and responses. Like all events, stimuli and responses can be members of many classes. . . . A classification that distinguishes stimuli from responses on the basis of their locus is meaningful, particularly when behavior analysts are trying to define their subject matter. . . .

In the context of describing the membership of an equivalence class, the distinction between stimulus and response, or even between controlling and controlled events, loses its significance. Ordered pairs of the events that comprise an equivalence class constitute the membership of an equivalence relation. How the individual events that make up these ordered pairs are classified in other contexts has no bearing on whether the pairs belong to an equivalence relation. An equivalence relation may contain stimulus–stimulus, response–response, stimulus–response, and response–stimulus pairs.

*Page 386*

In analyzing equivalence relations, then, we do not sometimes call an event a stimulus and

at other times a response. Rather, we discard both of those terms. Equivalence relations have their own defining characteristics, none requiring the stimulus/response dichotomy.

The independence of equivalence relations from the arrows of time and causality removes any need to distinguish between stimuli and responses when specifying the membership of an equivalence class. . . . We need not conceptualize equivalence relations in terms that are relevant to conditioning.

*Page 387*

Still remaining is the question of where equivalence relations come from. An important part of the answer to this question almost leaps out at us now that we have included in the equivalence relation all the elements of the analytic unit. Question: Where does that unit come from? Answer: the reinforcement contingency creates the unit and with it, the equivalence relation. The establishment of equivalence relations is, then, one of the outcomes of reinforcement contingencies.

*Page 390*

Reinforcement contingencies select the particular elements that constitute a unit of analysis. As we have seen, the equivalence relation consists of ordered pairs of the unit's elements. The analytic units and the equivalence relations that reinforcement has established comprise an individual's repertoire of acts and discriminations. . . . The facts that analytic units and equivalence relations are established at all, however, are species characteristics. . . . the extent of the generality of [equivalence relations] across species is as yet unknown.

*Pages 553–554*

*The origin of equivalence relations.* Elementary mathematical set theory describes the abstract properties of equivalence relations, and I have argued that behavioral phenomena like those described in this book exemplify the mathematical abstraction. But although set theory informs us how to find out whether any particular event pair belongs to an equivalence relation, it is silent about the origin of equivalence relations themselves. Questions about testing for equivalence relations are to be distinguished from questions about prerequisites for the development of equiva-

lence. I have dealt with the latter problem by treating equivalence relations parsimoniously as a natural product of reinforcement contingencies. One consequence of this treatment is that all of the variables that modulate the effects of reinforcement contingencies can be expected to be found relevant also to equivalence relations, even though those variables may have no place in the mathematical description.

A second consequence is that no additional experience on the part of the individual need be invoked in order to account for the observation that the components of a reinforcement contingency are related by equivalence. My suggestion (chapter 10) was that equivalence relations are a built-in effect of reinforcement contingencies but that the breakdown of particular equivalence relations is a product of contextual control which, in turn, comes about through experience [Bush, Sidman, & de Rose, 1989]. . . . The problem then becomes not how to explain the origin of equivalence relations in general but how to explain those particular instances in which some or all of the events involved in a reinforcement contingency fail to become members of the same equivalence class. Experience is responsible for the removal or preclusion of events from an equivalence class.

Others, however, have argued that special kinds of experiences are necessary precursors of equivalence relations.

#### LANGUAGE AS A PREREQUISITE FOR EQUIVALENCE

*Pages 362–364*

Dugdale (1988) and Dugdale and Lowe (1990) have advanced the strongest arguments for the necessity of vocal or subvocal naming in the establishment of equivalence relations. Their studies demonstrated clearly that equivalence relations can be facilitated by naming or by some aspect of the experimental procedures or instructions that lead a subject to name the stimuli. This distinction between the effect of naming per se and the variables that lead to naming has not yet been addressed experimentally. At present, therefore, it is not clear that linguistic naming by

itself has been the critical factor in these studies. . . .

In discussing the relation between naming and equivalence, Dugdale and Lowe (1990) proposed a distinction between *naming* and *labeling*. They recommended that we use the term *naming* only when the relation between the name and the thing named is symmetric. For example, true naming is demonstrated when a child not only says “boy” upon seeing a boy but, having said (or heard) “boy,” then points to a boy.

I find myself sympathetic to this suggestion . . . [but] I believe that Dugdale and Lowe’s . . . definition of true naming is a component of a larger picture. In chapter 10, I will propose that differential responses in the analytic unit be included also in the equivalence relation. Such inclusion will require the relation between names and the stimuli that occasion them to be not only symmetric but reflexive and transitive as well. The equivalence relation will then include not only stimulus–stimulus pairs, but stimulus–response, response–stimulus, and perhaps even response–response pairs. Including differential responses in the equivalence relation will remove the necessity for the distinction that Dugdale and Lowe had to make between stimulus–response symmetry and stimulus–stimulus symmetry. This revised conception of the equivalence relation will also establish theoretical grounds for the facilitation of equivalence by differential responses (pp. 413–414).

*Pages 306–307*

In spite the interpretive difficulties that arise when naming tests are given after equivalence relations have been demonstrated, it would be imprudent to dismiss the naming data in Table 8-3 [from Sidman, Willson-Morris, & Kirk, 1986] too quickly. It is not obvious that all of those data can be attributed to subjects’ misinterpretations of what they were being asked to do, or to other methodological features that might have caused the subjects to give different names than those they had applied to the stimuli during the earlier conditional-discrimination tasks. . . . [data review] . . . Taken together with earlier observations cited in the *Role of Naming* paper [Sidman et al., 1986] . . . and with later replications by Green (1990), these data cannot

easily be declared irrelevant to the question of whether common names are necessary to mediate equivalence relations.

*Page 511*

To say that verbal mediation is unnecessary for equivalence is not to say that verbal labels and rules are always irrelevant. To deny what Luria [as cited in Vocate, 1987, p. 135] has termed “the abstracting and generalizing, analyzing and synthesizing power of language” . . . would be contrary to everyday observation. But how does language help us to abstract, to generalize, to analyze, and to synthesize, and how does it come to do so? The mere acknowledgment of those powers does not explicate the role of verbalization. . . . It is possible, for example, that attaching labels to stimuli in a single-node equivalence class and expressing rules for relating those labels may help one subsequently to expand the class in accord with multinodal contingencies. Why verbalization should have such a facilitating effect is an interesting and important problem but its interest and importance are not restricted to equivalence phenomena. Surely, that facilitation involves something more than just the establishment of links in a causal stimulus–response chain (Skinner, 1957, pp. 107–129).

*Pages 364–365*

*Generalized symmetry?* Having postulated that true naming, a symmetrical stimulus–response relation, is necessary for stimulus equivalence, Dugdale and Lowe (1990) went on to ask where naming comes from. They pointed out that symmetrical stimulus–name relations arise naturally in the course of a child’s language development, when the child is taught to be both a speaker and a listener—to say words and to comprehend those same words when others say them. Hayes, too (1991), has argued that such a history is necessary (although not sufficient) for equivalence relations. Dugdale and Lowe (1990) and Hayes (1991), therefore, attempt in this way to derive equivalence relations from an individual’s linguistic experience. I believe, however, that they have overlooked a significant assumption that underlies their derivation. They assume that with enough name–event and event–name examples (which ordinarily occur extensively in a

child’s natural language community), a generalized relation of symmetry will emerge naturally.

As Hayes (1991) pointed out, the concern here is with arbitrary relations. . . . I can understand how a sufficient number of examples may give rise to generalized nonarbitrary relations like *larger*, *brighter*, *heavier*, *more*, and so on. But I do not understand how any number of examples can give rise to generalized arbitrary relations like *reflexivity*, *symmetry*, *transitivity*, and so on. Because the exemplars would possess no measurable feature in common, it is not at all evident that one might be able to generalize an arbitrary relation solely from exemplars. What aspect of several examples of symmetric event–name relations would permit a new example to be recognized or produced?

*Symmetry* is a complex verbal construction, involving preestablished classes like *names*, *things*, *self*, *others*, and so on. The mere exposure of a verbally unsophisticated organism like a child or a nonhuman to a number of exemplars that have, themselves, not yet been appropriately classified does not seem to me sufficient to explain the emergence of a generalized concept of event–name symmetry on the basis of any known behavioral principle.

The key here is “known behavioral principle.” . . . Mere exposure to exemplars may yet prove sufficient to yield a generalized concept of symmetry. . . . But, if classes defined by such relational properties can, like nonarbitrary classes, be generated merely by presenting exemplars to nonverbal or verbally unsophisticated individuals, this will itself define a new behavioral process, not derivable from anything more basic.

In attempting to derive equivalence relations from an individual’s behavioral history, therefore, “exemplar theory” does not fulfill its intended purpose; it does not avoid the need to specify a behavioral process that is itself not derivable from anything more basic.

*Pages 556–557*

The accomplishments of mathematicians show us that linguistically proficient organisms can indeed abstract the properties of arbitrary relations and come up with a list of features that other similarly proficient organisms (behavior analysts?) can look for in any specific instance. . . . As I have asked before

(pp. 364–365), however, what makes it possible for linguistically unsophisticated organisms (like young children, people with severe mental retardation, or nonhumans) to abstract the shared features from a set of specific instances of reflexivity, symmetry, and transitivity when those very words—sophisticated abstractions that define the relation—remain outside of their repertoires? . . .

A linguistically naive organism's abstraction of commonalities from a set of exemplars that share no physical feature requires more of an explanation than just a history of experience with the exemplars. It is certainly possible to teach specific equivalence relations nonlinguistically, like *sameness*, for example, and to teach other kinds of arbitrary relations, too, like *opposition* and *difference*; all of these involve control by physical characteristics of stimuli. If, however, we were to find that linguistically impoverished organisms could derive the concept *equivalence relation* just from a reinforcement history with paired elements that shared no feature beyond the relation itself, that very finding would require an explanation that is not currently available among the principles of behavior analysis.

#### CONTEXTUAL CONTROL

*Pages 512–513*

Although the sources of equivalence have been a matter of theoretical dispute . . . the contextual control of equivalence relations has been generally agreed to have an experiential basis. I have gone so far as to suggest that experience may be required not to make equivalence possible but rather, to break down or prevent specific equivalence relations. . . . Instead of asking, "Where does equivalence come from?" I have found it useful to ask instead, "What breaks down or precludes an equivalence relation?" To answer this question in any particular instance, look for contingency-engendered contextual control.

Without experientially based contextual control, simple and conditional discriminations and equivalence classes . . . would be impossible; multiple class membership, giving rise to class union, would take events that we had to discriminate and bring them instead into one large equivalence class where they

would all be treated alike. Everyday observation tells us, however, that events can belong to more than one class even while those classes remain independent of each other. In such instances, what breaks down or prevents class union?

*Pages 523–524*

The experiments I have just described [pp. 515–523] . . . show me clearly that contextual control does not create equivalence relations but rather, that context prevents lower level contingencies from generating potentially maladaptive equivalence relations, and breaks down equivalence relations that other contingencies have already generated.

*Page 530*

That is the background of the suggestion that was advanced in the *Contextual Control* paper [Bush et al., 1989] for a resolution of the problem of "why the context itself does not become a member of all the emergent classes and, by virtue of its common membership, condense all of the classes into one" (p. 507). Because equivalence relations are not directly specified in a reinforcement contingency, it is possible for a conflict to exist between the two outcomes of a contingency: (a) the creation of an analytic unit and (b) the formation of an equivalence relation. In a five-term unit, for example, . . . the contingency calls for differential control by discriminative stimuli in each three-term unit and by conditional stimuli in each four-term unit; on the other hand, the contingency creates equivalence classes containing a common contextual element that could wipe out differential control by bringing all discriminative and conditional stimuli together into a single class. Our suggestion was that creation of the unit takes priority. The explicit inclusion of differential stimulus control in the contingency counteracts the formation of equivalence classes; the latter are not only not explicitly included in the contingency but would actually prevent the conditions that are included from being met. And so, behavioral processes determine which aspect of the mathematically derived description is applicable; in this instance, whether control by context brings about class union or class intersection.

## THE DESCRIPTIVE SYSTEM

Pages 536–537

My own theorizing has been directed not so much at an explanation of equivalence relations but rather, at the formulation of a descriptive system—a consistent, coherent, and parsimonious way of defining and talking about the observed phenomena. Mathematical set theory contains tools that allow me to meet all of these goals. . . . My colleagues and I therefore adapted set theory's definition of the equivalence relation, a definition that has a large number of regularities already built in. That is to say, the regularities themselves define the equivalence relation. Any relation that is to be called an equivalence relation *must* show those regularities. This necessity gives the descriptive system one of the flavors of an explanatory theory; it permits us to make predictions. The predictions, however, are already incorporated in the definition. . . . That those regularities have been so reliably confirmed continues to astonish me. . . .

In the course of writing the present story, I also found other components of mathematical set theory to be useful for the description of equivalence phenomena. In particular, the fundamental concepts of set union and set intersection permit us to include within the same descriptive system behavioral phenomena that had previously seemed to require the postulation of a separate process—*transfer* of function. . . . Different classes that possess members in common may merge into a single class—set union—or may remain independent—set intersection. Contextual components of the contingency determine whether set union or intersection takes place. None of this requires more than a description of the events that make up an observable reinforcement contingency.

Still, there is more to equivalence relations than mathematical set theory can describe. If equivalence relations are a product of reinforcement contingencies, all behavioral variables that are relevant to reinforcement contingencies must be relevant also to equivalence relations. Beyond this truism, some investigators have suggested that the mathematically derived description of equivalence relations is incomplete because new—previously unknown—behavioral variables or theoretical principles are involved. Any dis-

covery of new variables is, of course, an exciting event. Nevertheless, such discoveries may introduce interpretive complexities and require difficult conceptual changes. It is wise, therefore, to follow a conservative course and search carefully for alternative explanations before trying to work a new variable into an existing formulation. Several instances in the literature on equivalence are worth noting.

*Structural determinants: Directionality* (pp. 537–538).

*Structural determinants: Nodal distance* (pp. 538–549).

*Class size as a variable* (pp. 549–550).

Page 550

*The mathematics and the behavior.* When applied to the analysis of behavior, the mathematical theory of sets seems to agree closely with behavioral reality. That this correspondence exists is in itself remarkable. How is it that purely mathematical conceptions fit observed behavioral phenomena so well?

The same question, of course, has been asked in physics and other natural sciences.

Page 553

Whitehead's conception that pure mathematics is concerned with general abstractions from matters of fact (see above) is also relevant to the sometimes expressed opinion that the mathematically derived behavioral definition of equivalence relations which I and my colleagues have offered is just "Sidman equivalence." The implication is that the definition, if not capricious, is, at the least, arbitrary, with no stronger a priori justification than any other definition. Far from being arbitrary, however, the mathematical definition of the equivalence relation possesses tremendous generality: "Equivalence relations are found not only in every corner of mathematics, but in almost all the sciences" (Gellert, Küstner, Hellwich, & Kästner, 1977). To adopt the mathematical definition is to take the position that behavior is included among the many real-world specifics that the abstractions of mathematical set theory encompass. This position, although conceivably incorrect, is hardly arbitrary. Given the general empirical support for the mathematical formulation, the a priori denial of its relevance to behavior

is considerably more arbitrary than its acceptance.

#### REFERENCES

- Bush, K. M., Sidman, M., & de Rose, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior*, 51, 29–45.
- Dugdale, N. A. (1988). *The role of naming in stimulus equivalence: Differences between humans and animals*. Unpublished doctoral dissertation, University of Wales.
- Dugdale, N. A., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 115–138). Hove, England: Erlbaum.
- Gellert, W., Küstner, H., Hellwich, M., & Kästner, H. (1977). *The VNR concise encyclopedia of mathematics*. New York: Van Nostrand Reinhold.
- Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation*, 95, 260–270.
- Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Reno, NV: Context Press.
- 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.
- Rumbaugh, D. M. (1995). Emergence of relations and the essence of learning: A review of Sidman's *Equivalence Relations and Behavior: A Research Story*. *The Behavior Analyst*, 18, 367–375.
- Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Erlbaum.
- Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 93–114). Hillsdale, NJ: Erlbaum.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities*, 6, 1–19.
- Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior*, 52, 261–274.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Vocate, D. R. (1987). *The theory of A. R. Luria*. Hillsdale, NJ: Erlbaum.

---

#### STANDARD PRINCIPLES, NONSTANDARD DATA, AND UNSOLVED ISSUES

FRANÇOIS TONNEAU AND MICHEL B. C. SOKOLOWSKI

CHARLES DE GAULLE UNIVERSITY, FRANCE

---

In their impressive article, Horne and Lowe (1996) reevaluate the theoretical import of stimulus equivalence. Their critical analysis seems largely correct. At one place, however, Horne and Lowe attribute the failure of equivalence explanations to the “artificial” character of match-to-sample procedures (Horne & Lowe, p. 238). In our view,

the fundamental problem with the equivalence framework stems less from its reliance on artificial situations than from a lack of clear theoretical principles (cf. Harzem, 1995<sup>1</sup>). Were such principles available, the artificiality of the procedures involved would not constitute a significant obstacle; after all, most behavioral concepts, including the concept of reinforcement used by Horne and Lowe, have been derived from experimental

---

Address correspondence to François Tonneau, Centro de Estudios e Investigaciones en Comportamiento, Universidad de Guadalajara, Avenida 12 de Diciembre 204, Col. Chapalita, CP 45030, Guadalajara, Jalisco, México (E-mail: ftonneau@udgserv.cencar.udg.mx).

---

<sup>1</sup> Harzem, P. (1995, May). *Natural contingencies*. Paper presented at the annual meeting of the Association for Behavior Analysis, Washington, DC.

situations no less artificial than match to sample.

### *The Naming Account*

What processes, then, underlie *stimulus equivalence*? Horne and Lowe develop various possibilities, but not all of them seem equally important or relevant, and some are only partly specified. In the section entitled *Other Verbal Behavior and Stimulus Equivalence*, for example (Horne & Lowe, pp. 221–222), complex utterances such as “circle goes with the open triangle” are said to determine performance on equivalence tests. However, the utterances in question could not be emitted unless equivalence outcomes had already emerged; emitting “green means up” after up → green forward pairings, for example (Horne & Lowe, p. 221), assumes some form of symmetry in the first place because the names “up” and “green” appear in the utterance in a reversed order.

The core of the naming account can in fact be found in the section entitled *Intraverbal Naming and Stimulus Equivalence* (Horne & Lowe, pp. 218–221). A first mechanism explains [AB, AC]:BC effects, where AB and AC are trained relations and BC is the matching relation observed on testing. According to Horne and Lowe, AB and AC trials establish two verbal sequences, “ab” and “ac.” Over a mix of AB and AC trials, the name “a” tends to evoke “b” and “c” at once, which yields a new chain “b” → “c.” This chain mediates performance on BC testing: B → “b” → “c” → choose C. A second mechanism deals with symmetry or [AB]:BA effects. The AB trials establish the verbal sequence “ab.” Repeating this sequence subvocally, “ababab,” yields spontaneous “ba” reversals, which in turn establish the verbal chain “b” → “a.” This chain mediates performance on BA testing: B → “b” → “a” → choose A.

Thus, according to Horne and Lowe, it is the involvement of the subject’s internal speech that explains the discrepancies between traditional behavioral principles and human performance on equivalence tests. On the other hand, the behavior of nonverbal animals should generally conform to standard molecular Pavlovian and operant laws (Horne & Lowe, p. 223).

### *Implications of Nonhuman Data*

Numerous results, however, contradict this last aspect of Horne and Lowe’s position. Holland (1981), for example, exposed rats to AB pairings between a neutral stimulus (A) and food (B); in a second stage (AC), A was paired with a toxin (C). During testing, the rats displayed an “emergent” aversion to B, as if the food itself (B) had been paired with C. Urcuioli, Zentall, Jackson-Smith, and Steirn (1989, Experiment 2) obtained analogous results with matching to sample by pigeons; subjects exposed to BB, AB, and AC trials showed clear positive transfer on a BC matching task.

The point is not that such data demonstrate something called “stimulus equivalence,” but merely that they and the [AB, AC]:BC effects obtained in humans (Horne & Lowe, Figure 2, p. 188) raise similar problems for standard behavioral principles. Hence, these principles should be revised or else supplemented with internal mediation. Revising or abandoning standard principles would be incompatible with Horne and Lowe’s approach, which is after all almost entirely built on such principles, conceived as “behavioral primitives” (Horne & Lowe, p. 239); Horne and Lowe should therefore follow the second course, and this even for nonhuman subjects.

A mediational explanation of Holland’s (1981) data has in fact been proposed by Holland himself (1981, 1990). Events A, B, and C are supposed to evoke the internal representations *a*, *b*, and *c*, respectively (Holland, 1981, could rule out peripheral mediation as a plausible explanation of his results). First the AB pairings establish an internal chain *a* → *b*. The AC stage then allows *b* and *c* to occur together, which yields a new internal chain *b* → *c*. This produces the result observed on B testing: B → *b* → *c* → aversion. Holland’s model obviously parallels Horne and Lowe’s account of [AB, AC]:BC effects (see above). The [AB, BC]:AC transitivity results of monkeys (D’Amato, Salmon, Loukas, & Tomie, 1985) and pigeons (Kuno, Kitadate, & Iwamoto, 1994) could be explained along similar lines.

Other data raise doubts about Horne and Lowe’s picture of nonhuman behavior as being “locked” in forward relations (Horne &

Lowe, p. 223). With a transfer design, Hogan and Zentall (1977, Experiment III) did find some evidence for [AB]:BA effects in pigeons (see their Figure 5, p. 12), although the obtained effects were both transient and small in magnitude. Richards (1988, Experiments 2 and 3) also observed [AB]:BA transfer in pigeons with conditional discrimination tasks. The effects were small in magnitude, but persisted over sessions.

Recently Zentall, Sherburne, and Steirn (1992) found strong evidence for what they termed *backward associations* in pigeons' matching to sample. Subjects exposed to  $X \rightarrow A \rightarrow B$  trials, where  $X$  is a sample,  $A$  is a comparison, and  $B$  is some differential outcome (food or no food), showed clear transfer on a  $B \rightarrow A$  matching task (for a related effect in the budgerigar, see Manabe, Kawashima, & Staddon, 1995, Experiment 3). Zentall et al. (Experiment 3) showed that Pavlovian mediation involving the  $X$  samples probably did not underlie the observed effects, which thus took place solely at the level of the  $A$  and  $B$  stimuli.

Hearst (1989) similarly found strong [BA]:AB effects with pigeons in an autoshaping situation. Subjects exposed to BA pairings between food ( $B$ ) and a neutral stimulus ( $A$ ) showed clear positive transfer in a second stage of AB forward conditioning. The point is not that the pigeons showed backward conditioning; in fact, during the initial BA stage they virtually did not peck at Stimulus  $A$ . The point is rather that the pigeons showed [BA]:AB transfer.

Such data cannot be accounted for on the sole basis of strictly unidirectional principles of behavior. Besides, Horne and Lowe insist that the effects of common naming on equivalence class formation depend on the bidirectional nature of speaker-listener skills, as opposed to the mere establishment of a common response to different stimuli (Horne & Lowe, p. 225). But pigeons do tend to match stimuli that have previously been paired with the same outcome (Edwards, Jagielo, Zentall, & Hogan, 1982). Following Horne and Lowe's logic (Horne & Lowe, p. 225), one should therefore attribute internal echoic skills to pigeons.

And why not, indeed? With its reliance on the contiguities afforded by echoics to produce internal chains (" $b$ "  $\rightarrow$  " $a$ ," etc.), the

naming account closely parallels cognitive models of association based on covert rehearsal (e.g., Atkinson & Shiffrin, 1968). Because numerous data (e.g., Wagner, Rudy, & Whitlow, 1973) suggest the existence of rehearsal processes in nonhuman species, Horne and Lowe should be willing to extend their approach to the nonverbal mediators of nonhuman subjects (cf. Urcuioli, Zentall, & DeMarse, 1995). Backward associations in nonhumans would be attributed to the subjects' rehearsal of their own perceptual activities (Holland, 1990; Horne & Lowe, p. 194), just as the naming account attributes human [AB]:BA effects to the subjects' covert verbal echoics (see Wagner, 1981).

#### *Toward a Nonmediational Account*

Any explanatory framework that is incomplete or inadequate at the overt behavioral level can be made to fit numerous data by positing a variety of internal mediators. We believe, however, that as behavioral theory internalizes its terms, it becomes increasingly indistinguishable from cognitive psychology (see Rachlin, 1988; Schnaitter, 1987). A possible alternative to mediational accounts would be to formulate nonstandard behavioral principles at a molar level; although such an alternative remains to be developed, some properties of the relevant laws can be easily understood by reference to Holland's [AB, AC]:BC experiment (1981). The two stages of Holland's basic training paradigm appear as follows, where time flows from left to right: AB . . . . . AC. Neglecting the  $A$  stimuli, it is clear that Holland's procedure includes a molar sequence, .B . . . . . C, that differs from a standard BC pairing only in being more extended in time. From a molar viewpoint, the effect of the added  $A$  stimuli in Holland's experiment is to potentiate direct conditioning by this molar sequence (cf. Himeline, 1977, p. 397). Because this sequence is similar to a compact BC sequence (or pairing) in terms of internal composition, that it also results in a conditional response to  $B$  is not surprising.

A similar formulation could be applied to the AB and AC trials of match to sample. In addition, we suspect that the covert verbal activities postulated by Horne and Lowe could be replaced by molar sequences (Tonneau, 1990) including the past *overt* responses that

current mediators allegedly represent (see Rachlin, 1985). In the naming framework, after all, each internal response (“a”) can be traced back to some prior overt behavior (“A”).

Substituting combinations of molar sequences to chains of covert mediators, however, will require behavioral laws that are more complex than those assumed by Horne and Lowe. Adequate molar laws will probably incorporate response-independent forward and backward stimulus correlations (cf. Stemmer, 1973; Tonneau, 1993), consistent with numerous data showing the sensitivity of operant performance to Pavlovian relations (e.g., Davidson, Aparicio, & Rescorla, 1988; Ross & LoLordo, 1987). “Contingencies of reinforcement” may thus lose their basic status to become components of a broader causal system (see Malone, 1975; Pierce & Epling, 1984); our main concern with Horne and Lowe’s approach is that it adheres to behavioral principles that are clearly in need of revision (Hinson, 1987; Malone, 1987).

#### REFERENCES

- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). New York: Academic Press.
- D’Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Journal of the Experimental Analysis of Behavior*, *44*, 35–47.
- Davidson, T. L., Aparicio, J., & Rescorla, R. A. (1988). Transfer between Pavlovian facilitators and instrumental discriminative stimuli. *Animal Learning & Behavior*, *16*, 285–291.
- Edwards, C. A., Jagielo, J. A., Zentall, T. R., & Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching to sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*, 244–259.
- Hearst, E. (1989). Backward associations: Differential learning about stimuli that follow the presence versus the absence of food in pigeons. *Animal Learning & Behavior*, *17*, 280–290.
- Hineline, P. N. (1977). Negative reinforcement and avoidance. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 364–414). Englewood Cliffs, NJ: Prentice Hall.
- Hinson, J. (1987). Interchange [Reply to Malone]. In S. Modgil & C. Modgil (Eds.), *B. F. Skinner: Consensus and controversy* (pp. 204–205). Barcombe, Lewes, England: Falmer Press.
- Hogan, D. E., & Zentall, T. R. (1977). Backward associations in the pigeon. *American Journal of Psychology*, *90*, 3–15.
- Holland, P. C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learning and Motivation*, *12*, 1–18.
- Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition*, *37*, 105–131.
- 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.
- Kuno, H., Kitadate, T., & Iwamoto, T. (1994). Formation of transitivity in conditional matching to sample by pigeons. *Journal of the Experimental Analysis of Behavior*, *62*, 399–408.
- Malone, J. C., Jr. (1975). William James and B. F. Skinner: Behaviorism, reinforcement, and interest. *Behaviorism*, *3*, 140–151.
- Malone, J. C., Jr. (1987). Interchange [Reply to Hinson]. In S. Modgil & C. Modgil (Eds.), *B. F. Skinner: Consensus and controversy* (pp. 205–206). Barcombe, Lewes, England: Falmer Press.
- Manabe, K., Kawashima, T., & Staddon, J. E. R. (1995). Differential vocalization in budgerigars: Towards an experimental analysis of naming. *Journal of the Experimental Analysis of Behavior*, *63*, 111–126.
- Pierce, W. D., & Epling, W. F. (1984). On the persistence of cognitive explanation: Implications for behavior analysis. *Behaviorism*, *12*(1), 15–27.
- Rachlin, H. (1985). Pain and behavior. *Behavioral and Brain Sciences*, *8*, 43–83. (Includes commentary)
- Rachlin, H. (1988). Molar behaviorism. In D. B. Fishman, F. Rotgers, & C. M. Franks (Eds.), *Paradigms in behavior therapy: Present and promise* (pp. 77–105). New York: Springer.
- Richards, R. W. (1988). The question of bidirectional associations in pigeons’ learning of conditional discrimination tasks. *Bulletin of the Psychonomic Society*, *26*, 577–579.
- Ross, R. T., & LoLordo, V. M. (1987). Evaluation of the relation between Pavlovian occasion-setting and instrumental discriminative stimuli: A blocking analysis. *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 3–16.
- Schnaitter, R. (1987). Behaviorism is not cognitive and cognitivism is not behavioral. *Behaviorism*, *15*, 1–11.
- Stemmer, N. (1973). Language acquisition and classical conditioning. *Language and Speech*, *16*, 279–282.
- Tonneau, F. (1990). From reflex to memory: Molar sequences in Pavlovian and instrumental conditioning. *The Psychological Record*, *40*, 587–607.
- Tonneau, F. (1993). Stimulus correlations in complex operant settings. *Behavioral and Brain Sciences*, *16*, 393–394.
- Urcuioli, P. J., Zentall, T. R., & DeMarse, T. (1995). Transfer to derived sample-comparison relations by pigeons following many-to-one versus one-to-many matching with identical training relations. *Quarterly Journal of Experimental Psychology*, *48B*, 158–178.
- Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., & Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*, 264–273.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear &

- R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Erlbaum.
- Wagner, A. R., Rudy, J. W., & Whitlow, J. W. (1973). Rehearsal in animal conditioning [Monograph]. *Journal of Experimental Psychology*, 97, 407–426.
- Zentall, T. R., Sherburne, L. M., & Steirn, J. N. (1992). Development of excitatory backward associations during the establishment of forward associations in a delayed conditional discrimination by pigeons. *Animal Learning & Behavior*, 20, 199–206.