JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1992, 58, 123-133 NUMBER 1 (JULY)

THE DEVELOPMENT OF FUNCTIONAL AND EQUIVALENCE CLASSES IN HIGH-FUNCTIONING AUTISTIC CHILDREN: THE ROLE OF NAMING

SVEIN EIKESETH AND TRISTRAM SMITH

UNIVERSITY OF KANSAS AND UNIVERSITY OF CALIFORNIA, LOS ANGELES

The development of functional and equivalence classes was studied in four high-functioning, preschool-aged autistic children. Initially, all subjects failed to demonstrate match-to-sample relations indicative of stimulus equivalence among two three-member classes of visual stimuli. Then, 2 subjects showed emergence of those relations after they were taught to assign the same name to all members in each class. Next, subjects were taught names for new stimuli outside the match-to-sample format. On subsequent match-to-sample tests, 2 subjects demonstrated untrained conditional relations among the stimuli given a common name. New, unnamed stimuli were then related via match-to-sample training to stimuli from sets of named stimuli. Tests for emergent conditional relations between the new unnamed stimuli and the named stimuli yielded positive results for 1 subject and somewhat mixed results for 3 subjects. Finally, without naming, 2 subjects developed stimulus equivalence among two new three-member classes of visual stimuli. These data suggest that naming may remediate failures to develop untrained conditional relations, some of which are indicative of stimulus equivalence.

Key words: equivalence classes, functional classes, naming, conditional discrimination, discrimination, matching to sample, autistic children

---

The research program of behavior analysis has often been criticized for ignoring behavior that is acquired without explicit training, and applied behavior analysis has been criticized for failing to produce such behavior in educational and therapeutic settings (e.g., Gardner, 1987). In fact, however, there is a growing behavior-analytic literature on untrained behavioral relations, much of which concerns the development of stimulus classes. A special case of stimulus classes, stimulus equivalence, has attracted particular interest: Stimulus equivalence occurs when a subject correctly matches any member of a class of stimuli with any other member of that class, despite having been trained on only a subset of the possible matches.

More precisely, stimulus equivalence is demonstrated when the stimuli in a conditional relation display the properties of reflexivity, symmetry, and transitivity (Sidman et al., 1982; Sidman & Tailby, 1982). *Reflexivity* is demonstrated when subjects, without explicit training, match any stimulus of a prospective class to an identical stimulus. *Symmetry* is demonstrated when the sample and comparison stimuli in a match-to-sample discrimination can be reversed and still yield the same function. *Transitivity* is demonstrated when a subject who is taught to match A1 to B1 and B1 to C1 then will match A1 to C1 without additional training.

Verbal human subjects as young as 2 years old readily develop equivalence classes (Devany, Hayes, & Nelson, 1986). To date, however, there have been no clear demonstrations of stimulus equivalence in nonhuman subjects. Unsuccessful attempts have been made with pigeons (D'Amato, Salmon, Loukas, & Tomie, 1985; Lipkens, Kop, & Matthijs, 1988: Sidman et al., 1982), monkeys, baboons, and other primates (D'Amato et al., 1985; Sidman et al., 1982). Two reports of successful attempts have appeared, one with monkeys (McIntire, Cleary, & Thompson, 1987) and the other with pigeons (Vaughan, 1988). The validity of these reports, however, has been disputed (Hayes, 1989; Saunders, 1989). For example, McIntire et al. (1987) trained monkeys to produce a common response (a name) to stimuli in prospective stimulus classes. During both training and testing, the monkeys were required to emit the appropriate name when presented with the sample and then to do so again when presented with the correct com-

The authors thank Donald M. Baer, Ivar Lovaas, Kathryn Saunders, and Joseph Spradlin for their comments at earlier stages of the preparation of this manuscript. Preparation was supported in part by a U.S. Office of Education Grant (H133G 80103) to O. Ivar Lovaas. Portions of this research were presented at the annual meeting of the Association of Behavior Analysis, May 1989, Milwaukee, Wisconsin. Correspondence should be sent to Svein Eikeseth, Department of Psychology, 405 Hilgard Ave., University of California, Los Angeles, Los Angeles, California 90024-1563.

parison stimulus. The monkeys showed mastery of all possible match-to-sample relations among stimuli in a class after being presented with a subset of these relations during training.

Hayes (1989) and Saunders (1989) have pointed out that the relations observed in McIntire et al.'s (1987) study may not have been indicative of stimulus equivalence, as defined by Sidman and Tailby (1982). This is because the definition requires that stimulus equivalence emerge without direct training, and it is not clear that the naming procedure used by McIntire et al. met this requirement. Rather than equivalence classes, they may have produced functional classes. Goldiamond (1962) defined a functional class as "a set of stimuli in the presence of which a similar response may occur" (p. 303). Stimuli assigned the same name constitute a functional class because the different stimuli in the class occasion the same response (i.e., the same name).

A number of investigators have been interested in the relation between verbal behavior (e.g., naming) and the emergence of equivalence classes. Some evidence for the importance of naming comes from the finding that subjects frequently apply a common name to the members of equivalence classes (e.g., Mackay & Sidman, 1984; Sidman, 1971; Sidman & Cresson, 1973). On the other hand, studies have shown that some subjects who develop equivalence classes do not necessarily apply a common name to all stimuli in the class when they are asked to provide names for the stimuli after completing the equivalence probes (Lazar, Davis-Lang, & Sanchez, 1984; Sidman & Tailby, 1982; Sidman, Willson-Morris, & Kirk, 1986). Such findings have led some authors to conclude that naming is neither necessary nor sufficient for the development of equivalence classes (e.g., Lazar et al., 1984; Sidman et al., 1986). However, the naming test simply assesses correlations between naming and equivalence (i.e., whether subjects who show equivalence also tend to show naming), not the functional relation between the two (cf. Skinner, 1950). Thus, future studies may best be directed toward an analysis of the contingencies that produce stimulus equivalence rather than merely probing for correlations between equivalence and verbal classes.

As a step in this direction, the present study explored whether naming facilitated the development of untrained conditional relations indicative of stimulus equivalence in subjects who previously failed to show such relations. The subjects were 4 high-functioning, preschool-aged autistic children. This population was considered an appropriate target for efforts to examine conditions under which equivalence relations develop because autistic children seem to be less likely to display untrained relations (such as equivalence relations) than other human populations (Lovaas, 1977).

METHOD

Subjects

Four subjects participated in the experiment. All subjects were concurrently enrolled in an intensive behavioral treatment program at the UCLA Autism Project (Lovaas & Smith, 1988), where they had been in treatment for 1 to 2 years. Before enrolling in the program, each subject had received two independent diagnoses of autism, based on DSM-III-R (APA, 1987) criteria, from psychiatrists or licensed clinical psychologists. When the subjects began treatment, they lacked expressive and receptive language, avoided interactions with others, failed to play appropriately with toys, displayed numerous types of self-stimulatory behavior (high-rate, stereotyped acts such as rocking back repeatedly), exhibited severe tantrums, and scored in the mentally retarded range on standardized tests of intellectual functioning. As part of treatment, all subjects were taught a variety of new skills. Three of these skills were particularly relevant to the present experiment: (a) generalized identity matching, (b) generalized imitation of verbal stimuli, and (c) expressive naming of visual stimuli (Lovaas, 1981). By the time of the experiment, all subjects had acquired some language, posed fewer management difficulties, and performed more satisfactorily on standardized tests. However, they continued to have some problems in each of these areas (particularly language), as will be discussed shortly. Thus, the subjects could be classified as high-functioning autistic children.

Immediately prior to the experiment, the subjects took the following tests: (a) the Leiter International Performance Scale (Leiter, 1959), which measures visual-spatial skills, with an emphasis on matching of visual stimuli; (b) the Peabody Picture Vocabulary Test—Revised (Dunn, 1981), which measures recep-

Table 1

Mental ages calculated from the Leiter International Performance Scale, the communication part of the Vineland Adaptive Behavior Scale (expanded version), and the Peabody Picture Vocabulary Test, together with each subject's chronological age, expressed in years and months.

Subject	Chronological age	Leiter	Vineland: Receptive	Vineland: Expressive	Vineland: Written	Vineland: Communication	Peabody
Trey	3-6	4-3	4-1	2-10	6-3	3-9	2-6
Joe	3-10	6-0	4-0	2-6	5-2	3-1	2-8
Danny	5-6	5-4	4-2	3-6	5-11	4-6	3-8
Rory	4-5	4-7	2-7	1-11	4-2	2-1	<2-4[a]

[a] Basal score not obtained.

tive vocabulary; and (c) the communication section of the Vineland Adaptive Behavior Scale (Sparrow, Balla, & Cichetti, 1984), which assesses the use in everyday situations of receptive language, expressive language, and writing skills as reported by a child's caregiver. Chronological ages and test results are presented for all subjects in Table 1.

The subjects ranged in chronological age from 3.5 to 5.5 years. They performed at average or above average levels on the Leiter and the writing section of the Vineland, indicating high levels of visual-spatial and writing skills. However, they tended to score below age level on tests for expressive and receptive language. Two subjects, Trey and Joe, scored in the normal range on the receptive language section of the Vineland, but they scored below normal on the Peabody, which also assesses receptive language. This discrepancy may be due to differences in the tests: The Vineland focuses on how subjects use receptive language to function adaptively as reported by a caregiver, whereas the Peabody directly samples how many receptive labels subjects know. In general, then, all subjects exhibited average or above average visual-spatial skills but displayed deficits in language.

Stimuli

The stimuli (shown in Figure 1) were Greek letters and their printed names presented visually to subjects on sheets of white paper (33 cm by 22 cm). Each sheet contained three rectangles (10 cm by 7 cm). When the long side of the sheet was placed horizontally, the sample stimulus was always in the top rectangle and the two comparison stimuli were side by side in the two bottom rectangles. There was one sheet for each possible arrangement of the stimuli.

Design

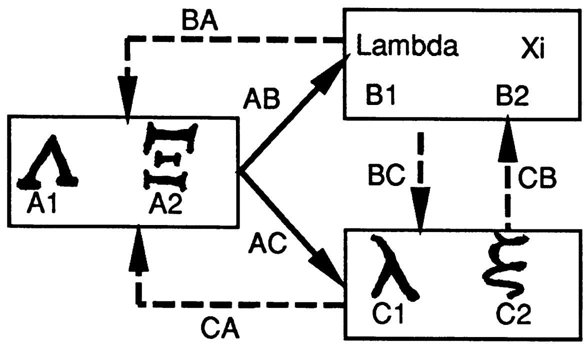
The basic design is summarized in Figure 1 and can be described as follows:

Phase 1. Would the autistic subjects develop conditional relations indicative of stimulus equivalence? After the AB and AC baseline conditional discriminations were taught, the BC and CB relations were assessed to test for stimulus equivalence (top of Figure 1). If subjects responded correctly on fewer than 90% of the equivalence probes (i.e., if subjects failed to show mastery), symmetry was tested by assessing the number of correct responses to the BA and CA conditional discriminations. Then, regardless of the result, the BC and CB relations were assessed again. Finally, a test for the AB and AC relations was conducted to assess any extinction of the baseline discriminations.

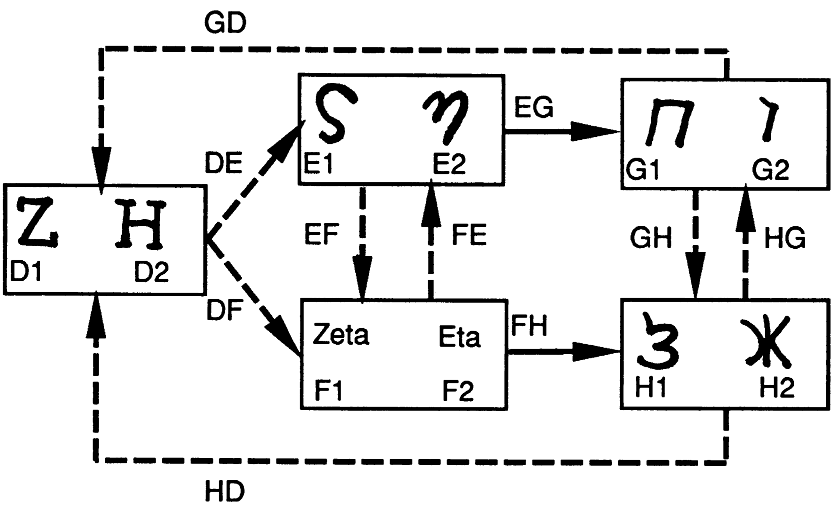
Phase 2. If equivalence relations, as assessed by correct responding to the BC and CB probes, failed to appear, would naming facilitate the development of those relations? Subjects who failed to show equivalence during Phase 1 were taught to speak a common name in response to each member of the first stimulus class (A1, B1, C1) and to speak another common name as a label for each member of the second stimulus class (A2, B2, C2). Then subjects were required to perform the AB and AC baseline discriminations while correctly labeling the stimuli to which they pointed. Upon mastery of this task, the subjects' performance on the BC and CB relations was reassessed.

Phase 3. Would naming alone lead to the development of conditional relations among new sets of stimuli (D, E, and F) after the

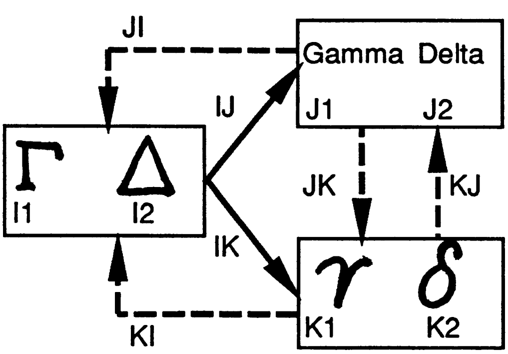
Phase 1 and 2



Phase 3 and 4



Phase 5



subjects were taught to speak a common name in response to each member of the first stimulus class (D1, E1, F1) and to speak another common name as a label for each member of the second stimulus class (D2, E2, F2)? After the subjects learned to label the stimuli correctly, the DE and DF conditional relations were assessed, followed by a test of the EF and FE conditional relations, without any prior match-to-sample training with these stimuli. Any subject who failed to develop the DE and DF relations was taught those relations before the EF and FE relations were assessed.

Phase 4. After teaching subjects to relate new unnamed stimuli to named stimuli via matching to sample, (a) would subjects develop any untrained conditional relations between other named stimuli and the new unnamed stimuli, and (b) would untrained conditional relations among the new unnamed stimuli develop? Two new stimuli (G and H) were related to each of the two classes of D, E, and F stimuli by teaching the EG and FH conditional discriminations (without teaching names for the G and H stimuli); then the GD and HD conditional discriminations were assessed, and finally the GH and HG relations were assessed.

Phase 5. Would subjects still fail to show evidence of stimulus equivalence in new conditional relations in which naming was not programmed? For new stimuli (I, J, and K) the IJ and IK conditional discriminations were taught without naming, and the JK, KJ, JI, and KI relations were assessed (bottom of Figure 1).

Summary. This sequence of training and testing formed an ABCDA design. Condition A (Phases 1 and 5) tested for equivalence after teaching conditional discriminations without naming; Condition B (Phase 2) added naming; Condition C (Phase 3) assessed the effects of naming without match-to-sample training; and Condition D (Phase 4) mixed unnamed and named stimuli.

Training Procedures

Training was conducted during subjects' regular therapy sessions in their homes. Behavior therapists with at least 6 months of supervised training from the UCLA Autism Project conducted the sessions. They were blind as to the purpose of the experiment. Prior to data collection, they held at least one practice session with the first author to familiarize themselves with the procedures. Throughout training, therapists reinforced correct responses with praise, smiles, and hugs, individually selected and demonstrated to be functional for each child. Incorrect responses resulted in an informational "no" and a quick removal of the training stimuli. They gave a loud "NO" if subjects engaged in an incompatible behavior such as self-stimulation.

Conditional discriminations. The procedures used in the present experiment were similar to standard match-to-sample procedures for establishing conditional discriminations in autistic children (Lovaas, 1981): Training sessions lasted between 20 and 30 min and consisted of approximately 120 trials. Three 2-min breaks were given during each session, always contingent upon correct responding. On some days, more than one training session took place. First, subjects were taught to select B1 and not B2 in the presence of A1 on 9 of 10 consecutive trials. Then subjects were taught to select B2 and not B1 in the presence of A2 on 9 of 10 consecutive trials. Next, these two tasks were mixed by presenting stimulus sheets from both sets together in random order (i.e., teaching the AB conditional discrimination). Once this discrimination was mastered (i.e., 9 of 10 consecutive trials correct), the AC conditional discrimination was taught in the same manner. Finally, the AB and AC conditional discriminations were mixed and presented in random order, with the position of the correct comparison stimulus counterbalanced across trials. Criterion for mastery was always 9 of 10 consecutive trials correct.

Each trial started when the therapist placed a sheet of stimuli in front of the subject and said "point to same." A correct response was defined as pointing first to the sample stimulus and then to the correct comparison stimulus within 5 s of the onset of the trial. Initially,

←

Fig. 1. Training and testing paradigms. Arrows point from sample stimuli to comparison stimuli. The solid arrows represent conditional relations that were explicitly taught. The broken arrows represent tests for untrained conditional relations.

all correct responses were reinforced. After subjects responded correctly on 9 of 10 trials on the mixed conditional discriminations, the reinforcement schedule was gradually thinned to a variable-ratio (VR) 3 schedule to prevent extinction during testing. Training continued until subjects responded correctly on 9 of 10 trials on the mixed conditional discriminations on a VR 3 schedule.

Position prompts (presenting the same stimulus set on consecutive trials until subjects responded correctly), pointing to the correct stimuli, or manual prompts (physically guiding subjects to emit a correct response) were used when subjects responded incorrectly. If a position prompt failed to evoke a correct response, pointing was used on the next trial. If pointing failed to evoke a correct response, manual prompts were used.

Naming. Subjects were trained to label verbally each member of the prospective stimulus classes with a common name. For example, subjects were taught to speak "lambda" when shown Λ, λ, and printed word *lambda.* The procedure for teaching naming was as follows. Each stimulus was presented by itself in a rectangle (10 cm by 7 cm) at the center of a sheet of paper (33 cm by 22 cm). Each training trial started when the therapist placed a stimulus in front of the child and asked, "What is it?" A response was scored correct if the subject labeled the stimuli with the correct name within 5 s of the onset of the trial. On the first training trial for each stimulus, the therapist prompted by saying the name of the stimulus (e.g., "lambda"). Reinforcement occurred if the subject repeated this name. Over the next 5 to 10 trials, the therapist faded out the prompt, either by saying it more and more softly or by leaving off more and more of the word (e.g., "lambda," then "lambd," followed by "la" and "l"). Once the prompt had been faded, no additional prompts were given unless a subject failed on two consecutive trials. When this occurred, the therapist gave a minimal prompt (the last prompt used before prompts were eliminated) on the next trial. If the subject again failed to respond correctly, the therapist gave the maximum prompt (e.g., saying "lambda").

At the onset of training with each subject, one visual stimulus from a prospective stimulus class was presented until the subject named this stimulus correctly on at least 9 of 10 consecutive trials. Then a second stimulus from the same prospective class was presented in the same way, and then the third. Next, the three stimuli were presented in random order until the subject again achieved at least 9 of 10 correct responses on consecutive trials. In the next phase of training, this procedure was repeated with a second class of stimuli. In the final phase of training, all members of both stimulus classes were presented in random order. The criterion for mastery of naming was correct responding on at least 9 of 10 trials during the final phase of training. All correct responses were reinforced.

Visual naming. Only one of the subjects (Rory) was taught visual naming. This subject did not develop conditional relations following the oral naming procedure attempted during Phase 2, even though he mastered the oral names. Therefore, a visual naming procedure was used during subsequent phases in an attempt to remediate this failure to develop untrained conditional relations. The visual names were block structures built on top of stimuli. Each stimulus was presented by itself in a rectangle at the center of a sheet of paper. For the first prospective stimulus class, the subject was taught to build a tower of two blue blocks (3 cm by 3 cm by 1.5 cm) when each stimulus was presented one at a time. For the second prospective stimulus class, he was taught to build a pyramid of three orange blocks (1.5 cm by 1.5 cm by 3 cm) on top of the stimuli when each stimulus was presented one at a time. Each training trial started when the therapist placed a stimulus and the blocks (two blue and three orange) in front of the child and asked, "What is it?" A response was scored correct if the subject built the correct structure of blocks on top of the stimulus within 5 s of the onset of the trial. On the first training trial for each stimulus, the therapist prompted the correct response by building the correct structure on top of the stimulus (e.g., building the tower with the two blue blocks). Reinforcement occurred if the subject repeated this structure. Over the next 5 to 10 trials, the therapist faded the prompt by leaving off more and more of the structure (e.g., two blocks, then one block, followed by pointing to the correct blocks). Once the prompt had been faded, no additional prompts were given unless a subject failed on two consecutive trials. When this occurred, the therapist gave a minimal prompt (the last prompt used before prompts were eliminated) on the next trial. If the subject again failed to respond correctly, the therapist gave the maximum prompt. All correct responses were reinforced.

Testing Procedures

Testing was conducted by the same therapists in the same setting as training. The criterion for scoring correct responses was the same as in training. However, no consequences followed correct and incorrect responses. When the stimuli were removed after each trial, reinforcement was delivered on a VR 3 schedule for good sitting, good working, and so on. A relation was considered mastered if a subject responded correctly on at least 90% of the test trials.

Conditional discriminations. Blocks of 10 trials were administered for one type of test at a time (e.g., conditional relations that had been trained [baseline], symmetry, or equivalence). Each type of test contained four discriminations (e.g., A1B1, A2B2, A1C1, A2C2, for baseline; B1C1, B2C2, C1B1, C2B2, for equivalence; and, B1A1, B2A2, C1A1, C2A2, for symmetry). The discriminations were presented in a random order, with the restriction that each of the four discriminations appear at least twice within the 10 trials. The position of the correct comparison stimulus was counterbalanced across trials.

Each test trial started when the therapist placed a sheet of stimuli in front of the subject and said "Point to same." The subject was never asked to name the stimuli during tests for the conditional relations.

Naming. Blocks of 10 trials were administered. The stimuli were presented randomly, with the restriction that all stimuli were presented at least twice within the 10 trials. Each test trial started when the therapist placed one of the visual stimuli in front of the child and asked, "What is it?" (When visual naming was used [Rory in Phases 3 and 4], each trial started when the therapist placed a stimulus and the blocks in front of the child and asked, "What is it?")

Recording and Reliability

Observers naive to the purpose of the experiment scored the number of correct and incorrect responses during all sessions of training and testing. At the same time, they recorded any vocalizations that the subjects made. To monitor interobserver reliability, two observers independently scored 40% of the test sessions for each child. Reliability was calculated as the percentage of trials on which the observers agreed in their scoring; mean interobserver agreement across children was 96% (range, 91% to 100%).

RESULTS

Phase 1. Figure 2 presents results for each of the 4 subjects across the five phases of this study. During Phase 1, subjects were taught AB and AC conditional discriminations and then were tested to see whether they showed stimulus equivalence (as indicated by mastery of the untrained BC, CB, BA, and CA conditional relations). Subjects required between 360 and 1,546 trials to master the AB and AC conditional discriminations. Having achieved mastery (always defined as at least 90% correct responding), all subjects maintained the AB and AC discriminations during the remainder of Phase 1.

However, none of the subjects developed the BC and CB conditional relations that were indicative of stimulus equivalence. A possible exception to this pattern was Joe (second from the top in Figure 2), who performed at 90% accuracy on the first block of 10 trials probing equivalence, although he performed at chance levels on the remaining blocks of trials. It is unclear whether this subject's performance during the first block of trials represented random fluctuation in performance or whether the subject initially had acquired equivalence, only to lose it over the course of testing.

Trey and Joe demonstrated mastery of the BA and CA relations, which were indicative of symmetry. Danny performed at chance levels. Rory performed above chance but below mastery. In sum, none of the subjects showed clear indications of stimulus equivalence, although 2 subjects demonstrated symmetry.

Phase 2. During Phase 2, subjects were taught names for the A, B, and C stimuli used in Phase 1. The subjects required between 86 and 234 trials to acquire these names (Figure 2). After mastering the names, all subjects continued to show mastery of the original conditional discriminations (AB and AC). Moreover, Joe and Danny showed complete and

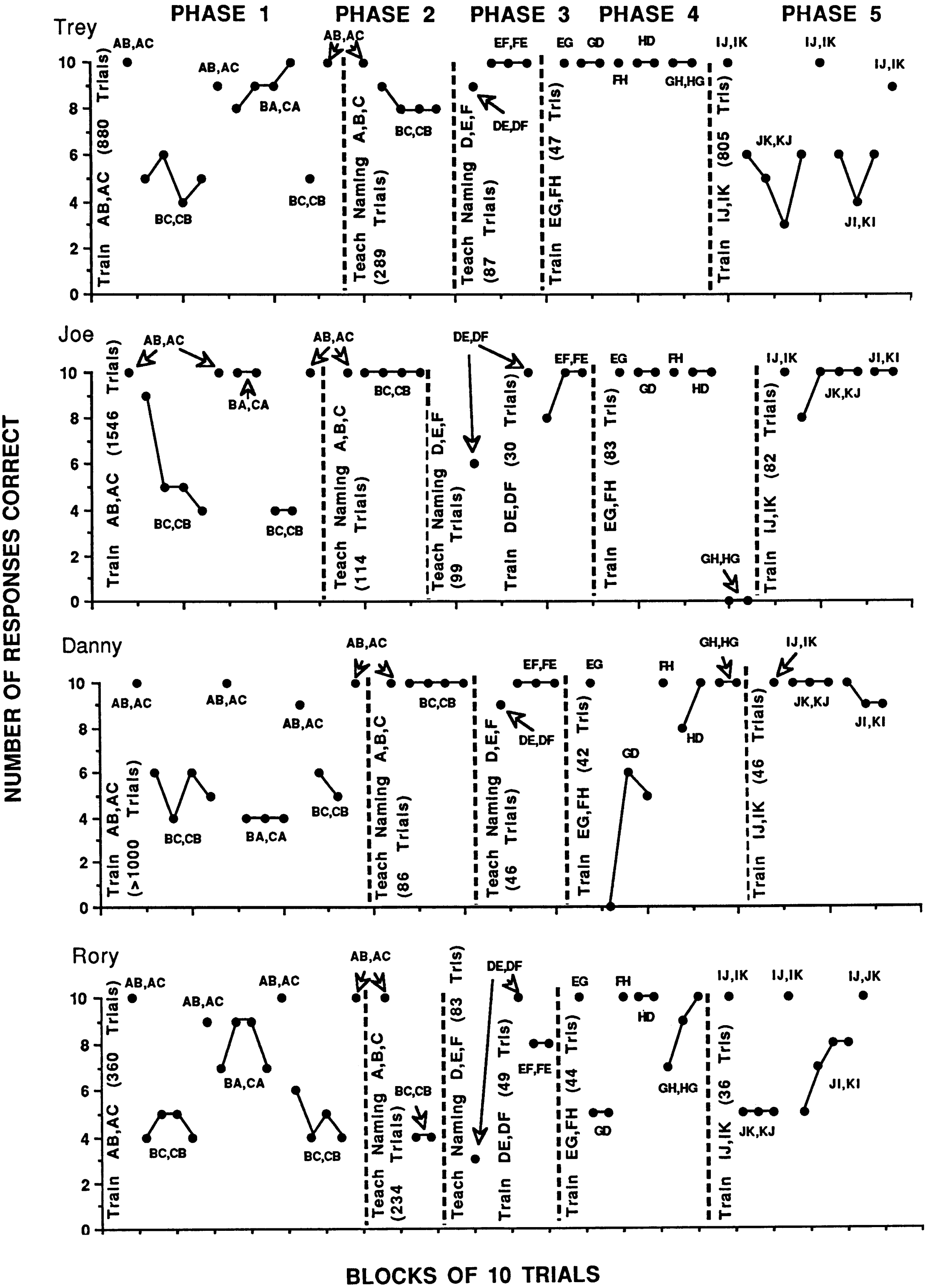


Fig. 2. Number of correct responses (vertical axis) across blocks of 10 trials (horizontal axis). The frequency polygons represent unreinforced test trials, with each dot representing 10 trials. The letters refer to the designations in Figure 1 and indicate which stimuli were involved as samples (first letter) and comparisons (second letter) in the tests.

immediate emergence of the untrained BC and CB relations (previously taken as evidence of equivalence), and Trey performed above chance but below mastery. Only 1 subject (Rory) continued to perform at chance levels on these probes. (Because this subject had lower levels of language skills than the remaining subjects but had good visual-spatial skills, it was decided to use the visual labeling procedure during Phase 3.) Thus, for at least 2 subjects, naming appeared to facilitate the emergence of the BC and CB relations that the subjects had previously failed to acquire without naming.

Phase 3. At the outset of Phase 3, subjects were taught to name new stimuli (D, E, and F) that had not previously been trained in a match-to-sample format. The subjects mastered the names in 46 to 99 trials (Figure 2). Without any additional instruction, Trey and Danny showed mastery of the DE, DF, EF, and FE conditional relations when presented in a match-to-sample format. However, Joe and Rory failed to acquire these emergent relations. Hence, they were trained to make the DE and DF conditional discriminations. Following this training, Joe acquired the untrained EF and FE relations, whereas Rory performed above chance but below mastery on tests for these emergent relations. Even though Rory's performance was still not perfect, he performed better after vocal naming was replaced with the visual labeling procedure. As in Phase 2, therefore, the naming procedure apparently facilitated the emergence of untrained conditional relations. This was particularly evident for Trey and Danny, but also occurred for Joe and Rory.

Phase 4. In Phase 4, subjects were trained to match stimuli from Phase 3 (E and F) with new, unnamed stimuli (G and H). Specifically, EG and FH conditional discriminations were trained. This training required 42 to 83 trials. Trey showed mastery of four conditional relations that had not been trained: GD, HD, GH, and HG. Joe showed mastery for GD and HD but not GH and HG. Joe consistently selected the wrong stimuli on tests of GH and HG, suggesting that perhaps some other unidentified source of stimulus control was responsible for his pattern of responding. Danny showed mastery of HD, GH, and HG but not GD. Rory showed mastery only of HD, although he approached mastery on tests of GH and HG (87% accuracy). In sum, probes for emergent conditional relations between unnamed stimuli and named stimuli yielded positive results for 1 subject (Trey) and somewhat mixed results for 3 subjects.

Phase 5. In this final phase, subjects were presented with new, unnamed stimuli (I, J, and K). All children were taught the IJ and IK conditional relations. Joe and Danny showed stimulus equivalence by demonstrating the emergence of the JK, KJ, JI, and KI relations reliably, but Trey and Rory did not (Figure 2).

Naming. During Phases 2 and 3, subjects were taught to label each stimulus in prospective classes with a common name. All subjects acquired the stimulus names rather quickly (Figure 2). During the match-to-sample probes in Phases 2, 3, and 4 (when naming had not been requested), Trey, Joe, and Rory continued to label the stimuli correctly, even when they responded incorrectly on tests of match-to-sample relations. For example, if subjects matched xi and lambda, they still labeled the stimuli "xi" and "lambda," rather than "xi" and "xi." Furthermore, in Phase 4, when unnamed stimuli were related to named stimuli via match-to-sample training, all 3 subjects labeled the new unnamed stimuli correctly while performing the match-to-sample tasks. However, none of the subjects used overt names during Phases 1 and 5 (i.e., when naming had not been programmed for any of the stimuli).

Danny did not use overt labels while performing the match-to-sample tasks. He labeled the stimuli only when required (i.e., when trained to name stimuli and tested for mastery of the names that had been trained).

DISCUSSION

The present study extended earlier attempts to examine the role of naming in the development of equivalence classes. We evaluated whether naming would facilitate the development of equivalence classes in subjects who previously failed to show evidence of equivalence without naming. We used naming in three ways: First, subjects were taught names for stimuli in conditional relations in which stimulus equivalence had failed to emerge. This tested whether naming could facilitate the development of the untrained conditional relations that failed to develop without naming. Second, subjects were taught names before stimuli were presented in the match-to-sample format. This tested whether untrained conditional relations would emerge among stimuli given the same name without any prior match-to-sample training with those stimuli. Finally, subjects were taught names for some of the members of a conditional relation. This tested whether untrained conditional relations would emerge between the named stimuli and the unnamed stimuli.

Trey showed the clearest relationship between naming and the development of untrained conditional relations. This subject consistently displayed mastery on all probes of untrained relations when naming had been programmed and consistently failed to do so when naming had not been programmed. The other subjects were also more likely to display untrained conditional relations when naming had been programmed, but they were not as consistent as Trey.

The naming procedure established the different stimuli given the same name as members of a functional class (Goldiamond, 1962). That is, the different stimuli came to evoke the same response (i.e., saying a certain name), which is a defining feature of a functional class. After a functional class had been established with the naming procedure, match-to-sample relations emerged among the stimuli in the class for 2 subjects. Thus, match-to-sample relations among members of functional classes may emerge even when the functional classes are established outside the match-to-sample task, a finding also reported by Sidman, Wynne, Maguire, and Barnes (1989).

At the end of the study, naming was apparently no longer required for 2 of the subjects, because they displayed equivalence properties for new conditional relations without being taught names for any of them. However, it is unclear whether this acquisition of stimulus equivalence was due to the naming procedure that had been in effect during earlier phases of this study or had developed merely as a function of confounding variables such as maturation or practice.

Stronger evidence for the relation between the naming procedure and the emergence of equivalence classes comes from the next-to-last phase of the study. In this phase, subjects were taught to relate new unnamed stimuli to members of an existing functional class of named stimuli via match-to-sample training. Subsequent probes were conducted to examine whether untrained conditional relations would emerge between the named stimuli and the unnamed stimuli, as well as between the different unnamed stimuli. According to Sidman et al. (1989, p. 267), such probes assess the emergence of stimulus equivalence. Trey showed perfect performance on all three different sets of those probes, and Danny and Rory mastered the two last sets of the same probes.

Four directions for further research are suggested by the present study. First, as has already been noted, most subjects' performance was inconsistent over the course of the study, and it was frequently difficult to make firm inferences from the data. Therefore, a more rigorous experimental design should be incorporated into subsequent studies. A multiple baseline design, rather than a reversal design, may be appropriate, because the results of the present study suggest that, for some subjects, the effects of naming may not be reversible.

Second, our findings indicate that the establishment of functional classes through naming facilitates the development of untrained conditional relations, some of which appear to be equivalence relations. Thus, functional and equivalence classes may be related. However, the relation between functional and equivalence classes is a complex issue that is not yet well understood. Sidman et al. (1989) reported what appears to be the first study on this topic. They concluded that, although functional and equivalence classes may be related, they do not constitute the same behavioral process.

Third, in the initial phases of the study, subjects required hundreds of trials to acquire the conditional discriminations that were trained. Therefore, a more efficient training procedure should be sought. Possible improvements over the procedure we used are suggested by Schilmoeller and Etzel (1977) and Saunders and Spradlin (1989).

Finally, it is unclear to what extent the results from the present investigation can be replicated in other populations. Autistic children have unusual learning characteristics that may set them apart from other populations (e.g., Lovaas, Koegel, & Schreibman, 1979). On the other hand, the study of autistic children offers some advantages. In particular, autistic children may be less likely to demonstrate stimulus equivalence, creating opportunities to establish it. Moreover, if autistic children frequently fail to develop stimulus equivalence, remediating their failure may be an important aspect of treatment. Further, in previous research, principles for establishing behavior in autistic children have been similar to those of other organisms (Lovaas & Smith, 1989) despite the extreme behavioral problems that these children present; thus, findings in this population may generalize to other populations.

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. rev.). Washington, DC: Author.

D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus appella*) and pigeons (*Columba livia*). *Journal of the Experimental Analysis of Behavior,* **44,** 35-47.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior,* **46,** 243-257.

Dunn, L. M. (1981). *Peabody Picture Vocabulary Test—Revised.* Circle River, MI: American Guidance Service.

Gardner, H. (1987). *The mind's new science.* New York: Basic Books.

Goldiamond, I. (1962). Perception. In A. J. Bachrach (Ed.), *Experimental foundations of clinical psychology* (pp. 280-340). New York: Basic Books.

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior,* **51,** 385-392.

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalence in children. *Journal of the Experimental Analysis of Behavior,* **41,** 251-266.

Leiter, R. G. (1959). Part one of the manual for the 1948 revision of the Leiter International Performance Scale: Evidence of the reliability and validity of the Leiter tests. *Psychology Service Center Journal,* **11,** 1-72.

Lipkens, R., Kop, P. F. M., & Matthijs, W. (1988). A test of symmetry and transitivity in the conditional discrimination performances in pigeons. *Journal of the Experimental Analysis of Behavior,* **49,** 395-409.

Lovaas, O. I. (1977). *The autistic child: Language development through behavior modification.* New York: Irvington Publishers.

Lovaas, O. I. (1981). *Teaching developmentally disabled children: The ME book.* Baltimore: University Park Press.

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin,* **86,** 1236-1254.

Lovaas, O. I., & Smith, T. (1988). Intensive behavioral treatment with young autistic children. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 11). New York: Plenum Press.

Lovaas, O. I., & Smith, T. (1989). A comprehensive behavioral theory of autistic children: Paradigm for research and treatment. *Journal of Behavior Therapy and Experimental Psychiatry,* **20,** 17-29.

Mackay, H. A., & Sidman, M. (1984). Teaching new behavior via equivalence relations. In P. H. Brooks, R. Sperber, & C. McCauley (Eds.), *Learning and cognition in the mentally retarded* (pp. 493-513). Hillsdale, NJ: Erlbaum.

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.

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

Saunders, K. J., & Spradlin, J. E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior,* **52,** 1-12.

Schilmoeller, K. J., & Etzel, B. C. (1977). An experimental analysis of criterion-related and noncriterion-related cues in "errorless" stimulus control procedures. In B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New developments in behavioral research* (pp. 317-347). Hillsdale, NJ: Erlbaum.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research,* **14,** 5-13.

Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalence in severe retardation. *American Journal of Mental Deficiency,* **77,** 515-523.

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior,* **37,** 23-44.

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.

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 classes. *Journal of the Experimental Analysis of Behavior,* **52,** 261-274.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review,* **57,** 193-216.

Sparrow, S. S., Balla, D. A., & Cichetti, D. V. (1984). *Vineland Adaptive Behavior Scale.* Circle Pines, MN: American Guidance Service, Inc.

Vaughan, W., Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes,* **14,** 36-42.

Received April 4, 1990
Final acceptance December 10, 1991