JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1983, 40, 69-78 NUMBER 1 (JULY)

THE DEVELOPMENT OF DERIVED STIMULUS RELATIONS THROUGH TRAINING IN ARBITRARY-MATCHING SEQUENCES

BRUCE WETHERBY, GEORGE R. KARLAN, AND JOSEPH E. SPRADLIN

JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE, PURDUE UNIVERSITY, AND UNIVERSITY OF KANSAS

Five-year-old children were taught three-stage sequences of arbitrary matching: A-C, B-C, A-D; A-C, B-D, B-C; or A-C, A-D, B-C. Each stage refers to a sample-comparison relation between stimuli. Unreinforced test probes revealed untrained arbitrary matches (B-D, A-D, and B-D, respectively), derivable by substitution of stimuli with a common sample or comparison function. Additional probes revealed further untrained sample-comparison relations derivable by substitution and identity, including the commuted relations D-B, D-A, and D-B, respectively. These processes may have relevance to conceptual and verbal behavior.

Key words: stimulus control, arbitrary matching to sample, derived stimulus relations, stimulus equivalence, stimulus classes, substitutability, language, children

Sidman (1971) and Sidman and Cresson (1973) demonstrated that severely retarded adolescents who were taught certain auditory-visual equivalences also exhibited derived visual equivalences that had not been trained (see Figure 1). At the beginning of the experiment, subjects could match comparison pictures (D) to sample spoken words (A) (*A-D matching*). Subjects were then taught to match printed words (C) to the same spoken words (A) (*A-C matching*). Without any additional training, subjects matched the printed words to their pictures (C-D) and vice versa (D-C).

This research is important as it demonstrated that if different visual stimuli were selected in response to the same spoken words, they then functioned as equivalent stimuli and could be matched to each other. However, these studies raised further questions. First, does the stimulus equivalence found in Sidman's study occur only between printed words and pictures? Second, would stimulus equivalence occur if arbitrary stimuli were used rather than stimuli from the subjects' natural language? Third, would stimulus equivalence occur if the matching tasks were visual-visual rather than auditory-visual?

Sidman, Cresson, and Willson-Morris (1974) demonstrated that stimulus equivalence could be established in severely retarded adolescents between spoken words (A) and pictures (D) if printed words (C) were matched previously with the spoken words and pictures. This para-

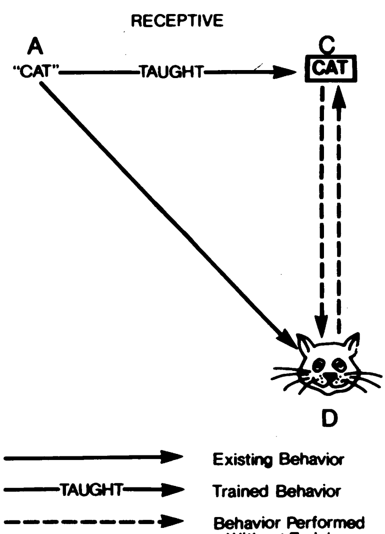

Fig. 1. The Sidman (1971) and Sidman and Cresson (1973) paradigm and results.

This research was supported by NICHD Grant HD 00870 and NICHD Research Service Award HD 07066 to the Kansas Center for Research on Mental Retardation and Human Development, University of Kansas, and NICHD Grant HD 00973 to the John F. Kennedy Center for Research on Education and Human Development, George Peabody College for Teachers, Vanderbilt University. Portions of this report were presented at the 1975 meeting of the American Psychological Association in Chicago, Illinois. The authors wish to express their appreciation to Drs. Howard Goldstein, Evalyn F. Segal, and Eleanor Criswell for providing critical comments of earlier drafts. Send reprint requests to Bruce Wetherby, Department of Psychology, John F. Kennedy Institute, Johns Hopkins Medical Institutions, Baltimore, Maryland 21205.

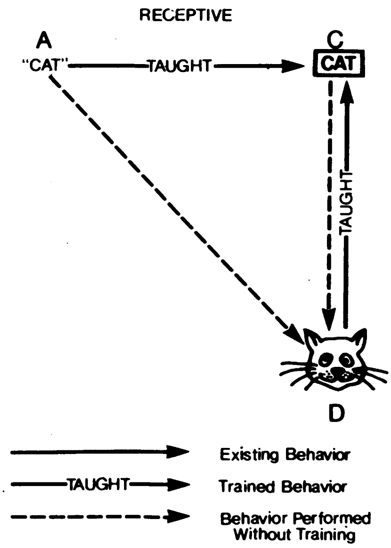

Fig. 2. The Sidman, Cresson, and Willson-Morris (1974) paradigm and results.

digm, shown in Figure 2, indicated that untrained matching performances could be developed between stimulus relations other than just printed words and pictures. The critical component of the work of Sidman and his associates appears to be relational training of either a single sample with two sets of comparisons or a single comparison with two sets of samples.

Spradlin and Dixon (1976) and Dixon and Spradlin (1976) were also interested in isolating the conditions responsible for the development of untrained sample-comparison relations. They demonstrated the development of untrained performances with arbitrary rather than natural-language stimuli. Their paradigm, shown in Figure 3, was used with moderately retarded adolescents. Results indicated that arbitrary auditory stimuli (A) could be brought to control the selection of arbitrary visual stimuli (D) as a function of establishing the matching relations A-C, C-D, and D-C. This work replicated the cross-modal findings of Sidman and associates.

More recently, Stromer and Osborne (1982) addressed the question of derived visual-visual matching. Twelve developmentally delayed adolescents were taught the visual-to-visual arbitrary matches shown in Figure 4. Results replicated the earlier demonstrations of the development of an untrained arbitrary matching relation and also indicated that the sample

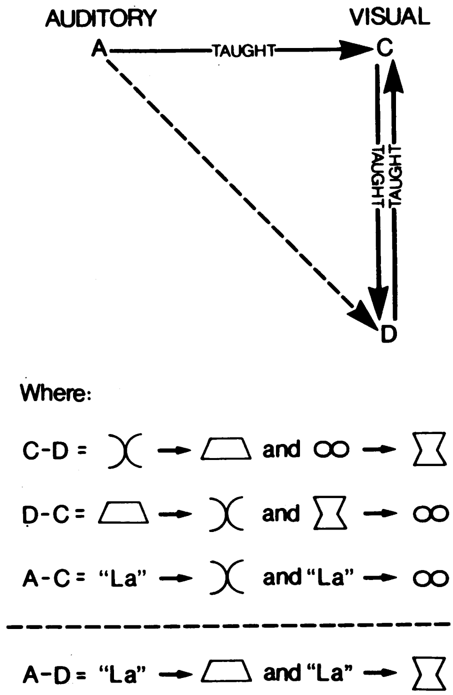

Fig. 3. The Spradlin and Dixon (1976) and Dixon and Spradlin (1976) paradigm and results.

and comparison functions of stimuli in both trained and derived relations were reversible (i.e., commutable). This work demonstrated that untrained arbitrary matching relations can be derived solely within the visual modality and are not confined to cross-modal stimulus relations. It also indicates that the functions of the terms in a stimulus relation can be reversed.

Sidman and Tailby (1982) recently added to our understanding of stimulus equivalence by demonstrating that classes of equivalent stimuli (i.e., stimuli that can be matched to one another in both directions) could be expanded, thereby multiplying the number of untrained (derived) stimulus relations subjects display. Eight normal children were taught the matching relations shown in Figure 5 and were tested on the remainnig relations. Training A-C and B-C matches established an ABC class of equivalent stimuli, as demonstrated by tests for derived relations A-B and C-B. Training A-D matches expanded the stimulus-equivalence class to ABCD, resulting in the appearance of additional derived stimulus relations B-D, D-B, C-D, and D-C. Note that the establishment of B-D and D-B relations did not follow the pat-

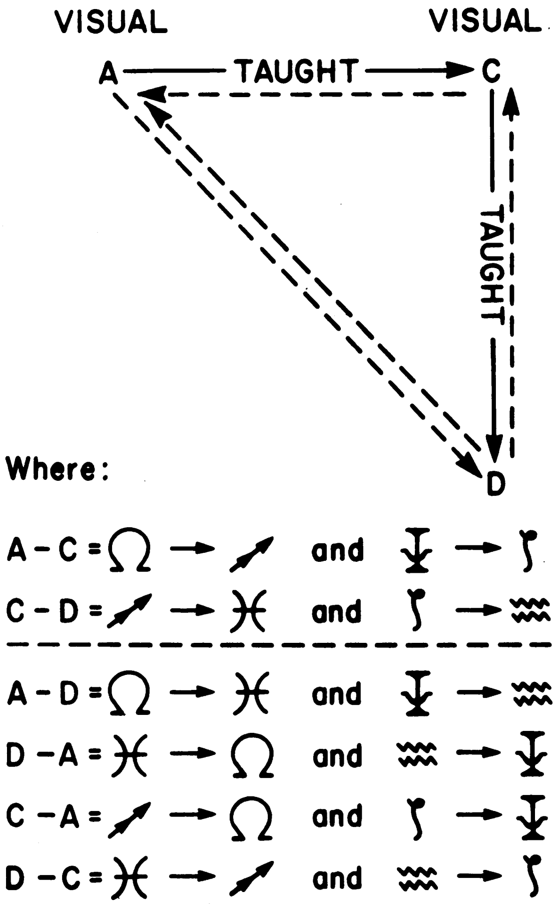

Fig. 4. The Stromer and Osborne (1982) paradigm and results.

tern of earlier experiments, in which stimulus equivalence between comparisons was established by matching one sample with both comparisons or stimulus equivalence between samples was established by matching one comparison with both samples.

The Sidman and Tailby (1982) experiment built on the work of Spradlin, Cotter, and Baxley (1973), who taught three moderately retarded adolescents the first three arbitrary-matching tasks (A-C, B-C, A-D) shown in Figure 6. Subjects were then tested on a fourth, derived, matching task (B-D). All subjects matched correctly in the B-D task. This finding indicates that derived matching relations may be established simply by matching two sets of comparison stimuli to a common set of samples and then matching one of the comparison sets to a new set of samples. Sidman and Tailby (1982, p. 8) suggested that this form of transfer

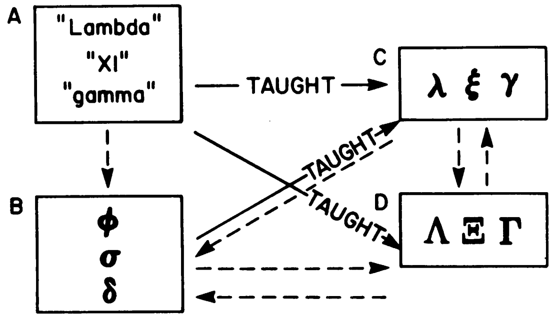

Fig. 5. The Sidman and Tailby (1982) paradigm and results.

implied the development of four sets of stimulus-equivalence classes. Thus, one does not have to train a common sample or comparison function in order to obtain a novel matching relation, as one might have inferred from the previous work of Sidman (1971), Sidman and Cresson (1973), Sidman, Cresson, and Willson-Morris (1974), Spradlin and Dixon (1976), Dixon and Spradlin (1976), and Stromer and Osborne (1982).

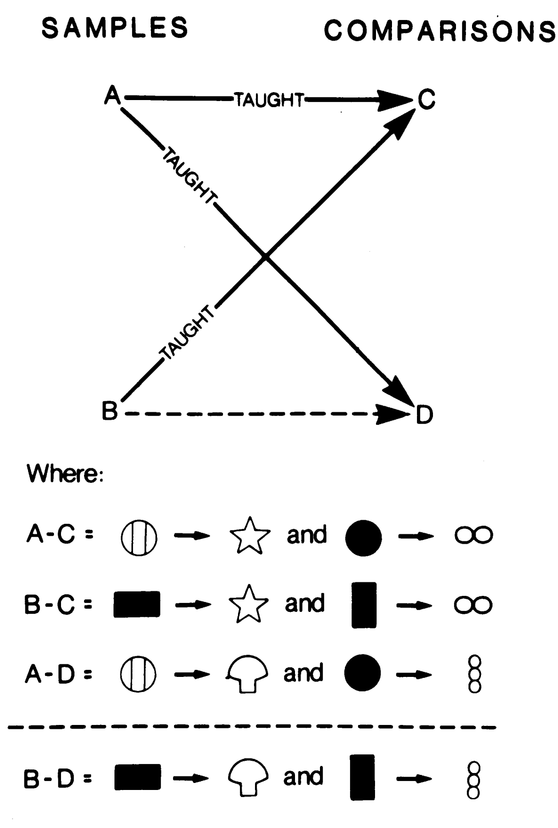

Fig. 6. The Spradlin, Cotter, and Baxley (1973) paradigm and results.

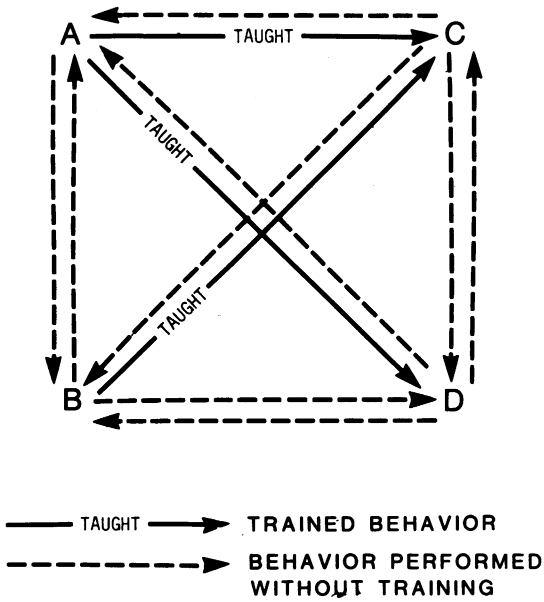

Fig. 7. Diagram of current paradigm.

The purpose of the present experiment was to expand the original Spradlin et al. (1973) work by testing for the existence of all the permutations of both three- and four-class matching-to-sample relations not tested in the previous work. These relations are shown in Figure 7. In particular, relations derived by substitution of stimuli matched to a common sample or comparison and relations derived by commutation of the sample-comparison functions were studied. This work extends Sidman and Tailby's (1982) by testing for the derivation of all possible equivalence relations developed among four classes of visual stimuli.

METHOD

Experimental and Control Subjects

Seven children (four male and three female) from the Edna A. Hill Preschool at the University of Kansas served in the experimental group. Their ages ranged from 4 years, 9 months to 5 years, 6 months. These children were among 13 children originally selected for the study. Of the six children who did not complete the study, three failed to learn the initial matching task, one was excluded because of insufficient time to complete the experiment, another for being uncooperative, and the sixth because of an experimenter error.

Four children (two male and two female) from the same preschool served as control subjects. Their ages ranged from 4 years, 10 months to 6 years, 6 months. Two other children, originally selected as controls, were excluded because they failed to learn the initial matching task.

Setting and Apparatus

The experimental space consisted of an experimental room and adjacent control room near the children's classrooms. The control room contained electromechanical programming equipment. The experimental room contained a match-to-sample display cabinet, a few small chairs, a table that held three carnival-type games (ring-toss, bean-bag, and fishing) used as reinforcing activities, and a box containing small toys.

The stimulus-display cabinet concealed a Carousel projector, Davis Model 310 Universal Feeder (dispensing tokens), a door chime, and a buzzer. The child was seated in front of the stimulus display, a screen of translucent Plexiglas, 25.40 by 20.32 cm, onto which slides were rear-projected. The lower 25.40 by 7.62 cm of the display area was divided equally into four comparison panels. Small opaque squares covered the two unused outside panels. The sample stimulus was displayed at child's eye level simultaneously with display of two comparisons below it. A red error light was located above and to the left of the stimulus display, and a dispenser aperture was located to the right through which tokens were dispensed into a clear plastic glass.

Touching a comparison panel closed a microswitch mounted behind the panel. During initial training, correct selections were followed immediately by the chime and delivery of a token, and incorrect selections by the red light and buzzer. After a 5-sec intertrial interval the next set of stimuli was presented. A solenoid-operated shutter blocked light from the projector during the intertrial interval. Maximum trial duration was 30 sec. If the child failed to select a comparison during this time, the program advanced to the intertrial interval and then to the next trial in the sequence.

Experimental Procedures

To avoid potential conflict with the subject's previous exposure to words and pictures,

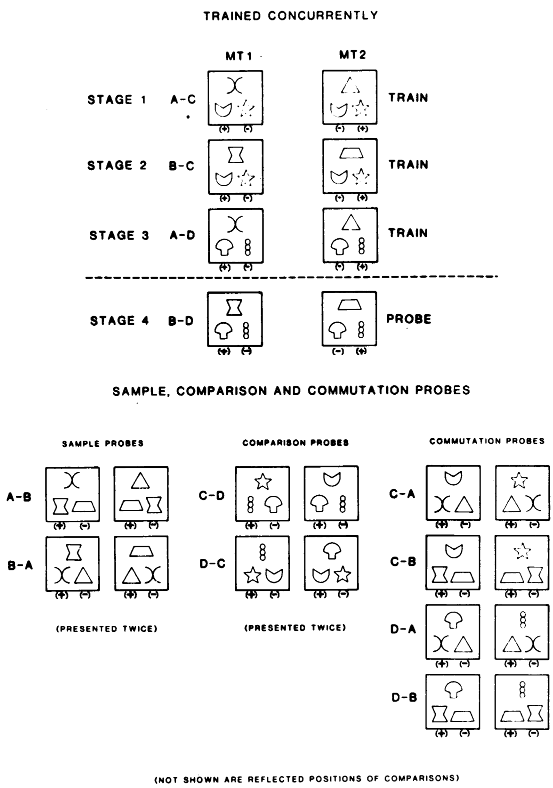

Fig. 8. Stimuli and training sequence used for NA, AB, PG, and TM.

abstract designs were used as visual stimuli, some borrowed from Spradlin et al. (1973) and some new. Figure 8 illustrates the stimuli and training sequence for four experimental subjects (NA, AB, PG, and TM) and the four types of probe trials used for these subjects.

Table 1 lists the order in which arbitrary matches were trained or probed for the seven experimental subjects. Four subjects received the training-and-testing sequence used by Spradlin et al., A-C, B-C, A-D; B-D. Two received the training-testing sequence A-C, A-D, B-C; B-D. One received the sequence A-C, B-D, B-C; A-D.

The subjects were brought into the experimental room one at a time and given a brief orientation. They were told that when a picture appeared in the top (sample) window, they were to pick one of the two bottom (comparison) windows that "went with" the sample. The subjects were further instructed to try to get as many correct as possible and that for each correct selection they would hear a bell

Table 1

Training sequence for experimental and control subjects.

Subject	*Training Sequence*		
Experimental			
NA, AB, PG, TM	Stage 1	A-C	Train
	Stage 2	B-C	Train
	Stage 3	A-D	Train
	Stage 4	B-D	Probe
CB, LV	Stage 1	A-C	Train
	Stage 2	A-D	Train
	Stage 3	B-C	Train
	Stage 4	B-D	Probe
MM	Stage 1	A-C	Train
	Stage 2	B-D	Train
	Stage 3	B-C	Train
	Stage 4	A-D	Probe
Control			
HL, SR, GG, BH	Stage 1	A-C	Train
	Stage 4	B-D	Probe

and a token would be delivered into the receptacle to the right of the panel. Sessions were conducted daily, Monday through Thursday.

Stage 1. The first task for all subjects was the set of A-C arbitrary matches shown in Figure 8, A1 (C1*, C2), A2 (C1, C2*). There were 10 trials per session with each sample, for a total of 20 trials. Selecting comparison C1 was rewarded given sample A1, and selecting C2 was rewarded given A2. When the subjects reached at least 90% correct in each of three consecutive sessions, they advanced to the second stage. Subjects were not allowed to count the tokens received at the end of the session, but were told they had done well and that the tokens could be exchanged for an opportunity to play one of the games and obtain a small toy. These interactions took place after every session.

Stage 2. Sessions included 20 Stage-1 trials randomly intermixed with 20 Stage-2 trials per session. Stage 2 tasks were B1 (C1*, C2), B2 (C1, C2*) for NA, AB, PG, and TM; A1 (D1*, D2), A2 (D1, D2*) for CB and LV; and B1 (D1*, D2), B2 (D1, D2*) for MM. Three sessions at 90% correct advanced the child to the next stage.

Stage 3. This stage consisted of 44 trials per session: 20 Stage-3 trials intermixed with 12 Stage-1 and Stage-2 trials. The Stage-3 tasks were A1 (D1*, D2), A2 (D1, D2*) for NA, AB, PG, and TM; and B1 (C1*, C2), B2 (C1, C2*) for CB, LV, and MM. Again, criterion was 90% correct over all trials for each of three days.

Probe-preparation condition (variable-ratio reinforcement). After completion of initial training, subjects were prepared for unreinforced probes by being given sessions of 36 trials (12 from each of the first three stages), on a random 1/3 of which selections received no feedback. Before each session, the subject was told that the apparatus would be silent for some of the slides, but that tokens would be given at the end of the session for the correct silent trials. After each session, the experimenter dropped a handful of about five or six tokens all at once into the token glass. As before, the subjects were not allowed to count the tokens but were simply told that they had done well. Any questions concerning the correctness of selections to the silent slides were either ignored or answered with some variation of "I can't tell you." When the subjects met the usual criterion of at least 90% correct for three successive sessions, they advanced to the probe stage.

Stage 4 (Probe Stage). There were four types of probes. Because of time restrictions, however, not all children received all types. All children were given the first type of probe, which consisted of 12 unreinforced substitution-derived trials (B-D or A-D, six trials with each of two derived matching tasks). These probes were intermixed with 24 reinforced training trials, eight from each of the first three stages, for a total of 36 trials per probe session. The parentheses below the Stage-4 arrays in Figure 8 indicate the selections that would be correct in Stage-4 probe trials if the three previous stages developed control of arbitrary matching by the derived sample-comparison relations.

After the Stage-4 probes, further derived sample-comparison relations were tested in probe trials. These probes employed stimuli both of which had functioned as samples or both of which had functioned as comparisons during initial training but now served (commutably) in a sample-comparison relation. These probes were presented in different sequences and were never intermixed with each other within a session. Sessions contained 40 trials, 16 probes intermixed among 24 training trials from the first three stages. *Sample* probes consisted of four presentations each of A1 (B1*, B2), A2 (B1, B2*), B1 (A1*, A2), and B2 (A1, A2*). *Comparison* probes displayed C1 (D1*, D2), C2 (D1, D2*), D1 (C1*, C2), and D2 (C1, D2*) four times each. *Commutation* probes displayed sample and comparison stimuli for each stage with commuted functions (former samples were now comparisons, former comparisons were now samples). When given, commutation probes were presented last.

Control Procedures

Without training on the first stages of a substitution paradigm, the probability that a subject would match correctly on unreinforced fourth-stage probes and maintain this performance without reinforcement should be very low. This hypothesis was tested by giving control subjects fourth-stage probes after they had been taught only first-stage (A-C) arbitrary matching.

Control subjects were given the same first-stage training as experimental subjects. Three control subjects were trained with the same stimuli as those used for NA, AB, PG, and TM. When the criterion of 90% was reached for no less than two consecutive days, the subject advanced to a probe-preparation (variable-ratio) condition analogous to that for experimental subjects. There were 24 training trials per session (all from the first stage) with eight of the trials unreinforced. The probe-preparation condition was followed by eight fourth-stage probes intermixed with 16 first-stage trials.

RESULTS

Accuracy scores are shown in Figures 9 and 10 for the seven experimental subjects. Generally, subjects had the most difficulty learning the first-stage matches. Only one subject, MM, was correct on 90% or more of the trials during the first session of the first stage. Accuracy for the remaining subjects increased gradually over six to twelve sessions.

Six of seven subjects reached criterion on the second-stage tasks within three to four sessions. Only one subject, NA, started out below 90%. His performance subsequently improved as in the first stage.

Learning of third-stage matching for all subjects was comparable to that during the

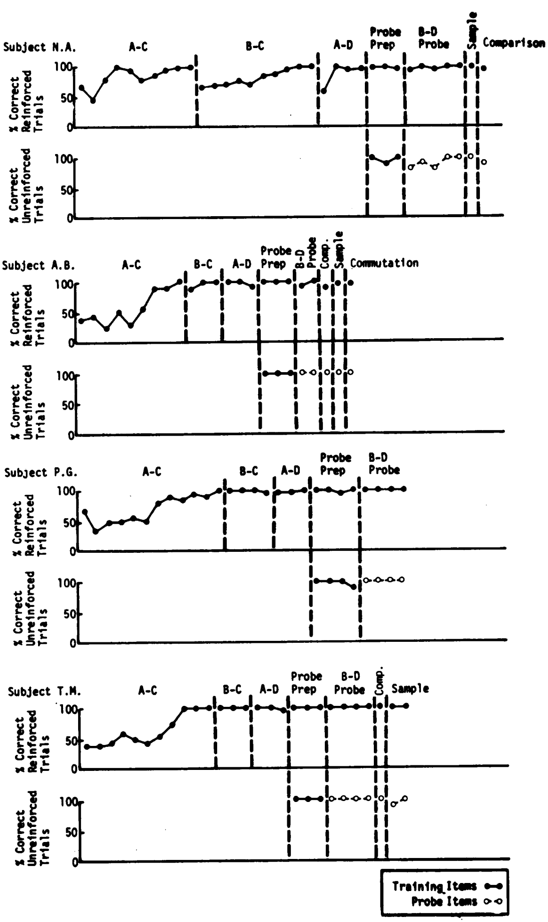

Fig. 9. Experimental training and probe results for NA, AB, PG, and TM.

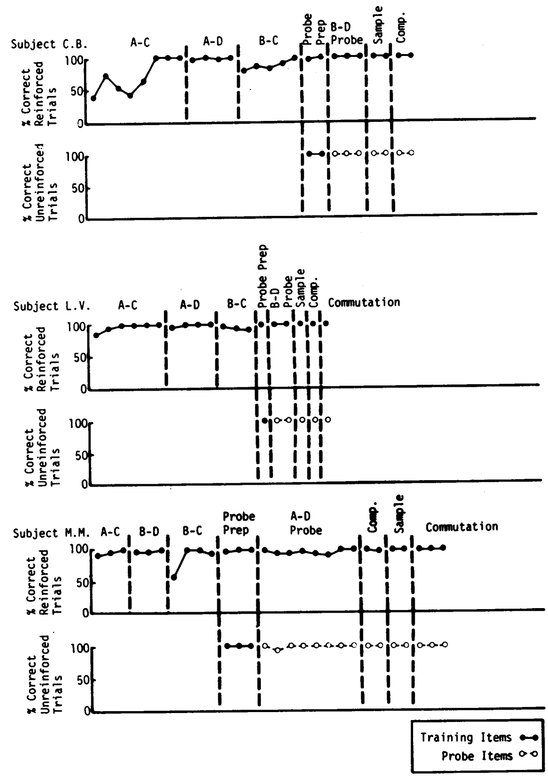

Fig. 10. Experimental training and probe results for CB, LV, and MM.

second stage. All but three subjects (NA, MM, and CB) performed at 90% or better during the first session, and all subjects reached criterion within five sessions. Once learning was complete for the third stage, all subjects continued to match perfectly or nearly perfectly when a random third of the trials were unreinforced during the probe-preparation condition.

Scores for the four control subjects are shown in Figure 11. In general, acquisition of the first matching tasks was comparable to that of the experimental subjects. Training was extended beyond criterion for Subject BH because of experimenter error. Performance by control subjects during the probe-preparation condition was similar to that of experimental subjects; no subject matched below 92%.

On fourth-stage substitution probe trials (B-D for NA, AB, PG, TM, CB, and LV; and A-D for MM), all experimental subjects performed perfectly or nearly perfectly. Only one subject, NA, started out below 100% correct, and his performance increased to 100% within three more sessions. The remaining subjects either continued their perfect performance on fourth-stage trials or varied slightly from this level during the remainder of these sessions.

In contrast, control subjects selected correctly on a considerably smaller percentage of fourth-stage probes. Three of four control subjects matched over 50% of the fourth-stage trials correctly during the first session of this condition but did less well in the subsequent one or two sessions. The remaining subject (SR) scored considerably below 50% in the first session and did not attain 50% correct in the subsequent three sessions.

Arbitrary matching on the three remaining types of probes (sample, comparison, and commutation) for those experimental subjects who were tested on them was perfect or nearly perfect in the predicted directions.

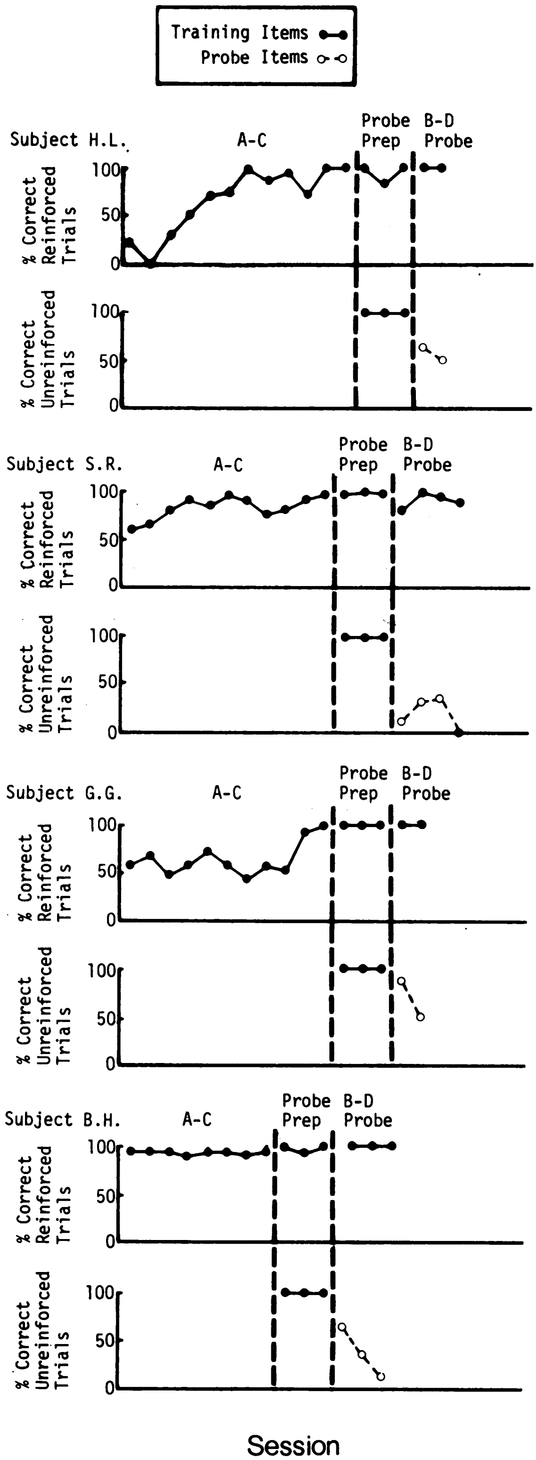

Fig. 11. Control training and probe results for HL, SR, GG, and BH.

DISCUSSION

The results of the current research replicate systematically the findings of Spradlin et al. (1973) and Sidman and Tailby (1982). When four subjects were taught to select a common comparison stimulus in the presence of two different sample stimuli and then to select a new comparison stimulus in the presence of one of the samples, they selected the same new comparison when the other sample was presented. When two subjects were taught to select two different comparison stimuli in the presence of a single sample stimulus and then to select one of those comparison stimuli in the presence of a new sample stimulus, they also selected the other comparison when the new sample was presented. The seventh experimental subject also matched appropriately in the fourth stage after a slightly different sequence of training. Essentially, each of these procedures trained the same three arbitrary-matching relations but in different orders. Apparently, order of training is not critical to the development of matching by derived sample-comparison relations.

In the original Spradlin et al. study, two of the comparison stimuli established as substitutable were a vertical chain-like figure and an infinity sign. Both these figures consist of joined curved enclosed lines (see Figure 6), and their physical similarity could have been a factor in the derived stimulus control obtained in that study. In the current study, a crescent was used in place of the infinity sign, thus eliminating physical similarity of the substitutable comparisons. Even so, the current study replicated Spradlin et al.'s results.

Four control subjects who were given training on only one arbitrary-matching task and then given the first probe test did not respond to the probe stimuli in the same consistent way as the experimental subjects. Two subjects, SR and BH, matched correctly on far less than the 50% of trials to be expected if the subject matched at random (See Figure 11). This is not surprising. When a subject is presented a two-choice arbitrary-matching task with no feedback, one of three patterns of stimulus control often emerges. First, the subject may adopt a position preference. Second, the subject may simply select the same comparison on each trial. Either of these types of stimulus control would result in a matching score of 50% over trials. Third, a subject who has been trained on prior matching tasks may consistently select one comparison in the presence of one sample and the other comparison in the presence of the second sample. This would result in matching scores approaching either 100% or 0%. This third type of control appears to apply to Subjects SR and BH. The fact that no control subject matched as the experimental subjects did supports the conclusion of Spradlin and Dixon (1976) and Dixon and Spradlin (1976) that stimulus control by derived sample-comparison relations requires training matching relations in sequences that foster substitutability.

The results of this study as well as those of the earlier stimulus-equivalence studies (Sidman, 1971; Sidman & Cresson, 1973; Sidman et al., 1974; Sidman & Tailby, 1982; Spradlin et al., 1973; Stromer & Osborne, 1982) are compatible with Spradlin and Dixon (1976) and Dixon and Spradlin's (1976) notion that substitutability is a primary determinant of the development of derived stimulus relations. In every case, original training in these studies involved establishing two or more stimuli as substitutable for each other within a specific sample or comparison function. For example, in Sidman's original (1971) training, pictures and printed words were substitutable for each other as comparisons in the presence of auditory sample words. When "cat" was spoken, for instance, selecting the comparison picture of the cat or the comparison word *cat* was reinforced. The test for the development of control by a derived (and commutable) sample-comparison relation involved matching pictures to words and words to pictures.

Probes determined whether substitutable sample stimuli could function in a derived sample-comparison relation. All six subjects who were given this sample probe matched correctly on over 88% of probe trials. Five of the six matched correctly 100% of the time.

Another probe presented substitutable comparison stimuli in a derived sample-comparison relation. All six subjects given this test matched correctly 100% of the time.

The final probe determined whether commuted stimulus relations would control matching. This commuted sample-comparison probe was given to three subjects. All matched correctly 100% of the time.

Figure 7 summarizes the derived stimulus relations obtained without explicit training in this experiment. Compared to the derived matching relations established in Spradlin et al. (see Figure 6), these data suggest that the potential range of derived stimulus relations that can control arbitrary matching without explicit training is quite diverse.

The concept of derived sample-comparison relations based on substitutability may also apply to the development of complex conceptual performances. For example, a child who on different occasions has been told to label tulips and daisies as *flowers* may come to treat both types of flowers equivalently. Therefore, if the child learns to provide the label *plant* for one flower, that label may also generalize to the other flower. Additional contingencies will also teach the child that although tulips and daisies may control a common response in one context, their specific labels are not freely substitutable. If they were, of course, there would not be two different labels.

This speculation involves very complex social exchanges, whereas the studies cited above involved the performance of human subjects on laboratory match-to-sample tasks. Nevertheless, it is unlikely that the development of derived stimulus relations by substitution is limited to contrived tasks.

REFERENCES

Dixon, M., & Spradlin, J. Establishing stimulus equivalences among retarded adolescents. *Journal of Experimental Child Psychology*, 1976, **21**, 144-164.

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

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

Sidman, M., Cresson, O., Jr., & Willson-Morris, M. Acquisition of matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior*, 1974, **22**, 261-273.

Sidman, M., & Tailby, W. Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 1982, **37**, 5-22.

Spradlin, J. E., Cotter, V. W., & Baxley, N. Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency,* 1973, **77,** 556-566.

Spradlin, J. E., & Dixon, M. H. Establishing conditional discriminations without direct training: Stimulus classes and labels. *American Journal of Mental Deficiency,* 1976, **80,** 555-561.

Stromer, R., & Osborne, J. G. Control of adolescents' arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior,* 1982, **37,** 329-348.

Received September 15, 1980
Final acceptance March 11, 1983