JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1989, 51, 393-396 NUMBER 3 (MAY)

REPLY TO SAUNDERS AND TO HAYES

Stimulus equivalence is defined by the reflexive, symmetrical, and transitive properties of a set of stimuli. Equivalence tests typically follow the training of stimulus pairs using a conditional match-to-sample procedure. Equivalence is tested by presenting the training stimuli in pairs that were not presented during training and/or presenting trained stimulus pairs in reverse temporal order. The properties of reflexivity, symmetry, and transitivity, and thus, equivalence, are shown when the members of the untrained, but tested, stimulus pairs are related to the subject's behavior in the same way as the trained pairs. These training/testing procedures have produced robust, reliable equivalence class formation by humans. However, prior to the study by McIntire, Cleary, and Thompson (1987), attempts to establish equivalence classes in nonhumans had failed (e.g., D'Amato, Salmon, Loukas, & Tomie, 1985). McIntire et al. (1987) demonstrated that 2 cynomolgous monkeys, each of which was trained to perform one of two topographically distinct responses in the presence of each stimulus of two three-member stimulus sets during conditional match-to-sample training, maintained the relationships among the members of the set when tested with novel pairings from each stimulus set and reverse temporal orderings of trained pairs. The test results showed reflexivity, symmetry, and transitivity (i.e., equivalence).

The central disagreement of both Saunders and Hayes with our interpretation is that the performance of our monkeys, Rang and Manley, did not demonstrate stimulus equivalence because the basis of their performance was fundamentally different from that underlying the establishment of stimulus equivalence by humans. We will present our comments in four sections: (a) a note on comparative research, (b) a note on issues of definition, (c) comments on procedures, and (d) concluding comments.

Comparative Research

Comparative research is designed to elucidate the similarities and differences among species. In addition, the use of nonhumans provides a degree of simplification and control often absent in studies of human behavior. The purpose of our study was not to show that equivalence with monkeys is as rich as human equivalence (i.e., that monkey equivalence is unconditionally equivalent to human equivalence). We specified in the introduction of our paper that establishing an appropriate mechanism (heuristically called *naming*) might produce responses to rearranged pairs of sample and comparison stimuli which, *if performed by a human,* would be called equivalence. We also stated that we had probably reached the limits of Rang and Manley's equivalence capacities with these procedures.

It is unlikely that the standard match-to-sample procedure will result in the spontaneous development of the requisite mechanism for establishing equivalence by nonhumans. Failure of a number of well-conducted attempts to obtain equivalence with nonhumans suggests that equivalence (responding accurately to rearranged pairs of sample and comparison stimuli) will not occur in nonhumans without special procedural approaches not usually used in human testing situations. On the other hand, nondisabled human subjects (even children) characteristically have rich natural histories of verbal behavior and equivalence class formation prior to being observed under the controlled conditions of the laboratory. Simply, the experimenter has little account of, and no control over, the complex histories that human subjects bring to the laboratory. The histories may be as useful in elucidating the mechanisms of equivalence class formation as the procedures used in the laboratory. Additionally, it seems unlikely that the mechanisms underlying the equivalence performance of even children (if such mechanisms are used in equivalence class formation) would be simple chains such as those of Rang and Manley. The value of additional comparative research lies in the ability to control the subjects' histories and to manipulate variables to an extent impossible with human subjects.

The authors thank Blaine Peden for comments on this paper. Correspondence and requests for reprints should be sent to Kenneth D. McIntire, Psychology Department, University of Wisconsin-Eau Claire, Eau Claire, Wisconsin 54702.

Even if future laboratory research using variations of the conditional match-to-sample procedure demonstrates additional mechanisms resulting in equivalence class formation in a variety of species, they will only be demonstrations under controlled conditions. It is possible that there are few natural contingencies related to the behavior of nonhumans resulting in the formation of equivalence performance. Analogously, apes don't show the complex variety of manual signs in the wild that they can perform after appropriate training by humans (ape signing as language is a different issue altogether). It is also possible that procedural variations will find limited domains of species-related abilities that are functionally adaptive.

The Matter of Definition

The most significant issue arising from our use of the term *equivalence* to describe the performance of Rang and Manley is the matter of class inclusion. That is, what is the equivalence class of equivalence class? This is a conceptual issue rather than an empirical issue. It is similar to the issue surrounding complex discriminative and signing performance by apes—is it language or not? However, the definitional issue with equivalence classes is potentially less controversial than language because the term *equivalence class* is not derived from the vocabulary of everyday discourse, as is the word *language*. Empirically, the problem is to compare and contrast the performance of a variety of species (including humans) under controlled conditions and to elucidate the threads of continuity underlying their behavior. Conceptually, the problem is to decide at which point the performance differences among species require excluding a performance from equivalence class membership. Because the term *equivalence* was derived from complex human discriminative performance, human performance should be the standard. As with *language,* it is possible to establish a definition of equivalence that excludes all nonhuman performance and makes irrelevant all work derived from different species and different conceptual origins. That has been the strategy some linguists have used in attacking Skinner's analysis of verbal behavior. However, to do so is to risk (a) defining away a breadth of potential species commonalities that may underlie complex discriminative performance, (b) developing unduly restrictive conceptual frameworks, and (c) accepting phenomena as "emergent" when they may be accountable by another set of principles. It would be a mistake to exclude the potential relevance of past research (human and nonhuman) on "acquired equivalence of cues," "mediated generalization," "response mediation," "cognition," and so on, simply because the name of the work differs from that which is being used here or because the procedures and underlying conceptual frameworks are not isomorphic with those currently used. Moreover, the inadequacies of earlier conceptual systems do not negate the potential relevance of the observations.

Procedural Comments

The principal criticism of both Saunders and Hayes is that the performances of Rang and Manley are considered more properly as instances of response chains than emergent equivalences. Thus, the relationships that were maintained among rearranged sample and comparison stimuli by Rang and Manley were established by a process different from that involved when humans show similar terminal performance. We established at the outset that we were establishing response chains in Rang and Manley, and in Figure 5 (McIntire et al., 1987) we specified what we believed to be the primary controlling relations in their performance. It is not clear that our analysis differs substantially from that of either Saunders or Hayes on this point; that is, the emission of the sample naming response was discriminative for the selection of the comparison stimulus and the emission of the second naming response. Additionally, it was implicit in the procedure that the maintenance of stimulus relations in the face of sample and comparison rearrangements would be a function of the naming responses.

Saunders and Hayes propose that Rang and Manley's performance was not equivalence for two reasons. First, the comparison stimulus selection was a function of the naming response, an S-R mechanism. Second, all S-R and R-S relations were trained. Thus, there were no emergent relationships in testing. Inherent in their analyses are two assumptions: (a) that human equivalence and conditional discrimination performance result from S-S relations, and (b) equivalence emerges from an undefined, human-specific property after conditional discrimination training. The current literature is inconclusive as to whether conditional discrimination performance and equivalence are best accounted for by S-R or S-S hypotheses. This issue appears to be a potentially rich and fruitful garden of future research. However, Rang and Manley demonstrated that a specifically trained S-R mechanism could result in the maintenance of stimulus relationships despite alterations in stimulus pairings and temporal order. The monkeys' equivalence performance would have appeared emergent had we not established the nature of the controlling S-R relationships. It is possible that human equivalence may also be accounted for by established S-R principles.

Saunders raises two interesting empirical questions based upon alterations of our basic procedure. The first involves an extension of our inferential statement that heterogeneous chains may be as effective in maintaining relations among test pairs as the homogeneous chains we used. The form of the model that is presented clearly results in a conflict in controlling variables that would be ambiguous and result in a loss of discriminative control. Although it is an interesting problem, it does not elucidate the results of the published study nor mediational hypotheses in general. The second empirical issue concerns Saunders' assertion that we might expect test performance similar to Rang and Manley's if we had trained only the six identity pairings (three odd and three even). Unfortunately, the data are not available to answer this question. We hope someone conducts this hypothetical experiment, perhaps with human subjects.

Concluding Comments

Whether stimulus equivalence is best accounted for by S-R or S-S principles (i.e., does language result *in* equivalence or result *from* equivalence—if either?) is not resolvable at present. However, Rang and Manley demonstrated that the presence of an S-R mechanism allowed relationships among stimulus class members to be maintained in the presence of novel pairings and reversed temporal order. We showed how these relationships could occur in nonhumans and may similarly occur in humans even though they are not explicitly trained or observed by the experimenter. Hayes presents the following analog to describe our procedures:

> . . . if an animal turned on an orange light by touching stimuli x, y, or z, and was trained to pick only x, y, or z in the presence of the orange light, the animal will naturally be able to go from any one of the three stimuli to any other. If x is the sample, and z and p are the comparison stimuli, the animal would respond to x and thus turn on an orange light and now pick z over p because of the explicit reinforcement history for selecting z given an orange light. This example captures precisely the experimental arrangement used by McIntire et al. (Wulfert & Hayes, 1988, p. 126)

It is arguable whether this procedure "captures precisely" our experimental arrangement. For example, our Stimuli 5 and 6 were used as sample stimuli only during testing. It is uncertain how the subjects in Hayes' hypothetical experiment would respond in testing if Stimulus z was not presented on the sample key during training but was followed by an orange light and reinforcement as a comparison stimulus. However, let us conduct Hayes' thought experiment in a chamber with an observation window on either side. Through Window A an observer can see the subject, the sample and comparison stimuli, and the orange light located above the stimuli. Through Window B an observer can see only the subject and the sample and comparison stimuli. Stimuli x, y, and z are trained as members of one set in a standard conditional match-to-sample procedure with the pairings used for Rang and Manley (e.g., x-x, x-y, y-z), and the orange light is presented after the response to the sample (and comparison). After the subject's behavior stabilizes, novel pairings and reverse temporal orderings of the stimuli are introduced (x-z, z-x, etc.). The observer looking through Window A would see the orange light functioning as part of a chain to maintain the test stimulus pairings (i.e., the equivalent of Rang and Manley's naming responses). The observer looking through Window B, unable to see the orange light, would see emergent equivalence relations. It is possible that Saunders and Hayes are looking at human equivalence through Window B and are unable to see the light visible through Window A.

Kenneth D. McIntire
University of Wisconsin–Eau Claire

James Cleary
University of Minnesota

Travis Thompson
University of Minnesota

REFERENCES

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.

McIntire, K. D., Cleary, J., & Thompson, T. (1987). Conditional relations by monkeys: Reflexivity, symmetry, and transitivity. *Journal of the Experimental Analysis of Behavior,* **47,** 279–285.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior,* **50,** 125–144.