JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1984, 41, 251-266 NUMBER 3 (MAY)

THE FORMATION OF VISUAL STIMULUS EQUIVALENCES IN CHILDREN

RONALD M. LAZAR, DEBORAH DAVIS-LANG, AND LISETTE SANCHEZ

QUEENS COLLEGE OF THE CITY UNIVERSITY OF NEW YORK

Four normal children were presented a series of matching-to-sample tasks, using five sets of visual stimuli designated A, B, C, D, and E. Stimulus equivalences were established by matching stimuli from one set to those from another set. Each set consisted of three stimuli, so matching set A to set D meant that each stimulus in set A served as a sample with all three stimuli in set D as comparisons. Subjects were first taught AD and DC matching and were then able to perform AC/CA matching without additional training. After ED was taught directly, CE/EC and AE/EA performances emerged. Following CB training, three new equivalences were demonstrated: AB/BA, EB/BE, and DB/BD. Oral naming of each stimulus showed that subjects had not assigned a common label to stimuli in the same class, indicating that naming is not necessary for the formation of stimulus equivalences. The absence of response mediation suggests that matching to sample can form direct stimulus-stimulus associations. The data also provide support for the notion that generative performances are outcomes of existing stimulus-control relationships.

Key words: stimulus equivalences, stimulus classes, matching to sample, generative behavior, key press, children

In his discussion of "arbitrariness" in human language, Hockett (1960) noted that the relationship between a meaningful linguistic element and its denotation is independent of any physical resemblance between the two. For example, there is little about a dog that requires it be called "dog," or its written name to be spelled *dog*, yet we treat the words as "representations" of the animal. Although there are instances in most languages in which there are common acoustic properties between a spoken work and its referent (onomatopoeia), their occurrence is quite rare (see Taylor & Taylor, 1965).

One behavioral approach to the study of such arbitrary relationships between stimuli has been within the framework of matching-to-sample procedures. In arbitrary matching, subjects are presented sample and comparison stimuli that are all physically different (see Cumming & Berryman, 1965). The correct comparison stimulus for a given sample is designated on a random basis; in principle, each of the comparison stimuli has the potential to be the correct choice. Arbitrary matching tasks have also been described as "symbolic" or "nonidentity" performances (Carter & Eckerman, 1975; Santi, 1978).

Experiments with arbitrary matching to sample with human subjects have shown that the relationship between samples and their correct comparisons can be more complex than the formal correspondence between the matching procedure and a conditional discrimination (if sample A1, select comparison B1; if sample A2, select comparison B2). For example, after learning to match B (pictures) to A (dictated words) and C (printed words) to A (dictated words), retarded subjects were then able to match B to C, and C to B without explicit training (Sidman, 1971; Sidman & Cresson, 1973). Oral naming of the printed words also emerged in the absence of direct teaching. It has been suggested that because printed words and pictures were interchangeable with corresponding dictated words, each

This research was supported by Grant PSC-CUNY 13496 from the Research Foundation of the City University of New York. Some of these data were presented at the Eastern Psychological Association meeting, Baltimore, 1982. Reprints may be obtained from Ronald M. Lazar, Department of Psychology, Queens College, Flushing, New York 11367.

picture/printed-word/dictated-word triad represented a class of equivalent stimuli (Lazar, 1977).

The formation of stimulus equivalences with matching to sample has been demonstrated with retarded adolescents (Lazar & Mackay, in press; Mackay & Sidman, in press; Sidman, Cresson, & Willson-Morris, 1974; Stromer & Osborne, 1982), younger retarded children (Spradlin, Cotter, & Baxley, 1973; Spradlin & Dixon, 1976), and normal children (Lazar & Kotlarchyk, 1984).

Sidman and Tailby (1982) have extended the analysis of stimulus equivalences by demonstrating that if a new stimulus is made equivalent to one member of an already existing class, that new stimulus can become equivalent to all members of that class. Their experiment presented arbitrary-matching tasks in which child subjects were taught to match four sets of stimuli (designated A, B, C, and D) in a systematic sequence. In this research, "set" referred to stimuli that served common, *experimenter-defined* functions as samples or comparisons during training, and not to the stimulus control relations demonstrated by subjects as a consequence of training. For example, when a subject was taught AB matching, each of the three stimuli in Set A served as the sample on different trials, and all three stimuli in Set B were presented together as comparisons. Given A1 as the sample, the correct choice was B1; when A2 was the sample, B2 was correct, and so on. Subjects were first taught AB and AC matching, the establishing conditions for three classes (a class being, for example, A1, B1, and C1). The question was whether establishing an additional performance, DC, would produce expanded classes, each consisting of four members (e.g., A3, B3, C3, and D3). The crucial evidence for the enlarged classes was derived from the subjects' ability to perform BD and DB matching without additional training, inasmuch as these tasks required equivalence among stimuli from all four sets. Oral naming of stimuli in Sets B, C, and D, consistent with the dictated names in Set A, emerged for all subjects who also demonstrated matching-to-sample equivalences.

The major aim of the present experiment was to explore further the nature of class expansion. In particular, would subjects demonstrate the enlargement of class membership with matching tasks involving purely visual stimuli? Previous studies have consistently shown the relationship between dictated names as members of stimulus classes and the emergence of the same labels for visual stimuli that were class members (Lazar & Mackay, in press; Sidman, 1971; Sidman & Cresson, 1973; Sidman et al., 1974; Sidman & Tailby, 1982). If naming serves an important mediating function in the establishment and expansion of equivalence classes, then the presentation of purely visual matching tasks would require subjects to assign their own names because none are provided by the experimenter. The purposes of this study were therefore twofold: first, to determine whether stimulus classes could be established and progressively enlarged in the absence of dictated names as class members; and second, to ascertain whether subjects who demonstrated equivalences would supply their own mediating labels.

Spradlin et al. (1973), Stromer and Osborne (1982), and Wetherby, Karlan, and Spradlin (1983) have previously shown that subjects could demonstrate purely visual classes. Our research sought to extend these earlier findings by establishing the conditions for larger classes, determining the role of naming, and testing all the requisite, intermediate equivalences.

METHOD

Subjects

Four normal children, all male except Subject A.W., participated in the experiment. Their ages (years-months) at the time they completed training and underwent their first tests for equivalence relations were: 5-9 (G.H.), 5-3 (A.W.), 6-1 (C.G.), and 7-2 (M.C.). Subject G.H. attended kindergarten, Subjects A.W. and C.G. were in first grade, and Subject M.C. was in second grade. Each child, recruited by signs posted in the community, came to the laboratory several days a week with a parent or sitter who was paid after each session. The total number of sessions, each 25 to 45 min long, varied because the children needed differing amounts of training to learn and maintain baseline matching performances.

Apparatus, Reinforcers, and Stimuli

The subject sat before a panel containing seven circular windows mounted on a modified teaching machine (Behavioral Controls, Inc., Model SR-400). The windows, each 3.9 cm in diameter, were arranged in a circle of six with the seventh in the middle (see Figure 1). The display diameter, from outer edge to outer edge, was 13.6 cm; the center-to-center distance was 4.9 cm between adjacent windows on the perimeter, and 4.9 cm between the center window and each of the others. Stimuli were mounted on continuous fan-folded paper, and the movement of the paper was synchronized with a paper-tape reader that controlled the contingencies for each trial. Each of the seven windows was coated with a transparent, electrically conductive surface. When the subject touched the window, a capacitance-sensing system operated the solid-state control and recording equipment. Embedded within each window was a layer of liquid crystal that functioned as a shutter, so that stimuli could be made to appear and disappear. In the present experiment, the windows to the left and right of the center window were nonfunctional and remained dark (shaded in Figure 1). Penny reinforcers were delivered by a universal dispenser (Gerbrands, Model 120). The advancement of the program paper and tape reader, the liquid-crystal shutters, the universal dispenser, and the recording equipment were controlled by solid-state equipment.

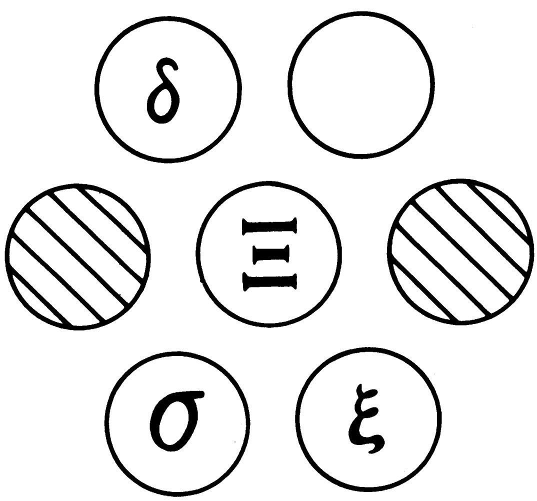

Fig. 1. Schematic representation of the stimulus-response panel, illustrating a choice display from AD matching. Shaded windows were dark throughout the experiment.

The stimuli consisted of 15 upper- and lower-case Greek letters and script Hebrew letters appearing as black line drawings on a white background, each 1.3 cm high and 0.6 to 1.2 cm wide. The reason for using these stimuli was the necessity that stimulus classes be defined solely within the context of our experimental procedures. We presumed that these stimuli were unfamiliar to child subjects and that even if the children assigned labels during the course of experimentation, it was unlikely that they would use the actual letter names. These novel stimuli therefore precluded the need for extensive pretests.

General Procedures

Matching to sample. Each trial began with the presentation of the sample stimulus on the center window. Each sample remained throughout the trial, and trial durations had no limit. The subject had to press the sample window to bring comparison stimuli onto the outer windows. No trial presented more than three comparison stimuli; thus, at least one of the four functional comparison windows was blank on each trial, as indicated in Figure 1. Positions of correct, incorrect, and blank comparison windows varied across trials. For sequences that included only two different sample stimuli, no more than three consecutive trials could occur with the same sample. In addition, all four functional comparison windows were scheduled correct before any window could be correct again.

After the comparison stimuli appeared,

the subject had to press a comparison window. Correct choices were followed by the delivery of a penny, unless the procedure involved extinction conditions (see below). All stimuli disappeared after any choice response, and after a 2-s intertrial interval a new sample began the next trial. No reinforcer was delivered after an incorrect choice. The children kept all the pennies they received.

Any window press during the intertrial interval postponed the next sample for an additional 2 s. A response to an incorrect choice or a blank choice window, or simultaneous comparison-window responses (defined as the onset of the touch on one key within 0.25 s of the touch on another) were treated as errors.

Oral naming. The oral-naming test was administered in two parts. First, the child was presented only with sample stimuli and was asked, "Tell me what this is," or "What is it?" No pressing responses were required. After each oral response, the experimenter pressed a switch to remove the sample stimulus and initiate the intertrial interval, and to advance the program paper for the next trial, 2 s later. No feedback was given regarding the adequacy of the naming response. In the second part of the test, the subject was given an opportunity to label stimuli in the context of actually performing baseline (see below) and equivalence matching tests. The directions were: "Don't touch; just point to them and tell me what it is." Here, too, the experimenter provided no feedback regarding the child's verbal response and advanced the program paper to the next trial. The entire oral-naming test was recorded by the experimenter and by audiotape; later confirmation checks showed perfect correspondence between the tape and our hand-scored records.

Preteaching and Pretesting

All subjects were magazine trained after the delivery of a few pennies. The matching-to-sample procedure was then taught with color (hue) identity matching, using a procedure derived from Sidman, Rauzin, Lazar, Cunningham, Tailby, and Carrigan (1982) and adapted by Lazar (1983).

In the first stage, a single colored disk appeared on one of the windows. We simply told the child, "Touch that." When the subject complied with the instruction, a penny was delivered. On the next trial a different color appeared on a different window, and when the subject pressed the appropriate window, another penny was given. This stage was completed when the subject pressed only the window with a color in a 24-trial set.

Each trial in the next step began with a color on the center (sample) window. By pressing the sample window, the subject produced the same color on one of the four functional outer windows, but no reinforcer. Pressing the outer window then produced a penny and the intertrial interval. At this point, no incorrect comparison stimulus was present; only one comparison window had a stimulus on it at any time. The learning criterion was perfect responding in a 24-trial set.

The next step introduced one incorrect color along with the correct color after the subject pressed the sample window. The three colors we used (red, yellow, blue) appeared as samples and correct choices in varied sequences. After the subject met a criterion of 22 correct in a 24-trial set, we proceeded to the last step. Now there were three choices on each trial: one correct and two incorrect comparison stimuli. The same learning criterion in a set of 24 trials was also in effect for this task.

We then tested the subject's ability to perform 3-choice, generalized identity matching with number stimuli, the Arabic numerals "1," "2," and "3." After this task was mastered at the criterion level, we proceeded to the final pretest task, assessing the subject's competence to perform identity matching with the Greek and Hebrew letters illustrated in Figure 2. On each trial, the sample and correct comparison were the same, and the two incorrect choices were always the other members of the same set (A, B, C, D, or E). This task evaluated the skills of (1) discriminating the experimental stimuli serving sample and comparison functions, and (2) matching these stimuli to themselves (reflexivity), a prerequisite for membership in equivalence classes (see Sidman et al., 1982).

Phase 1: Teach AD and DC Matching

The experiment was designed as a series of teaching/testing cycles, using five sets of stimuli (identified as A, B, C, D, and E as shown in Figure 2). The stimuli have been assigned abstract labels for purposes of exposition. Each box in Figure 2 represents a set of three stimuli; the adjacent numbers indicate that for a given sample stimulus, the correct choice was a letter with the same number. Solid lines indicate performances that were explicitly taught; dashed lines show performances that were tested. Arrows point from sample to comparison stimuli. In some cases, we departed from the experimental designed outlined below because

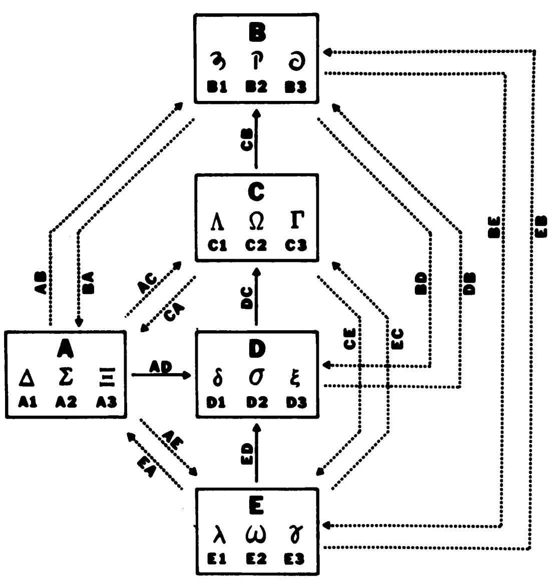

Fig. 2. The equivalence paradigm. Boxes, designated A, B, C, D, and E, represent sets of three stimuli. Beneath each stimulus is an abstract designation for purposes of exposition. The number subscripts indicate that for a given sample stimulus, the correct choice was a stimulus designated by the same subscript. Arrows point from sample to comparison stimuli: Solid arrows represent matching tasks taught directly; broken arrows represent equivalence performances.

subjects were unable to perform a given equivalence task. Deviations from the general plan will be described later, in the context of the results.

All subjects were first taught AD matching (Table 1 and Figure 2) in a sequence of steps, similar to those used by Sidman and Tailby (1982) to establish performance. In Step 1, the subject had to match only two samples, A1 and A2, to the comparison stimuli D1 and D2. The same two comparison letters, one correct and the other incorrect, appeared on every trial. Subjects went through different sets of 24 trials until they scored at least 22 correct in a set. In Step 2, children were taught to match A1 to D1, and A3 to D3 in consecutive sets of 24 trials until the same criterion was met. Because sample A1 and comparison D1 had already been paired with sample A2 and comparison D2, and with sample A3 and comparison D3 in separate sets, it was now possible to advance to step 3 in which A2D2 (i.e., sample A2-correct comparison D2) and A3D3 (i.e., sample A3-correct comparison D3) would be presented. The same learning criterion was in effect. Step 4 was comprised of the six trial types from the first three steps, thereby introducing all three sample-correct comparison relations in single, 24-trial sets. The last teaching segment of AD training

Table 1
Sequence of Teaching and Testing for each Subject

	Subject			
	A.W.	*M.C.*	*C.G.*	*G.H.*
Phase 1: (Teach) 1)	AD	AD	AD	AD
2)	DC	DC	DC	DC
Phase 2: (Test)	AC/CA	AC/CA	AC/CA	AC/CA
Phase 3: (Teach)	ED	ED	ED	ED
Phase 4: (Test) 1)	AE/EA	AE/EA	AE/EA	CE/EC
2)	CE/EC	CE/EC	CE/EC	AE/EA
Phase 5: (Teach)	CB	CB	CB	CB
Phase 6: (Test) 1)	DB/BD	AB/BA	DB/BD	EB/BE
2)	AB/BA	EB/BE	AB/BA	AB/BA
3)	EB/BE	DB/BD	EB/BE	DB/BD

(Step 5) consisted of one correct and two incorrect comparison stimuli, the performance designated in Figure 2. When the children achieved 22 correct in a 24-trial set, they advanced to DC training.

The DC performance was then taught in the same five-step sequence as that for AD. Whereas Set D stimuli had served as comparisons in AD matching, they now became samples with Set C stimuli as comparisons.

Each session in which teaching occurred began with a review of all the learned performances up to that point, and subjects had to meet the acquisition criterion on each previously established task before progressing to new learning. The AD performance, for example, was reviewed and maintained at criterion accuracy each session during the course of DC training.

After subjects met the learning criterion for AD and DC in separate 24-trial sets, both tasks were combined into single, 48-trial sets. To complete this first phase of training, the children had to score at least 45 correct in a 48-trial set, with the additional restrictions that there could be no more than one error on a trial type (i.e., for each sample-correct comparison match) and no more than two errors on AD or DC trials.

Phase 2: Test AC/CA Equivalences

After subjects met the learning criterion for the combined AD and DC tasks, they received versions of the 48-trial set, but *without* programmed consequences following their choices of the comparison stimuli. The first such set was preceded by the following instructions: "You won't get any pennies now, but you will have a chance to get some later." Under conditions of no feedback, the subject's response produced only the intertrial interval. The instructions were always repeated before the presentation of a set in which no reinforcers were delivered. At the end of sessions containing no-feedback trials, subjects were given a comparable number of identity trials (hue-hue, number-number, or shape-shape matching) to ensure a high density of reinforcement for the entire session. When the session was over, the child received additional pennies corresponding to the number of correct responses on the trials without reinforcement.

Following the demonstration of the learning criterion without feedback (at least 45 correct in 48 trials), the formation of stimulus equivalences was assessed by determining whether A and C stimuli, both matched to D stimuli during training, would now be matched to each other (see Figure 2). Because, for example, D1 had been matched to A1 and to C1 during training, would subjects now be able to match A1 with C1? An equivalence would be defined by a subject's ability to match both sample A1 to comparison C1, *and* sample C1 to comparison A1. (See Sidman et al., 1982, for a detailed discussion of the equivalence relation.)

The test consisted of 48 trials: 24 baseline matching trials (four presentations each of the six trial types in AD and DC tasks shown in Table 2) intermixed with four presentations each of the AC/CA equivalence-probe trials shown in Table 3, Phase II. On the AC probes, Set A stimuli (A1, A2, A3) served as samples with Set C stimuli (C1, C2, C3) as comparisons. On CA probes, Set C stimuli were samples and Set A stimuli were comparisons. The entire 48-trial test was presented *without* programmed consequences and was therefore preceded by the no-feedback instructions. Subjects received two 24-trial sets of identity matching after the test. At the end of the session, they received

Table 2
Trial Types for the Matching-to-Sample Baseline

Task	Sample	*Comparisons* Correct	Incorrect	Incorrect
AD	A1	D1	D2	D3
	A2	D2	D3	D1
	A3	D3	D1	D2
DC	D1	C1	C2	C3
	D2	C2	C3	C1
	D3	C3	C1	C2
ED	E1	D1	D2	D3
	E2	D2	D3	D1
	E3	D3	D1	D2
CB	C1	B1	B2	B3
	C2	B2	B3	B1
	C3	B3	B1	B2

Table 3

Trial types for each of the experimental phases (left column) in which equivalence performances (right column) were assessed. The corresponding baseline tasks are noted in the middle column. Each equivalence-probe trial consisted of a sample stimulus and three comparison stimuli, one correct and two incorrect. Equivalence trials were never reinforced.

Experimental Phase	*Baseline*	*Equivalence-Probe Trials*									
				Comparisons					*Comparisons*		
			Sample	*Cor*	*Incor*	*Incor*		*Sample*	*Cor*	*Incor*	*Incor*
II	AD,DC	*AC/CA*									
		AC:	A1	C1	C2	C3	CA:	C1	A1	A2	A3
			A2	C2	C3	C1		C2	A2	A3	A1
			A3	C3	C1	C2		C3	A3	A1	A2
IV	AD,DC,ED	*AE/EA*									
		AE:	A1	E1	E2	E3	EA:	E1	A1	A2	A3
			A2	E2	E3	E1		E2	A2	A3	A1
			A3	E3	E1	E2		E3	A3	A1	A2
		CE/EC									
		CE:	C1	E1	E2	E3	EC:	E1	C1	C2	C3
			C2	E2	E3	E1		E2	C2	C3	C1
			C3	E3	E1	E2		E3	C3	C1	C2
VI	AD,DC,ED,CB	*AB/BA*									
		AB:	A1	B1	B2	B3	BA:	B1	A1	A2	A3
			A2	B2	B3	B1		B2	A2	A3	A1
			A3	B3	B1	B2		B3	A3	A1	A2
		EB/BE									
		EB:	E1	B1	B2	B3	BE:	B1	E1	E2	E3
			E2	B2	B3	B1		B2	E2	E3	E1
			E3	B3	B1	B2		B3	E3	E1	E2
		DB/BD									
		DB:	D1	B1	B2	B3	BD:	B1	D1	D2	D3
			D2	B2	B3	B1		B2	D2	D3	D1
			D3	B3	B1	B2		B3	D3	D1	D2

pennies corresponding to the number of correct responses on baseline-test trials, but received *no* money for their performance on equivalence-probe trials.

Because the AC/CA performance served as a prerequisite for larger classes to be formed later in the experiment, we set the equivalence criterion at 22 correct in 24 trials with no more than one error on a trial type. The high criterion was also established because it was possible to score 78% correct on an equivalence test by responding correctly in the presence of two equivalence-probe samples and randomly to the third.

Phase 3: Teach ED Matching

If subjects demonstrated the equivalence between stimuli in Sets A and C, ED was then added to the matching baseline (see Figure 2 and Tables 1 and 2). Here, stimuli in Set E (E1, E2, E3) served as samples and Set D stimuli were comparisons. The ED performance was taught in a manner similar to that for AD. In the beginning of each session in which ED was being established, 24 trials of AD and DC matches were presented to maintain these performances.

After subjects learned ED (22 correct in 24 trials), AD, DC, and ED were combined into single, 72-trial sets. Each set of 72 trials consisted of eight blocks of the nine trial types in the matching baseline (see Table 2) so that all trial types occurred in equal numbers before any were repeated. The order of presentation varied across blocks. The third phase of the experiment was completed when the child scored 70 correct in a 72-trial set, with no more than one error on a given trial type.

Phase 4: Test AE/EA and EC/CE Equivalences

When the subjects had met the learning criterion on the mixed AD, DC, and ED tasks, they were read the no-feedback instructions and then were given similar 72-trial sets without programmed consequences. As before, we gave subjects the opportunity to earn additional money after a no-feedback set by presenting them with a comparable number of identity-matching trials.

After subjects were able to meet the learning criterion without feedback (70 correct), we tested for the expansion of the classes. If stimuli in E and D had become equivalent, and D stimuli were already members of A-C-D classes, would E stimuli become equivalent to those in A and C as well? Thus the equivalence tasks of interest were AE/EA and CE/EC (see Figure 2).

Subjects A.W., M.C., and C.G. were first presented the AE/EA test. It consisted of 60 trials: 36 baseline matching trials (four presentations each of the nine trial types in the AD, CD, and ED tasks shown in Table 2) mixed with four presentations each of the AE and EA equivalence-probe trials shown in Table 3, Phase IV. On AE trials, Set A stimuli were samples and Set E stimuli were comparisons. Conversely, EA probes consisted of E stimuli as samples and A stimuli as choices. The test was presented without feedback and was therefore preceded by the oral instructions. After the test, subjects received three 24-trial sets of identity matching. When the session was over, they were given reinforcers corresponding to the correct choices on baseline-test trials only.

If equivalences were demonstrated by a subject, the CE/EC test was administered. The test consisted of the same number of AD, DC, and ED baseline trials, but the CE and EC probes replaced those for AE and EA. Consistent with our assessment procedure, no feedback was given.

To determine whether there was a difference if we first measured AE/EA or CE/EC performances, the fourth subject, G.H., was given the two kinds of tests in reversed order (first CE/EC, and then AE/EA) as depicted in Table 1.

Phase 5: Teach CB Matching

CB matching was taught after testing had shown equivalences as a consequence of the AD, DC, and ED baseline. In this new task, Set C stimuli were now samples with Set B stimuli as comparisons (see Figure 2 and Table 2). The CB match was trained in the standard five-step sequence. At the start of a session, the AD, DC, and ED tasks were maintained by requiring 22 correct in 24 trials before commencing with CB training.

After a subject met the learning criterion on CB (22 correct in 24 trials), AD, DC, ED, and CB were combined into 72-trial sets. Each was comprised of six blocks of 12 trial types (Table 2) so that all trial types were repeated equally often before any were presented again. The acquisition criterion was 70 correct in a single set.

Phase 6: Test AB/BA, EB/BE, and DB/BD Equivalences

Following the attainment of the 70-correct performance on an AD, DC, ED, and CB set, subjects were required to meet the same criterion under no-feedback conditions. The children were then given the final series of stimulus-class tests. We wanted to determine whether establishing an equivalence between C and B would extend each class by an additional number, indicated by a subject's ability to match appropriate stimuli in Set B with those in Set A (AB/BA), Set E (EB/BE), and Set D (DB/BD).

Each equivalence test consisted of 72 trials: 48 baseline matching trials (four presentations each of the 12 trial types shown in Table 2) interspersed with four presentations each of AB/BA, EB/BE, or DB/BD (see Table 3). Table 1 also shows that subjects received the three tests in a varied order. As usual, no programmed consequences followed the subjects' behavior throughout the test, and no pennies were ever given for correct responses on the equivalence-probe trials

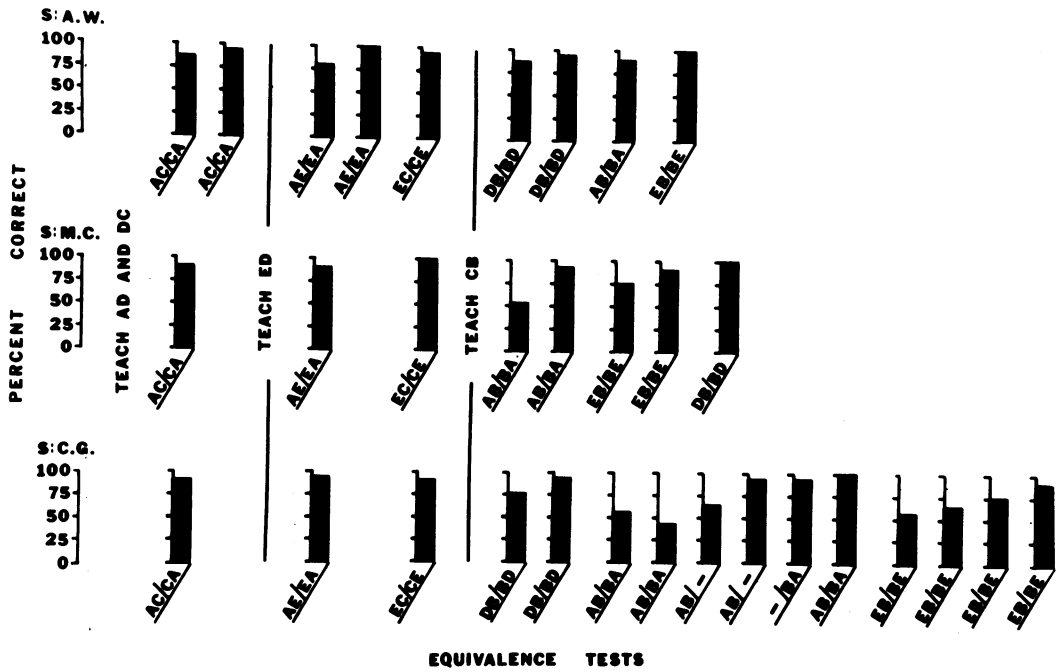

Fig. 3. Equivalence data for subjects A.W., M.C., and C.G. Each row of bars represents successive scores on equivalence tests for one child. The data are for equivalence probes only; baseline data have been omitted. Each bar gives the outcome of the particular test designated beneath it.

after the session was over.

Phase 7: Oral-Naming Test

After all matching tasks were completed, the oral-naming test was administered to determine whether names had served as mediators in the formation of equivalences. Did classes form because subjects had assigned a common name to stimuli comprising the same class?

In the first part of the test, subjects were asked to name the A, B, C, D, and E stimuli aloud. The 15 stimuli were presented in random order and no feedback was given as to the adequacy of the label.

The second half of the assessment consisted of the presentation of nine trials from an EB/BE equivalence test. Subjects were required to name each stimulus aloud as they pointed to it in the course of matching. The names for the E and B stimuli were derived from three equivalence-probe trials; the names for the A, C, and D stimuli were recorded on six baseline-matching trials. Neither naming nor matching responses were followed by programmed consequences.

RESULTS

Equivalence Tests

Each row in Figure 3 represents the teaching sequence and the corresponding scores on equivalence tests for subjects A.W., M.C., and C.G., respectively. The results for Subject G.H. were different and will be described later. Each bar represents only the performances on the equivalence-probe trials denoted beneath it; the baseline-matching data have been omitted because accuracy was always at or near 100% correct. Although there was no programmed feedback throughout testing (and thus no designations of "correctness" to the subjects), the percentages in the figure were derived on the basis of the predicted (and thus "correct" to the experimenter) equivalence classes. A subject's score had to exceed 90% correct for us to conclude that the classes had formed.

When a given equivalence was assessed more than once, multiple versions of the test were used.

After learning AD and DC matching, these three subjects indicated equivalence between stimuli in sets A and C—Subjects M.C. and C.G. in the first test session and Subject A.W. in the second session. Following the addition of the ED performance to the overall matching baseline, there were two new equivalences that could form. The AE/EA relation was assessed first, and, as before, Subjects M.C. and C.G. met the equivalence criterion in one session and A.W. in two. All subjects, however, showed the CE/EC equivalence in one test.

The last task to be trained was CB matching, so that the baseline now consisted of AD, DC, ED, and CB. The relevant equivalence tests evaluated the subjects' ability to match each stimulus in Set B to corresponding stimuli in Set A (AB/BA), D (DB/BD), and E (EB/BE); Subject A.W. scored 88% and 92%, respectively, on DB/BE, 92% on AB/BA, and 100% on EB/BE equivalence probes. Subject M.C. needed two tests (54% and 96%) to achieve the AB/BA performance, and two tests (75% and 92%) for EB/BE, but he scored perfectly in the first administration of DB/BD.

Subject C.G., however, required a variation in the test sequence before he was able to demonstrate all the instances of expanded stimulus classes. On the DB/BD equivalence, he first scored 79% and then 96%. His performance on the AB/BA probes, however, were 58% and 46%, respectively, in the initial two administrations of the test. Because he scored higher on the AB probes than on those that assessed BA, we presented a new version of the test consisting of the conventional 60 baseline trials, but with 24 AB probes only. Subject C.G. scored 66% and 96% in consecutive tests with AB trials, and when we presented him with a comparable test with only BA probes, he achieved 96%. He then performed perfectly on the standard version of AB/BA.

Subject C.G. also experienced difficulty on the EB/BE test. He required four administrations (58%, 66%, 75%, and 92%) to reach a criterion score.

The test outcomes for the fourth subject, G.H. ultimately yielded positive results, although the route to the emergence of the enlarged classes was different from that of his counterparts. The results, including baseline scores, are shown in Table 4. The first column on the left shows the matching performances that were directly trained, the second column lists test numbers, the next six headings (AC/CA, AE/EA . . . EB/BE) indicate the particular equivalence relation under consideration for the corresponding test, and the four columns to the right show the score on the baseline performances for each test.

Consistent with the order of training outlined in Table 1, we first taught the AD and DC matching tasks to Subject G.H. He then scored perfectly on the AC/CA equivalence test. After learning the ED performance, G.H. achieved a perfect score on the EC/CE equivalences. On the initial presentation of the AE/EA test, he scored only 62% correct; but his baseline performance on ED matching was weak (83%, or 10 of 12 correct). On the next AE/EA test, his equivalence score dropped to 58% despite a criterion baseline score. His accuracy improved by one additional correct response in each of three succeeding tests (Tests 5 through 7), but dropped to 58% in the sixth session in which AE/EA equivalences were being assessed. We then retested CE/EC to ensure that the performance had not been disrupted.

Based on our findings from Subject C.G., we separated the AE and EA probes into two different tests with 24 equivalence trials each, but with the same AD, DC, and ED baseline (36 trials). Three tests measuring the AE performance all yielded identical 66% scores: When sample A1 was presented, he correctly chose comparison stimulus E1, but comparison E3 was chosen both in the presence of sample stimuli A2 and A3. After the EC/CE equivalence was retested and found intact, two tests consisting of EA trials (Tests 14 and 15) also produced 66% scores.

Table 4

Subject G.H.: scores (percentage correct) during successive tests. A task noted in parentheses indicates that probes of only one type were assessed and is represented by the adjacent score.

		Equivalence						Baseline			
Teach	*Test Number*	*AC/CA*	*AE/EA*	*CE/EC*	*DB/BD*	*AB/BA*	*EB/BE*	*AD*	*DC*	*ED*	*CB*
AD,DC	1	100	–	–	–	–	–	100	100	–	–
ED	2	–	–	100	–	–	–	100	100	100	–
	3	–	62	–	–	–	–	100	100	83	–
	4	–	58	–	–	–	–	100	100	92	–
	5	–	62	–	–	–	–	100	100	92	–
	6	–	66	–	–	–	–	100	100	100	–
	7	–	70	–	–	–	–	92	100	100	–
	8	–	58	–	–	–	–	100	100	100	–
	9	–	–	96	–	–	–	100	100	100	–
	10	–	66(AE)	–	–	–	–	100	100	96	–
	11	–	66(AE)	–	–	–	–	100	92	100	–
	12	–	66(AE)	–	–	–	–	92	100	100	–
	13	–	–	100	–	–	–	100	100	100	–
	14	–	66(EA)	–	–	–	–	100	100	100	–
	15	–	66(EA)	–	–	–	–	100	100	100	–
	16	–	71(EA)	–	–	–	–	100	–	96	–
	17	–	88(EA)	–	–	–	–	83	–	100	–
	18	–	83(EA)	–	–	–	–	100	–	100	–
	19	–	100(EA)	–	–	–	–	92	–	92	–
	20	–	100	–	–	–	–	100	–	100	–
	21	–	100	–	–	–	–	100	100	100	–
CB	22	–	–	–	–	–	71	100	100	100	92
	23	–	–	–	–	–	58	100	100	100	100
	24	–	–	–	–	–	33	100	100	100	100
	25	–	–	–	–	–	67(EB)	100	100	100	100
	26	–	–	–	–	–	71(EB)	100	100	100	92
	27	–	–	–	–	–	92(EB)	100	100	100	100
	28	–	–	–	–	–	79(BE)	100	100	100	100
	29	–	–	–	–	–	100(BE)	100	100	100	100
	30	–	–	–	–	–	87	100	100	100	92
	31	–	–	–	–	–	100	100	100	100	100
	32	–	–	–	–	87	–	100	100	100	100
	33	–	–	–	–	92	–	100	100	100	100
	34	–	–	–	96	–	–	100	100	100	100

We then reasoned that only AD and ED matching were necessary prerequisites for the AE/EA equivalence (see Figure 2). If we temporarily removed the DC performance from the test baseline, would AE/EA matching emerge? Test 16 consisted of EA probes in the context of an abbreviated baseline (AD and ED). The number of baseline trials was reduced to 24 trials; there were still 24 equivalence probes. On four consecutive tests, he scored 71%, 88%, 83%, and 100%. With the modified baseline, G.H. then achieved a perfect score on EA and AE probes. In Test 21 we returned to the original baseline (AD, DC, ED), and he scored 100% correct in the AE/EA probes.

Following the addition of CB to the learned matching baseline, G.H. was presented with three tests for EB/BE equivalences, scoring 71%, 58%, and 33%, respectively (Tests 22 through 24). We then presented a test for EB matching (with the AD, DC, ED, CB baseline), and in three sessions his accuracy rose to 67%, 71%, and finally to a criterion-level 92%. We then assessed BE in the same manner and found that he required two tests before scoring 100%. Two additional tests (Tests 30 and 31) with

Table 5
Oral-Naming Data for each Subject

	Subject A. W.		*Subject M. C.*		*Subject C. G.*		*Subject G. H.*	
Stimulus	*Sample Condition*	*Matching Condition*	*Sample Condition*	*Matching Condition*	*Sample Condition*	*Matching Condition*	*Sample Condition*	*Matching Condition*
A1	Triangle	Triangle	Body of a girl	Triangle	Triangle	Capital A	Triangle	Triangle
B1	W	W	Three	Three	Three	Three	Triangle	Small M
C1	V	V	Upside-down V	Upside-down V	A	A	Half an A	Triangle
D1	Hook	Hook	B	Hanger	A	A	S	Hanger
E1	Y	Y	Upside-down Y	Upside-down Y	K	Y	Upside-down Y	Upside-down Y
A2	Three	Three	Three	Three	Three	P	Z	Z
B2	P	P	I	Moon	P	P	One if top erased	One
C2	U	U	Upside-down U	Horseshoe	N	N	Upside-down U	U
D2	Circle	Circle	Script O	Hanger	O	O	O	Circle
E2	W	W	W	W	Q	O	Small N	Small N
A3	Lines	Lines	Three	Straight lines	I	I	Half an E	E almost
B3	Six	Six	Six	Six	G	Six	Upside-down nine	Six
C3	L	L	Upside-down L	Hangman	R	R	Upside-down L	L
D3	E	E	Script E	Script E	Foot	None Given	E	E
E3	Circle	Circle	Tool	Two	Two	Two	Skinny O	Pretzel

combined EB and BE trials were needed for him to attain a 100% score.

G.H. scored 87% correct in the initial AB/BA test and 92% in the second. He required only one session to score 96% on the DB/BD performance.

Naming Tests

The naming data, obtained after the completion of equivalence testing, are shown in Table 5. The names have been listed for each child according to equivalence-class membership of the labeled stimulus, with Class 1 (A1, B1, C1, D1, E1) listed first, Class 2 (A2, B2, C2, D2, E2) second, and Class 3 (A3, B3, C3, D3, E3) third. In no case did a subject assign a common label to all stimuli that had been demonstrated to be members of the same class. This finding held both for the sample-naming condition and for naming in the context of performing matching-to-sample tasks. The accuracy on the matching performance, itself, was perfect. The closest approximation to a name-mediated equivalence was context-naming for Class 1 for Subject C.G., who responded "A" to A1, C1, and D1. Most of the labels used by subjects consisted of letters and number that shared common physical properties with the experimental stimuli. The word "triangle," for example, was used by all subjects to label upper-case delta; "three" was a common response to upper-case sigma.

The two methods for obtaining labels also yielded differing naming patterns across subjects. Subject A.W. provided the same names under both conditions; two subjects (C.G. and G.H.) were consistent for nine stimuli; and Subject M.C. gave consistent names for seven stimuli.

DISCUSSION

The results of this experiment show that normal children can demonstrate purely visual equivalences with matching-to-sample procedures, and that classes can be expanded by adding new performances to the matching baseline. Subjects were first taught AD and DC matching (see Figure 2 and Table 1) and were then able to perform AC/CA matching without additional training. After ED was taught directly, AE/EA and EC/CE performances emerged. Following CB training, three new equivalences emerged: AB/BA, EB/BE, and DB/BD. Programmed consequences for matching accuracy *never* occurred during the course of testing, for either baseline-matching trials or equivalence-probe trials. These data support the notion that it is not necessary for class membership to include auditory stimuli.

The second major finding of the present study is that equivalences can be formed in the absence of mediating names. The potential importance of auditory stimuli to the formation of equivalences had been suggested by the correlation between the demonstration of equivalences and the emergence of a common oral name for stimuli in the same class (Lazar & Mackay, in press; Sidman, 1971; Sidman & Cresson, 1973; Sidman et al., 1974). In each case, the oral name corresponded to the auditory member of the class (e.g., dictated picture or color names). Our oral-naming tests revealed that subjects had not assigned the same label to stimuli in the same class. This was found both in the sample-naming test and in naming during the course of performing the matching task.

Sidman et al. (1974) found that a few equivalences emerged before subjects were able to name stimuli, and Sidman and Tailby (1982) showed that one subject (E.W.) hesitated and expressed doubts when providing appropriate class names, suggesting that he had not labeled stimuli prior to the oral-naming test. In this latter study, however, Subject J.O.'s data showed that he could provide a consistent name to each member of a class but could *not* demonstrate equivalence in the context of matching to sample. Our findings, in conjunction with those of Sidman and Tailby, provide strong evidence that naming is neither necessary nor sufficient for the formation of stimulus equivalences.

This conclusion corroborates a recent view of the controlling stimuli when matching-to-sample techniques produce equivalences. As noted by Sidman and Tailby (1982), a conditional discrimination, such as matching to sample, does not require a different response to each stimulus; the subject need only touch or point to choose among stimuli. Inasmuch as it has now been demonstrated that naming is not a prerequisite to equivalence formation, we can take the position that human matching is governed by a relation between the sample and its corresponding correct comparison. In the case of arbitrary matching, the critical variable controlling choice responses is a relation between two physically different stimuli.

One implication of relational stimulus control is a functional approach to the generative properties associated with human language (see Chomsky, 1965; Fodor, Bever, & Garrett, 1974). Sidman (1971) and Sidman and Cresson (1973) showed that the emergence of picture/printed-word equivalences resembled semantic performances. Lazar (1977) and Lazar and Kotlarchyk (1984) found that adults and children can respond in untrained sequences that can be predicted on the basis of equivalences. Is it therefore possible that a specific language *performance* may be unique, but the prerequisite semantic and syntactic *relations* have an explicit training history? An affirmative answer would support a behavioral analysis of the conditions that produce linguistic competence. Moreover, the difficulty that animals have in performing true matching (see Sidman et al., 1982, for discussion) presents the possibility that nonhuman species may be limited in their capacity to respond on the basis of stimulus-stimulus relations, and perhaps this may account for their difficulty in learning the patterns and relationships of human language (see Terrace, 1979).

The data from this experiment also extend the findings of Sidman and Tailby (1982), Spradlin et al. (1973), and Wetherby et al. (1983) regarding class expansion. Sidman and Tailby taught AB, AC, and DC matching and then found that child subjects could demonstrate BD and DB preformances, documenting the presence of three 4-member classes. Our results show that equivalence classes can be expanded in two ways.

First, we trained AD and DC matching (see Figure 2) and then tested for the ability to match A and C stimuli. The emergence of this untrained performance can be regarded as an instance of 3-stage equivalence (see Sidman & Tailby, 1982) inasmuch as there were three sets of stimuli (A, D, and C) involved in the prerequisite conditions for the AC/CA performance. At this point there were three stimuli in each of three classes (A1, B1, C1; A2, B2, C2; A3, B3, C3). When we added ED matching to the subject's learned baseline, two additional equivalences became possible, AE/EA and CE/EC. These new performances also constituted 3-stage equivalences because the prerequisite conditions in each case still required three sets of stimuli: AE/EA depended upon class membership among A, D, and E stimuli, and CE/EC required membership among elements in C, D, and E. Despite the 3-stage data from each test alone, the combined results showed that the three classes had been expanded to four members, consisting of A, D, C, and E. Thus, the number of stages encompassed by an equivalence test between class members does not by itself necessarily indicate the number of stimuli in that class, and it is possible to expand class size without increasing the number of stages between class members.

In contrast, the consequences of teaching CB matching were to enlarge the number of members in each class and to increase stage size. Four-stage equivalences could be found both in the AB/BA (*A*-D-C-*B*) and EB/BE (*E*-D-C-*B*) performances. It was also demonstrated that subjects could perform the necessary DB/BD, 3-stage equivalence, thereby providing internal consistency to the stimulus class notion. Combined, these final tests showed the presence of three 5-member classes. What has yet to be determined is the functional relation between the type of class expansion, and the properties of the stimulus classes and their members.

Two related features of our data, however, pose unresolved questions regarding the emergence of stimulus equivalences. First, accuracy on an equivalence test can increase in successive administrations, even without programmed consequences on test trials. The most extreme cases occurred with Subjects C.G. and G.H. The EB/BE performance for Subject C.G. increased monotonically from 58% to 92% correct in four sessions. For Subject G.H., accuracy on the EA matching task rose from 71% to 100% (Table 5, Tests 16 through 19), and the EB performance reached 92% in three sessions (Tests 25 through 27). The gradual emergence of EA for Subject G.H. cannot be attributed to the function of stimuli in Set E serving sample and comparison roles for the first time because he had already demonstrated EC/CE equivalences. Overall, every subject required multiple testing on at least two equivalence tasks, and no subject demonstrated the first equivalence following CB training upon initial testing. There was, however, variability across subjects in terms of particular tasks that needed repeated presentations. These data replicate those of Lazar (1977), Lazar and Kotlarchyk (1984), and Spradlin et al. (1973), and suggest variables affecting the demonstration of class formation that have not yet been identified.

The second feature concerns the special set of testing conditions for Subjects C.G. and G.H. In C.G.'s case, the inability to perform the AB/BA equivalence was remedied by assessing AB and BA tasks in separate tests. It is possible that isolating these two performances simply permitted additional numbers of trial presentations, first for AB and then BA matching, respectively, suggesting a quantitative variable. It must also be considered that there are prerequisites to a successful equivalence performance not assessed prior to the actual test. For C.G. the AB task required transitivity (see Sidman et al., 1982) across the learned baseline (e.g., if AD and DC and CB, then AB), but BA matching required transitivity and symmetry (e.g., if AD and DC and CB, then AB; if AB then BA). Because AB had not emerged, BA was not possible; after AB was established, criterion accuracy on BA was achieved in one session.

The data from Subject G.H. provide a more convincing example for the analysis of prerequisites. Following ED training, he scored perfectly on the CE/EC test, but performance in six consecutive tests on the AE/EA equivalence (Table 5, Tests 3 through 8) yielded no evidence for expanded classes. Separating AE and EA matching was also unsuccessful. It was not until we reduced the test baseline to AD and ED matching, the two learned tasks necessary for the AE performance, that accuracy levels began to rise. We then tested the AE/EA equivalence with the modified baseline before assessing it in the context of the full baseline (AD, DC, and ED). The success of these remedial procedures suggests that Subject G.H. had learned the corresponding baseline tasks for the demonstration of the AE/EA equivalence, but that *testing* the prerequisites was in some way important in establishing them as part of his repertoire. These data for G.H., as well as those for all subjects showing gradual emergence of equivalences, indicate the need to clarify the role of testing as teaching in class formation and expansion.

REFERENCES

Carter, D. E., & Eckerman, D. A. (1975). Symbolic matching by pigeons: Rate of learning complex discriminations predicted from simple discriminations. *Science,* **187,** 662-664.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge: MIT Press.

Cumming, W. W., & Berryman, R. (1965). The complex discriminated operant: Studies of matching-to-sample and related problems. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 284-330). Stanford, CA: Stanford University Press.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language.* New York: McGraw-Hill.

Hockett, C. D. (1960). The origin of speech. *Scientific American,* **203**(3), 89-96.

Lazar, R. (1977). Extending sequence-class membership with matching to sample. *Journal of the Experimental Analysis of Behavior,* **27,** 381-392.

Lazar, R. M. (1983). Facilitating symbolic matching in preschool children. Manuscript submitted for publication.

Lazar, R. M., & Kotlarchyk, B. J. (1984). Formation of second-order equivalences in children. Manuscript submitted for publication.

Lazar, R. M., & Mackay, H. A. (in press). Teaching spelling via stimulus equivalences. In *Exceptional child education resources.* Washington, D.C.: National Institute of Education.

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

Santi, A. (1978). The role of physical identity of the sample and correct comparison stimulus in matching-to-sample paradigms. *Journal of the Experimental Analysis of Behavior,* **29,** 511-516.

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

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

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

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 discriminations vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior,* **37,** 5-22.

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

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

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

Taylor, I. K., & Taylor, M. M. (1965). Another look at phonetic symbolism. *Psychological Bulletin,* **64**, 413-427.

Terrace, H. S. (1979). Is problem-solving language? *Journal of the Experimental Analysis of Behavior,* **31**, 161-175.

Wetherby, B., Karlan, G. R., & Spradlin, J. E. (1983). The development of derived stimulus relations through training in arbitrary-matching sequences. *Journal of the Experimental Analysis of Behavior,* 1983, **40**, 69-78.

Received April 25, 1983
Final acceptance January 22, 1984