JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1982, 37, 5-22 NUMBER 1 (JANUARY)

CONDITIONAL DISCRIMINATION VS. MATCHING TO SAMPLE: AN EXPANSION OF THE TESTING PARADIGM

MURRAY SIDMAN AND WILLIAM TAILBY

NORTHEASTERN UNIVERSITY

A subject's performance under a conditional-discrimination procedure defines conditional relations between stimuli: "If A1, then B1; if A2, then B2." The procedure may also generate matching to sample. If so, the stimuli will be related not only by conditionality, but by equivalence: A1 and B1 will become equivalent members of one stimulus class, A2 and B2 of another. One paradigm for testing whether a conditional-discrimination procedure has generated equivalence relations uses three sets of stimuli, A, B, and C, three stimuli per set. Subjects learn to select Set-B and Set-C comparisons conditionally upon Set-A samples. Having been explicitly taught six sample-comparison relations, A1B1, A1C1, A2B2, A2C2, A3B3, and A3C3, subjects prove immediately capable of matching the B- and C-stimuli; six new relations emerge (B1C1, B2C2, B3C3, C1B1, C2B2, C3B3). The 12 stimulus relations, six taught and six emergent, define the existence of three three-member stimulus classes, A1B1C1, A2B2C2, and A3B3C3. This paradigm was expanded by introducing three more stimuli (Set D), and teaching eight children not only the AB and AC relations but DC relations also—selecting Set-C comparisons conditionally upon Set-D samples. Six of the children proved immediately capable of matching the B- and D-stimuli to each other. By selecting appropriate Set-B comparisons conditionally upon Set-D samples, and Set-D comparisons conditionally upon Set-B samples, they demonstrated the existence of three four-member stimulus classes, A1B1C1D1, A2B2C2D2, and A3B3C3D3. These larger classes were confirmed by the subjects' success with the prerequisite lower-level conditional relations; they were also able to select Set-D comparisons conditionally upon samples from Sets A and C, and to do the BC and CB matching that defined the original three-member classes. Adding the three DC relations therefore generated 12 more, three each in BD, DB, AD, and CD. Enlarging each class by one member brought about a disproportionate increase in the number of emergent relations. Ancillary oral naming tests suggested that the subject's application of the same name to each stimulus was neither necessary nor sufficient to establish classes of equivalent stimuli.

Key words: conditional discrimination, matching to sample, stimulus equivalence, stimulus classes, stimulus control, key press, children

Given two discriminative stimuli, B1 and B2, a subject selects B1 if a conditional stimulus, A1, is present, and selects B2 if the conditional stimulus is A2. This conditional discrimination provides a procedural definition of conditional relations between stimuli. In their simplest form, the conditional relations are: If A1, then B1; if A2, then B2. Although additional considerations may lead to a more precise description of an "if . . . then" relation, its existence is directly observable by reference to the subject's ongoing interactions with the procedure. Testing for the existence of a conditional relation requires no modification of the establishing procedure.

A well-established conditional discrimination is often assumed to demonstrate not just conditional relations between stimuli, but equivalence relations also. In addition to their "if . . . then" relation, Stimuli A1 and B1 are also supposed to be equivalent; similarly, A2 and B2 are supposed to be related both by conditionality and equivalence. Investigators who have assumed, explicitly or implicitly, that the conditional-discrimination procedure generates equivalence relations often call the subject's performance "matching to sample." Stimuli B1 and B2 become comparisons, to be compared to samples A1 and A2 and matched appropriately. When all sample and comparison stimuli are physically different,

This research was supported by Grants MH30063, HD05124, and NE-G-00-3-0011. We are grateful to F. Garth Fletcher for apparatus design, construction, and maintenance. Reprints may be obtained from Murray Sidman, Department of Psychology, 282 Nightingale Hall, Northeastern University, Boston, Massachusetts 02115.

equivalence is suggested by calling the performance "nonidentity," "arbitrary," or "symbolic" matching.

Unlike conditionality, equivalence is not definable solely by reference to the subject's ongoing interactions with the establishing procedure. To determine whether a performance involves something more than conditional relations between sample and comparison stimuli requires additional tests. Suppose, for example, a relation, R, between stimuli *a* and *b* fulfills the procedural requirements for conditionality, or "if *a*, then *b*." Is the relation between *a* and *b*, *a*R*b*, also an equivalence relation? The subject's current performance gives no clue. Appropriate tests can, however, be derived from the three properties that modern elementary mathematics texts specify as the definition of the equivalence relation: reflexivity, symmetry, and transitivity.

To determine that the conditional relation, R, is reflexive, one must show that each stimulus bears the relation to itself; *a*R*a* (if *a*, then *a*) and *b*R*b* (if *b*, then *b*) must hold true. Reflexivity can therefore be tested by an identity-matching procedure that requires the subject to match stimulus *a* to itself, and *b* to itself.

It will not suffice to *teach* the conditional relations, *a*R*a* and *b*R*b*. One does not know whether a subject to whom the relations had to be taught is matching each stimulus to itself (if matching is involved at all), or is perhaps controlled by one feature of the sample and another feature of the correct comparison. For example, instead of matching "red" to "red," and "green" to "green," a subject might be matching "red" to "bright," and "green" to "dark." Only if the subject matches each new stimulus to itself without differential reinforcement or other current instructions can one be certain that identity is the basis for the performance. Given a subject who is familiar with the stimuli and procedures, the proof of reflexivity is generalized identity matching.

To demonstrate that the relation, R, is symmetric, one must show that both *a*R*b* and *b*R*a* hold true. A subject who matches a sample *a* to comparison *b* is then required, without further training, to match sample *b* to comparison *a*, reversing "if *a*, then *b*" to "if *b*, then *a*." Given a subject who is familiar with each of the stimuli separately, both as sample and comparison, the proof of *a*R*b* symmetry is functional sample-comparison reversibility (Sidman, Rauzin, Lazar, Cunningham, Tailby, & Carrigan, 1982).

To determine whether R is transitive requires a third stimulus, *c*. Once "if *a*, then *b*" and "if *b*, then *c*" have been established, transitivity requires "if *a*, then *c*" to emerge without differential reinforcement or other current instructions. Given a subject who has learned two conditional relations, *a*R*b* and *b*R*c*, with the comparison in the first serving as the sample in the second, the proof of transitivity is the emergence of a third conditional relation, *a*R*c*, in which the subject matches the sample from the first relation to the comparison from the second.

Calling a conditional relation "matching to sample," then, requires proof that the relation possesses all three properties of an equivalence relation, as listed in the upper section of Table 1. Successful generalized matching will prove the relation reflexive, a property that must hold for each stimulus. Sample-comparison reversibility (Lazar, 1977) will prove symmetry, a property that must hold for each pair of related stimuli. Emergence of a third relation, in which the subject matches the sample from one of two prerequisite relations to the comparison from the other, will prove transitivity, a property that must hold for at least three interrelated stimuli.

Table 1

The Equivalence Relation

Equivalence relations must be:	
1. Reflexive:	*a*R*a*
2. Symmetric:	If *a*R*b*, then *b*R*a*
3. Transitive:	If *a*R*b* and *b*R*c*, then *a*R*c*
Combined tests for symmetry and transitivity are:	
A. Teach *a*R*b* and *a*R*c*.	Test *b*R*c* and *c*R*b*.
B. Teach *b*R*a* and *c*R*a*.	Test *b*R*c* and *c*R*b*.

Symmetry and transitivity can be evaluated simultaneously. To prepare for one kind of combined test (A, in the lower section of Table 1), first teach the subject two relations, *a*R*b* and *a*R*c*, that share the same sample. Then test for the emergence of the conditional relations, *b*R*c* and *c*R*b*. If R is symmetric, so that both *a*R*b* and *b*R*a* hold true, then *b*R*a* and *a*R*c* will yield *b*R*c* *via* transitivity. Similarly, if both *a*R*c* and *c*R*a* hold true by symmetry, the combination of *c*R*a* and *a*R*b* will yield *c*R*b* via transitivity. Emergence of the new conditional relations, *b*R*c* and *c*R*b*, requires the explicitly established relations, *a*R*b* and *a*R*c*, to be both symmetric and transitive.

In early experiments that used these tests to determine whether conditional-discrimination procedures had generated matching-to-sample performances (Sidman, 1971; Sidman & Cresson, 1973), retarded youths first proved capable of generalized identity matching, thereby meeting the reflexivity criterion. They then learned (or demonstrated that they were already able) to select pictures (comparison stimuli) conditionally upon any of 20 dictated picture names (sample stimuli); AB in Figure 1 represents 20 conditional relations (A1B1, A2B2 . . . A20B20). A determination of whether the AB relations involved equivalence, in addition to conditionality, required another set of conditional discriminations. The subjects learned to select printed names conditionally upon the same 20 dictated names; AC in Figure 1 represents the 20 new relations (A1C1, A2C2 . . . A20C20). At the completion of their AB and AC training, the subjects could select any of 20 pictures or printed picture names conditionally upon a dictated name.

It was then possible to find out whether AB and AC were equivalence relations by giving a combined test for symmetry and transitivity (A, in Table 1). Proof of equivalence required the subjects to select an appropriate printed word conditionally upon a picture sample

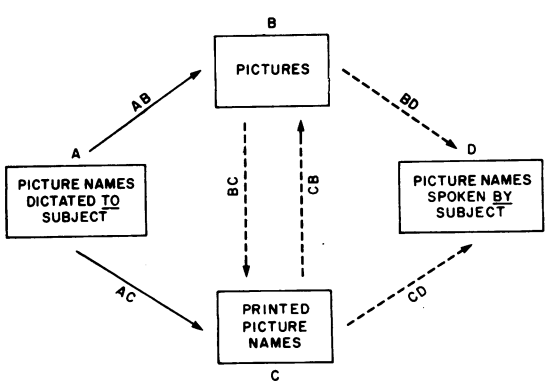

Fig. 1. A basic equivalence paradigm. Boxes A, B, and C represent stimuli, and Box D represents oral naming responses. Arrows AB, AC, BC. and CB point from sample to comparison stimuli and represent sets of conditional relations. Arrows BD and CD point from stimuli to naming responses. Solid arrows represent conditional relations that are explicitly taught to the subjects. Broken arrows represent conditional or oral naming relations that emerge after others have been explicitly taught.

(BC), and to select an appropriate picture conditionally upon a printed word (CB). The retarded subjects did relate the pictures and printed names correctly, even though they had not been explicitly trained to do so, and had demonstrated in pretests that they were unable to do so before learning both the AB and AC conditional discriminations. The conditional relations therefore met the criteria of reflexivity, symmetry, and transitivity, and the performances could be called "matching to sample."

The subjects' emergent ability to do two new sets of matching tasks, BC and CB, confirmed the creation of 20 three-member classes of equivalent stimuli: A1B1C1, A2B2C2 . . . A20B20C20. One of the classes, for example, contained the dictated word "boy," pictures of boys, and the printed word *boy;* another contained the dictated word "car," pictures of cars, and the printed word *car.* A necessary consequence of the establishment of classes of equivalent stimuli was the subject's ability to match members of a class to each other even without ever having done so before. The testing process itself can therefore cause new matching-to-sample performances to emerge without explicitly teaching them. Stimulus-class formation permits an impressive economy and efficiency in teaching and learning. The direct teaching of 40 conditional relations (20 AB and 20 AC) caused 40 more to emerge (20 BC and 20 CB). Actually, the spinoff was considerably greater; subjects also became capable of naming the pictures (BD) and the printed words (CD) aloud. Although matching to sample does not require subjects to name the stimuli, and they usually did not, oral naming emerged when tested. The original teaching of 40 conditional relations created 40 new conditional relations and 40 naming relations—80 new performances.

A major purpose of the present experiment was to add one more stimulus to each class and thereby test the power of equivalence relations to generate a larger network of interchangeable stimuli. Figure 2 shows the stimuli and illustrates the experiment's rationale. Previous studies using English language symbols had required extensive pretests to ensure that the subjects could not already do the critical matching and naming. In order to eliminate both the time required for pretests and the problems created by giving children tasks they are unable to perform, and to avoid the methodological dilemma of whether or not to reinforce correct responses during pretests, this experiment used Greek letters and letter names, stimuli that could be presumed unfamiliar to the subjects and therefore not requiring pretests at all. Reducing the number of stimuli from 20 to three in each set achieved additional simplification.

The upper "triangle" ABC in Figure 2 represents a smaller replication of the experiments summarized in the left triangle of Figure 1. Subjects learned to select letters in Sets B and C conditionally upon dictated names (A). If the conditional relations, AB and AC, were also equivalence relations, Test A (Table 1) would reveal the subjects' ability to match Set-B and Set-C letters to each other (BC and CB). Three three-member stimulus classes should emerge, each containing a dictated letter name, a Set-B letter and a Set-C letter.

Subjects were also taught to select comparison letters in Set C conditionally upon samples from Set D. Would the new DC relations expand the ABC classes? If they did, each three-member class would gain a fourth member from Set D: upper-case phi would join the "lambda" class; lower-case sigma, the "xi" class; and lower-case delta, the "gamma" class. Subjects would be able to match each stimulus to any other in its class. Having explicitly been taught only the AB, AC, and DC relations, they would then prove capable of matching samples from Set B to comparisons in Set D (BD), or samples from Set D to comparisons in Set B (DB). The first goal of the present experiment was to carry out these tests for the emergence of the four-member stimulus classes, ABCD.

Spradlin, Cotter, and Baxley (1973) successfully demonstrated the emergence of relations analogous to DB in Figure 2 but did not test for BD. Because the DB relations did not require DC symmetry, but the BD relations did (a point to be elaborated later), both DB and BD were required for proof that the procedures had generated four-member classes. The present experiment also extended the Spradlin, Cotter, and Baxley findings by testing the subjects' ability to name the stimuli and to perform other conditional discriminations besides DB and BD. For example, teaching AC and DC would establish the lower triangle ACD in Figure 2, once more making it possible to test for equivalence by evaluating symmetry and transitivity simultaneously. Equivalence tests in the lower and upper triangles, however, differed in at least two respects. First, the explicitly taught relations in the upper triangle, AB and AC, shared the same samples; analogous relations in the lower triangle, AC and DC, shared the same comparisons, like Test B in Table 1. The logic of both combined tests, however, was similar. If the DC relations were symmetric, so that conditionality also held for CD, then AC and CD could yield AD by transitivity; subjects would prove capable of matching auditory samples from Set A to comparison letters in Set D. Therefore, in addition to testing for four-member ABCD classes, the present experiment also tested for the emergence of two sets of three-member classes, ABC and ACD. Tests for BC and CB relations evaluated the ABC classes; tests for AD relations evaluated the ACD classes.

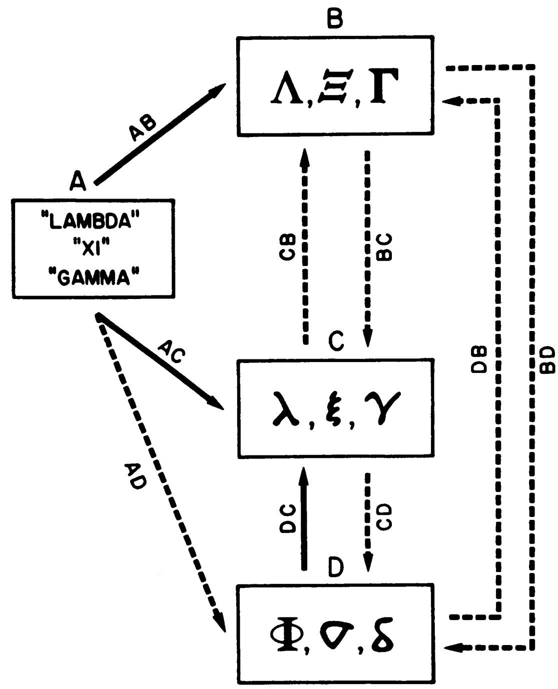

Fig. 2. The equivalence paradigm in the present experiment. The stimuli are a set of dictated Greek letter names (Set A) and three sets of printed Greek letters (Sets B, C, and D), three letters in each set. Arrows point from sample stimuli (only one presented at a time) to comparison stimuli. The solid arrows AB, AC, and DC represent conditional relations that are explicity taught to the subjects. The broken arrows CB, BC, AD, CD, BD, and DB represent conditional relations that are tested after others have been explicity taught.

A second difference between the upper and lower triangles in Figure 2 stemmed from the inability to present several auditory comparison stimuli simultaneously without altering their individual intelligibility. This technical feature of the conditional-discrimination procedure used here precluded a validation of AB- and AC-symmetry that was independent of the combined symmetry/transitivity tests. Testing BA or CA would have required the presentation of consecutive rather than simultaneous auditory comparison stimuli. Such a major procedural modification, although feasible, would at that time have disrupted the continuity of the experimental program. In the lower triangle, the DC relations involved only visual stimuli and could be tested for symmetry independently of the combined test. The present experiment, therefore, not only tested for equivalence relations by means of combined symmetry/transitivity tests, but also provided the independent evaluation of symmetry that the lower triangle in Figure 2 permitted.

The development of equivalence in all three sets of explicitly taught conditional discriminations—AB, AD, and DC—should create six sets of new matching-to-sample performances, summarized by broken arrows in Figure 2. The experimental tested all of these possibilities. Subjects were also tested for oral naming. Accurate naming of the letters in Sets B and C was perhaps to be expected, since the subjects were explicitly taught conditional relations between these letters and their auditory counterparts. Set-D naming was less surely predictable, since the subjects were never taught any direct relation between Set-D letters and the dictated names in Set A.

METHOD

Subjects

Eight normal children, all male except Subject A.D., participated in the experiment. Their ages (years-months) at the time they completed training and underwent their first tests for equivalence relations were: 5-0 (A.D.), 5-9 (J.L. and J.O.), 6-3 (D.B.), 6-5 (E.M.), 6-11 (D.W.), 7-4 (E.W.), and 7-5 (I.C.). Subjects A.D., J.L., D.B., and E.M. attended kindergarten, and Subjects J.O., D.W., E.W., and I.C. were in first grade. No other educational or test data were available. All children had at least one parent who completed high school, and all except Subjects A.D. and J.O. had at least one parent who had gone beyond high school. Each child, recruited by local newspaper advertisements, came to the laboratory several days a week with a parent who was paid after each session. The total number of sessions, each 15 to 30 minutes long, varied because the children differed in the amounts of training they required to learn and maintain the baseline conditional discriminations.

Apparatus and General Procedures

The child sat before a stimulus-response matrix of nine translucent hinged keys (windows) onto which stimuli were projected from the rear. Stimuli for a given trial had all been photographed on a single slide, and solenoid-operated shutters were mounted between projector and windows. The windows, each 7.3 cm in diameter, were arranged in a circle of eight, with the ninth in the center (Figure 3). The display diameter was 32.7 cm; the center-to-center distance was 9.5 cm between adjacent windows on the perimeter, and 12.7 cm between the center window and each of the others. Sample stimuli always appeared in the center, and comparison stimuli in the outer windows, but the present experiments did not use the three uppermost windows (shaded in Figure 3). Whenever the child pressed a window a limit switch signaled the solid-state programming equipment, impulse counters, and a 20-pen operations recorder.

Each trial began with a sample stimulus. Visual samples, black line drawings of Greek

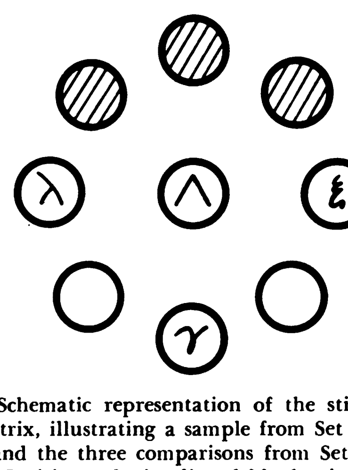

Fig. 3. Schematic representation of the stimulus-response matrix, illustrating a sample from Set B (center window) and the three comparisons from Set C (outer windows). Positions of stimuli and blank windows varied from trial to trial, but stimuli never appeared in the three uppermost windows.

letters on a white background, appeared on the center window. Auditory samples, Greek letter names, were dictated by a master tape and were repeated at 2-sec intervals by a continuous tape loop (Fletcher, Stoddard, & Sidman, 1971); the center window was illuminated but blank. Each sample, visual or auditory, remained throughout the trial (simultaneous matching), and trial durations had no limit. Visual and auditory samples never occurred in the same trial.

In matching-to-sample trials, after the sample was presented the subject had to press the center window to bring comparison stimuli onto the outer windows. Even dictated samples required this "observing response" but pressing the blank center window could not produce the comparisons until at least one complete sample word had sounded. No trial presented more than three comparison stimuli; at least two of the five functional comparison windows were blank on each trial, as indicated in Figure 3. Positions of correct, incorrect, and blank comparison windows varied from trial to trial. In sequences that included only two different sample stimuli, no more than three trials with the same sample could occur consecutively. With more than two samples, all possible trial types—sample-comparison combinations—had to occur before any could repeat. Also, all five functional comparison windows had to be scheduled as correct before any window could be correct again. With those exceptions, all trial types and all correct windows were equally probable on successive trials.

After comparison stimuli appeared, the subject had to press a comparison window. Unless a reduced reinforcement probability or an extinction probe (see below) precluded reinforcement, correct choices were followed by the sound of chimes, the disappearance of all stimuli, the delivery of a penny into an open receptacle below and at the left of the window matrix, and a 1.5-sec intertrial interval. Neither chimes nor penny followed an incorrect choice. The children kept all pennies they received.

Any window press during the intertrial interval postponed the next sample for 1.5 sec. Once the sample had come on, the subject could not produce comparison stimuli by pressing sample and comparison windows simultaneously; after any such simultaneous response the subject had to release both windows before a sample press could be effective. Once the comparisons had come on, sample presses no longer had any programmed consequences, but if the subject pressed correct, incorrect, or blank comparison windows simultaneously the trial was treated as incorrect. The programming apparatus arbitrarily specified simultaneity by waiting to define any response until the subject had released the window for .2 sec; pressure on another window during the preceding press, or within .2 sec of a release, defined a simultaneous response.

In oral naming tests, the subject had simply to respond to the instruction, "Tell me what you see", or "What is it?". The child pressed no windows, and only samples were presented, with no comparisons. After each oral response the experimenter, seated behind the child, pressed a handswitch to initiate the intertrial interval and the slide change. No chimes or pennies were presented, even for correct names (see below). Test sessions were recorded on magnetic tape, and the experimenter also recorded the child's naming responses during the tests. A secretary's transcription of each tape, in the absence of the visual stimuli to which the subject had been responding, never differed from the experimenter's record by more than one naming response in any 90-trial test.

Teaching and Testing Phases

The general plan of the experiment (details are provided below) was, A: first, acquaint the children with the procedures and confirm their ability to do visual-visual and auditory-visual conditional discriminations by pretesting them with stimuli that were presumably famliar, hues and hue names. Then, test identity matching with all of the Greek letters to be used in the experiment. These letter-letter (sample-comparison) tasks served as generalized matching-to-sample tests for reflexivity. B: next, teach the children the three sets of relations denoted in Figure 2 by the solid arrows, AB, AC, and DC. This established the baselines to be tested for symmetry and transitivity. C: finally, without additional teaching, evaluate the children's performance on the six sets of relations denoted in Figure 2 by the broken arrows, DB, BD, AD, BC, CB, and CD. Emergence of these new relations, or their failure to emerge, would prove whether or not the explicitly taught conditional relations were also equivalence relations, and whether the performances could be called "matching to sample." Also, test the children's oral responses to the B-, C-, and D-stimuli.

Table 2 outlines the planned sequence. All subjects went through Phases A (Pretests) and B (Teaching) in the indicated order. At the end of Phase B the children were performing the three sets of baseline conditional discriminations, AB, AC, and CD, at high levels of accuracy, although only a small proportion of their correct trials terminated with reinforcement (see below). Then, in Phase C (Final Tests), unreinforced probe trials, inserted among these infrequently reinforced baseline trials, constituted the tests for symmetry and transitivity, thereby evaluating the formation of equivalence relations in the baseline.

Pretests

After the delivery of two or three pennies had sufficed to accomplish magazine training, the experimenter provided nonverbal instruction by pressing windows appropriately for several hue-hue trials (hue sample and hue comparisons, with the sample hue the same as the correct comparison hue). Each subject then imitated the experimenter, meeting a criterion of 90% correct in a set of 25 hue-hue trials (Phase A1), followed by a similar criterion performance in matching dictated hue-name samples to the hue comparisons (Phase A2). These first tests with hues established the children's ability to cope with the experimental procedures. In the final pretest procedure (Phase A3) the stimuli were the Greek letters illustrated in Figure 2. On individual trials the sample and correct comparison were the same, and the two incorrect comparisons were always the other members of the same set (B, C, or D); comparisons from different sets never appeared together. This identity matching evaluated generalized matching to sample and thereby served as a test for reflexivity.

Table 2

Sequence of teaching and testing phases. The identifiers, AB, AC, DC, etc., refer to the sample-comparison relations diagrammed in Figure 2.

A. Pretests—Matching to Sample
  1. Hue samples and hue comparisons (Identity)
  2. Hue-name samples (dictated) and hue comparisons
  3. Greek-letter samples and comparisons (Identity)

B. Teaching—Matching to Sample
  1. AB: Set-A samples (dictated) and Set-B comparisons
  2. AC: Set-A samples (dictated) and Set-C comparisons
  3. AB and AC: Trials from Teaching Phases 1 and 2, mixed
  4. DC: Set-D samples (visual) and Set-C comparisons
  5. AB, AC, and DC: Trials from Teaching Phases 3 and 4, mixed
  6. Gradual lowering of reinforcement probability

C. Final Tests—Matching to Sample; Oral Naming
  1. DB: 4-stage equivalence probes, in baseline of AB, AC, and DC
  2. BD: 4-stage equivalence probes, in baseline of AB, AC, and DC
  3. AD: 3-stage equivalence probes, in baseline of AC and DC
  4. BC: 3-stage equivalence probes, in baseline of AB and AC
  5. CB: 3-stage equivalence probes, in baseline of AB and AC
  6. CD: Symmetry probes of DC relations, in baseline of DC
  7. B-, C-, and D-naming

Teaching and Maintenance Procedures

Teaching the three sets of baseline conditional discriminations went according to the sequence summarized in Table 2, Phases B1 through B6. Each teaching phase itself, however, was composed of a carefully programmed series of steps.

Teaching steps. Figures 4 and 5 depict the sample and comparison stimuli that comprised the trial types in each step. Starting with the AB relations, the children first had to match only two dictated samples, A1 and A2, to the comparison letters B1 and B2. The same two comparison letters, one correct and the other incorrect, appeared on every trial, as illustrated in the uppermost frame of Figure 4 (Task AB12). The children rotated through six balanced sets of 20 trials until they achieved at least 19 correct trials in a set. Then, in the second step, they went through sets of 20 trials with another pair of samples and correct comparisons, matching A1 and A3 to B1 and B3 (Figure 4, Task AB13) until they met the same criterion. Both steps included Sample A1 and Comparison B1. In the third step, the children learned to match the final pair of samples, A2 and A3, and Comparisons B2 and B3 (Task AB23). Each of these combinations—sample and correct comparison—had already been presented along with a different partner in Step 1 or 2.

Trials in these first three steps presented only two comparison stimuli (balanced with

TASK	SAMPLE	COMPARISONS CORRECT	COMPARISONS INCORRECT
AB12	"LAMBDA"	Λ	Ξ
	"XI"	Ξ	Λ
AB13	"LAMBDA"	Λ	Γ
	"GAMMA"	Γ	Λ
AB23	"XI"	Ξ	Γ
	"GAMMA"	Γ	Ξ
AB123	"LAMBDA"	Λ	Ξ, Γ
	"XI"	Ξ	Λ, Γ
	"GAMMA"	Γ	Λ, Ξ
AC12	"LAMBDA"	λ	ξ
	"XI"	ξ	λ
AC13	"LAMBDA"	λ	γ
	"GAMMA"	γ	λ
AC23	"XI"	ξ	γ
	"GAMMA"	γ	ξ
AC123	"LAMBDA"	λ	ξ, γ
	"XI"	ξ	λ, γ
	"GAMMA"	γ	λ, ξ
AB1	"LAMBDA"	Λ	Ξ, Γ
AC1	"LAMBDA"	λ	ξ, γ
AB2	"XI"	Ξ	Λ, Γ
AC2	"XI"	ξ	λ, γ
AB3	"GAMMA"	Γ	Λ, Ξ
AC3	"GAMMA"	γ	λ, ξ

Fig. 4. Each line depicts a trial type—sample stimulus, with correct and incorrect comparison(s). Each of the three main sections (bounded by solid lines) shows the trial types used in Teaching phases B1, B2, and B3 (Table 2)—AB relations first, then AC, and then both together. Frames within sections show the trial types used in each consecutive step. The numbers 1, 2, and 3 refer to the auditory stimuli shown from top to bottom in Figure 2, and to the visual stimuli shown from left to right.

respect to window positions), one related conditionally to the current sample, and the other to the second sample in the current pair. The fourth AB teaching step (Task AB123) mixed all three samples in sets of 30 trials, and presented all three comparisons, one correct and two incorrect, on every trial. When the children met a criterion of at least 29 correct trials in a set, they went on to learn the AC relations (Phase B2 in Table 2).

The central frames of Figure 4 show teaching Phase B2 broken down into steps like those of Phase B1. The children first learned pairs of AC trial types with only two comparisons per trial (Tasks AC12, AC13, and AC23). Then, sets of trials were given with all three samples, three comparisons per trial (Task AC123). Phase B3 then mixed the six AB and

TASK	SAMPLE	COMPARISONS CORRECT	COMPARISONS INCORRECT
DC12	φ	λ	ξ
	σ	ξ	λ
DC13	φ	λ	γ
	δ	γ	λ
DC23	σ	ξ	γ
	δ	γ	ξ
DC123	φ	λ	ξ, γ
	σ	ξ	λ, γ
	δ	γ	λ, ξ
AB1	"LAMBDA"	Λ	Ξ, Γ
AC1	"LAMBDA"	λ	ξ, γ
DC1	φ	λ	ξ, γ
AB2	"XI"	Ξ	Λ, Γ
AC2	"XI"	ξ	λ, γ
DC2	σ	ξ	λ, γ
AB3	"GAMMA"	Γ	Λ, Ξ
AC3	"GAMMA"	γ	λ, ξ
DC3	δ	γ	λ, ξ

Fig. 5. Each line depicts a trial type—sample stimulus, with correct and incorrect comparison(s). The frames show the trial types used in each consecutive step of Teaching Phases B4 and B5 (Table 2). DC trial types were taught first (upper four frames) and were then combined with the previously taught AB and AC trial types (lowest frame). The numbers 1, 2, and 3 refer to the auditory stimuli shown from top to bottom in Figure 2, and to the visual stimuli shown from left to right.

AC trial types, as depicted in the bottom frame of Figure 4, into balanced sets of 30 trials. By meeting a criterion of 29 correct trials in a set of 30, the children demonstrated that they could select either Set-B or Set-C letters conditionally upon dictated letter names from Set A.

Phase B4 (Table 2) consisted of a similar sequence in which the children learned the DC relations. The upper three frames of Figure 5 show the three pairs of DC trial types the children learned first, and the fourth frame shows the three-comparison DC trial types. Finally, Phase B5 mixed the nine AB, AC, and DC trial types (bottom section of Figure 5) in balanced sets of 45 trials. To complete Phase B5 the children had to meet a criterion of 44 correct trials in a set of 45. Later, sets of these nine trial types would constitute a baseline into which probe trials would be inserted to evaluate the formation of equivalence relations.

Reinforcement probability. Before inserting unreinforced probes, the reinforcement probability on correct baseline trials was gradually reduced (Phase B6) from 1.00 to .20. The sequence of probability steps was 1.00, .75, .50, .40, .30, and .20; subjects had to meet the 90% baseline accuracy criterion at each step before the probability could be lowered. It was occasionally necessary to halt a decreasing accuracy trend by increasing the reinforcement probability and then reducing it again. Since no probe trials would be reinforced, the reinforcement probability on baseline trials during subsequent tests was increased sufficiently to maintain the overall probability at .20. Probabilities were controlled by a probability gate (BRS/LVE PP-201), its output modified to permit no more than nine consecutive baseline trials to go unreinforced.

Before any set of trials in which the reinforcement probability was to be less than 1.00, the subject was told, "You won't always get a penny from now on, but you will have a chance to get more later." Then, at the end of the session, the child was given enough hue-hue matching trials, with all correct choices reinforced, to make up for the earlier correct baseline trials that had gone unreinforced.

Maintenance and review. Each new teaching session began with a review of the most advanced performance the subject had achieved, and the subject had to meet the learning criterion again before going on to the next step. For example, Tasks AB12, AB13, and AB23 (Figure 4) were reviewed at the beginning of each teaching session until the child had learned AB123; then, AB123 was reviewed at the beginning of all sessions in which the subject was learning the next four tasks; since the final task (bottom frame of Figure 4) included all AB and AC trial types, this was all that needed review during the next teaching phase; etc. Finally, after the reduction in reinforcement probability, each session in which a test was scheduled began with a review of the trial types that would serve as the sparsely reinforced baseline for the insertion of probes (each of these baselines is described below). To proceed with the test, subjects had to meet the usual accuracy criterion on this review; if they did not, reviews continued for the rest of the session. A child who failed frequently to meet a review criterion returned to an earlier teaching step and advanced through the sequence again.

Tests

Phases C1 through C7 in Table 2 summarize the tests with reference to the relations outlined in Figure 2.

Four-stage equivalence. Phase C1 inserted DB trials (samples from Set D, comparisons from Set B) as unreinforced probes into a sparsely reinforced baseline of AB, AC, and DC trials. Phase C2 inserted BD probes into the same baseline. Because emergence of the DB and BD performances required conditional relations within four sets of stimuli (A, B, C, and D), Phases C1 and C2 in Table 2 are called "four-stage equivalence probes." This "n-stage" terminology is consistent with a similar usage in conjunction with analogous paradigms that have been studied largely by means of paired-associate techniques (e.g., Jenkins, 1963).

Figure 6 shows the trial types (sample-comparison combinations) in the four-stage tests, with each of the three kinds of DB probes below its prerequisite AB, AC, and DC trial types. For example, emergence of the DB1 (probe) relation (Sample D1, Comparison B1) required the subjects to have learned the AB1, AC1, and DC1 relations that are shown just above DB1. In Phase C2, the three BD probes depicted at the bottom of Figure 6 replaced the DB trials. Each DB or BD test had 120 trials, 10 of each trial type (90 baseline and 30 probe trials).

TASK	SAMPLE	COMPARISONS CORRECT	COMPARISONS INCORRECT
AB1	"LAMBDA"	Λ	Ξ, Γ
AC1	"LAMBDA"	λ	ξ, γ
DC1	φ	λ	ξ, γ
DB1(PROBE)	φ	(Λ)	(Ξ, Γ)
AB2	"XI"	Ξ	Λ, Γ
AC2	"XI"	ξ	λ, γ
DC2	σ	ξ	λ, γ
DB2(PROBE)	σ	(Ξ)	(Λ, Γ)
AB3	"GAMMA"	Γ	Λ, Ξ
AC3	"GAMMA"	γ	λ, ξ
DC3	δ	γ	λ, ξ
DB3(PROBE)	δ	(Γ)	(Λ, Γ)
BD1(PROBE)	Λ	(φ)	(σ, δ)
BD2(PROBE)	Ξ	(σ)	(φ, δ)
BD3(PROBE)	Γ	(δ)	(φ, σ)

Fig. 6. Each line depicts a trial type—sample stimulus, with correct and incorrect comparisons. The upper section shows the AB, AC, and DC baseline trial types and the four-stage DB probes used in Final Test Phase C1 (Table 2). Each DB probe is just below its three prerequisite baseline trial types. The three lines in the lower section show the four-stage BD probes that were inserted in place of DB trials during Final Test Phase C2 (Table 2). The numbers 1, 2, and 3 refer to the auditory stimuli shown from top to bottom in Figure 2 and to the visual stimuli shown from left to right. Because probe trials were never reinforced, correct and incorrect probe comparison stimuli are enclosed within parentheses.

TASK	SAMPLE	COMPARISONS CORRECT	COMPARISONS INCORRECT
AC1	"LAMBDA"	λ	ξ, γ
DC1	φ	λ	ξ, γ
AD1(PROBE)	"LAMBDA"	(φ)	(σ, δ)
AC2	"XI"	ξ	λ, γ
DC2	σ	ξ	λ, γ
AD2(PROBE)	"XI"	(σ)	(φ, δ)
AC3	"GAMMA"	γ	λ, ξ
DC3	δ	γ	λ, ξ
AD3(PROBE)	"GAMMA"	(δ)	(φ, σ)

Fig. 7. Each line depicts a trial type—sample stimulus, with correct and incorrect comparisons—used in Final Test Phase C3 (Table 2). Each three-stage AD probe is just below its two prerequisite AC and DC baseline trial types. The numbers 1, 2, and 3 refer to the auditory stimuli shown from top to bottom in Figure 2 and to the visual stimuli shown from left to right. Because probe trials were never reinforced, correct and incorrect probe comparison stimuli are enclosed within parentheses.

Three-stage equivalence. Phases C3, C4, and C5 in Table 2 summarize the three-stage probes—those requiring subjects to have learned conditional relations within three sets of stimuli. AD probes (Phase C3) involved stimulus sets A, C, and D (Figure 2). Because emergence of the AD relations required subjects to have learned only AC and DC, AD probes were inserted into a baseline of only AC and DC trials. Figure 7 shows the trial types in the three-stage AD tests, with each AD probe below its prerequisite AC and DC trial types. AD tests had 90 trials, 10 of each trial type (60 baseline and 30 probe trials).

The three-stage BC and CB probes (Phases C4 and C5) required subjects to have learned conditional relations within Stimulus Sets A, B, and C (Figure 2). Because the emergence of BC and CB demanded only AB and AC as prerequisites, Phases C4 and C5 were combined into a single test, with probes inserted into a baseline of AB and AC trials. Figure 8 shows the trial types, with each pair of BC and CB probes located below the prerequisite AB and AC trial types. These combined BC-CB tests had 120 trials, 10 of each type (60 baseline and 60 probe trials—30 BC and 30 CB probes).

Symmetry. In Phase C6 (Table 2), CD probes tested the DC relations for symmetry. Figure 9 shows the trial types in this symmetry test with CD probes inserted into a baseline that contained only DC trials. CD tests had 60 trials, 10 of each type (30 baseline and 30 probe trials).

Oral naming. The final tests, Phase C7 of Table 2, were oral naming, in which the subjects named the B-, C-, and D-stimuli aloud. Naming tests had 90 trials, with 10 presentations of each visual stimulus.

Test sequence. Table 3 shows the sequence of tests for each subject. Six went through Phase C with minor variations in the order

TASK	SAMPLE	COMPARISONS CORRECT	COMPARISONS INCORRECT
AB1	"LAMBDA"	Λ	Ξ, Γ
AC1	"LAMBDA"	λ	ξ, γ
BC1(PROBE)	Λ	(λ)	(ξ, γ)
CB1(PROBE)	λ	(Λ)	(Ξ, Γ)
AB2	"XI"	Ξ	Λ, Γ
AC2	"XI"	ξ	λ, γ
BC2(PROBE)	Ξ	(ξ)	(λ, γ)
CB2(PROBE)	ξ	(Ξ)	(Λ, Γ)
AB3	"GAMMA"	Γ	Λ, Ξ
AC3	"GAMMA"	γ	λ, ξ
BC3(PROBE)	Γ	(γ)	(λ, ξ)
CB3(PROBE)	γ	(Γ)	(Λ, Ξ)

Fig. 8. Each line depicts a trial type—sample stimulus, with correct and incorrect comparisons—used in Final Test Phases C4 and C5 (Table 2), which were combined into a single test. Each pair of BC and CB probes is just below the two prerequisite AB and AC baseline trial types. The numbers 1, 2, and 3 refer to the auditory stimuli shown from top to bottom in Figure 2 and to the visual stimuli shown from left to right. Because probe trials were never reinforced, correct and incorrect probe comparison stimuli are enclosed within parentheses.

given in Table 2. Two subjects, J.O. and J.L., required major departures from this sequence; their tests and data will be noted separately.

TASK	SAMPLE	COMPARISONS CORRECT	COMPARISONS INCORRECT
DC1	Φ	λ	ξ, γ
CD1(PROBE)	λ	(Φ)	(σ, δ)
DC2	σ	ξ	λ, γ
CD2(PROBE)	ξ	(σ)	(Φ, δ)
DC3	δ	γ	λ, ξ
CD3(PROBE)	γ	(δ)	(Φ, σ)

Fig. 9. Each line depicts a trial type—sample stimulus, with correct and incorrect comparisons—used in Final Test Phase C6 (Table 2). Each symmetry probe is below its prerequisite DC baseline trial type. The numbers 1, 2, and 3 refer to the visual stimuli shown from left to right in Figure 2. Because probe trials were never reinforced, correct and incorrect probe comparison stimuli are enclosed within parentheses.

Table 3

Sequence of Tests given to Each Subject

	Four-stage		*Three stage*		*Symmetry*	
Subject	*DB*	*BD*	*AD*	*BC-CB*	*CD*	*Naming*
E.W.	1	2	3	4	–	5
A.D.	1	2	3	4	5	–
D.B.	2	1	3	4	5	6
E.M.	1	2	4	5	–	3
I.C.	1	2	5	4	6	3
D.W.	1	2	3	4	5	6
J.O.	1	2	(see text)			
J.L.	1		2	3	(see text)	

RESULTS

Each row of bars in Figure 10 represents the matching-to-sample test and baseline scores for one of six subjects. Column headings (BD . . . DC) refer to the relations diagrammed in Figure 2. Although there had been no differential reinforcement on probe trials, responses indicative of equivalence relations were arbitrarily designated as correct in summarizing probe-trial results. The three bars at the extreme right side of each row combine the scores for each indicated kind of baseline trial across all tests that included it (Figures 6 to 9). The first six bars (only five for Subjects E.W. and E.M.) each represent a probe-trial score for one four-stage, three-stage, or symmetry test.

Baseline performances were uniformly excellent. A finer analysis of each baseline relation into its three individual trial types (for example, those listed in the lower section of Figure 5) showed that these children rarely made as many as two errors, and never more, in 10 baseline trials of any type. Inserting unfamiliar probe trials among the explicitly taught trial types did not disrupt the baselines.

The children also behaved remarkably consistently in probe trials. In four-stage probes for the relations BD and DB, only Subject I.C. selected an "incorrect" letter is as many as four out of 30 probe trials. All subjects distributed their errors relatively evenly among the trial types. In three-stage AD tests, five of the children ranged from zero to two errors in the 30 probe trials, but Subject I.C. did give an indication here of incompletely established equivalence relations. Although he made only five errors in the 30 AD probes, four came in the 10-trial set with "gamma" as the dictated sample, and three of those were selections of

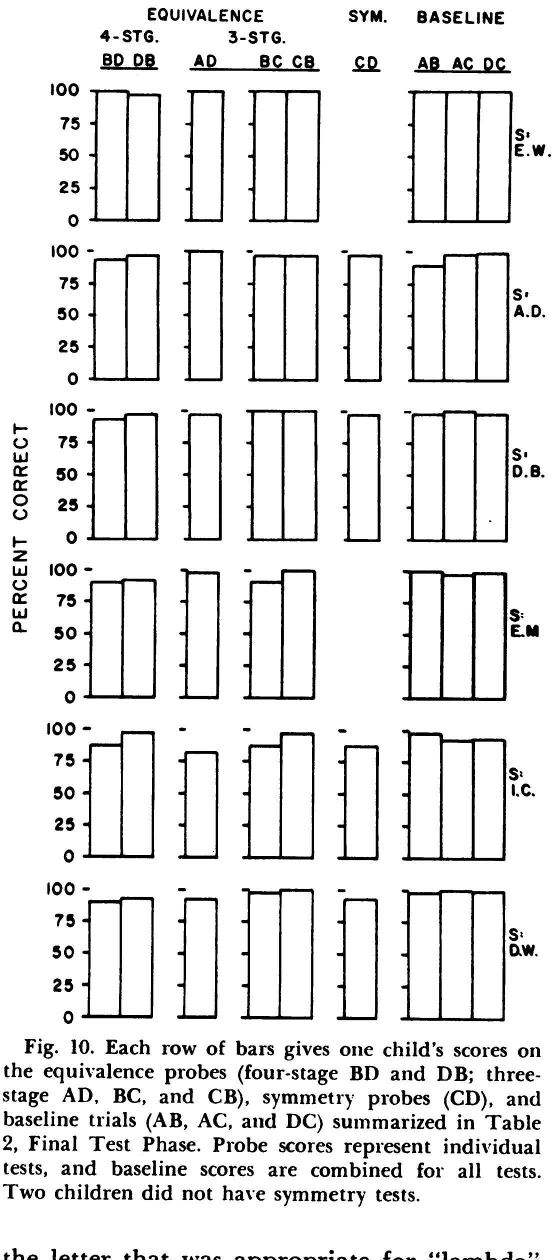

Fig. 10. Each row of bars gives one child's scores on the equivalence probes (four-stage BD and DB; three-stage AD, BC, and CB), symmetry probes (CD), and baseline trials (AB, AC, and DC) summarized in Table 2, Final Test Phase. Probe scores represent individual tests, and baseline scores are combined for all tests. Two children did not have symmetry tests.

the letter that was appropriate for "lambda" (Figure 7). In the three-stage probes for the BC and CB relations, and in the CD symmetry probes, Subject I.C. was again the only one of these six children to make as many as four errors in any set of 30 trials, but these did not concentrate in any particular trial type.

When asked to identify the letters, the subjects named them almost perfectly, none making more than two errors in the 90-trial test. Within each set of letters shown in Figure 2 (B, C, and D), the children consistently called the one at the left, "lambda," the center one, "xi," and the letter at the right, "gamma." (The correlation between class membership and position of the letters in Figure 2 is simply an expository device; during naming tests the letter was always in the center window, and during matching-to-sample tests the positions of comparisons letters varied from trial to trial.) This consistency was perfect in trials with the D-stimuli, even though the children's selection of D-stimuli conditionally upon dictated names (AD) had never been reinforced.

Subject E.W., however, differed from all the others in his manner of responding to D-stimuli. Table 4 reproduces the first half (45 trials) of his oral-naming test, including the instructions. (Unlike the actual transcript, Table 4 does identify the stimulus on each trial.) Although Subject E.W. gave all the expected names, his responses indicated that the D-stimuli differed from the B- and C-stimuli; it is possible to scan Table 4, even without observing which trials had D-stimuli, and to pick out the D-trials. On Trial 2, the first presentation of a letter from Set D, the child expressed doubt, not even venturing a guess until prompted. Although later responses were not as lengthy, he expressed some doubt about the name of each subsequent D-stimulus. In contrast, with the exception only of Trial 15, he named B- and C-stimuli without qualification. In Trials 40 and 41, the child even emphasized the contrast himself. During the second half of the test, 45 more trials, Subject E.W. continued consistently to indicate uncertainty in naming the D-stimuli.

If the four-member classes, revealed by the four-stage BD and DB tests, arose through equivalence relations, the data must also possess certain internal consistencies. Reference to Figure 2 will help clarify the prerequisites for four-member classes. Given the explicitly taught relations, AB, AC, and DC, the following transitivity paradigm (TR1) constituted the simplest route for the emergence of the DB relations, in which the children matched Set-D samples to Set-B comparisons:

(TR1) If DC and CB, then DB.

Any child able to do DB must also have been capable of DC and CB. The explicitly taught DC relations were easily verified in the baseline, but CB, never explicitly taught, had to

Table 4

Subject E.W.'s responses on his first 45 oral-naming trials, including the pretest interaction between Subject (S) and Experimenter (E). Trials are listed consecutively, and the stimulus is identified with reference to its portrayal in Figure 2 as the left (1), center (2), or right (3) member of Set B, C, or D.

S: What's it going to be like?
E: O.K. Now this time, ----, there aren't going to be any pennies.
S: Yes.
E: No pennies this time.
S: Yes.
E: I'd like you to keep your hands in your lap like they are.
S: Mmmm.
E: And I don't want you to push anything, O.K.?
S: O.K.
E: Now when something comes on, you tell me what it is that you see.
S: O.K.
E: O.K.?
S: O.K.
E: Now remember, don't push anything; you just tell me what it is that you see.
S: O.K.

Trial	*Stimulus*	*Responses*
1	1C	S: ....LAMBDA. E: You don'—You don't have to yell into the microphone like that. It'll—it'll hear you alright, O.K.? S: O.K.
2	3D	S: .... Hmmmmmmmmmm mmm. Oh, somethin' like somethin', but—I don't know. It's kinda—Well I—don't really have any idea—One little idea, but I don't know what it is. E: What's that? S: Well,—I have an—a slight idea that it might be GAMMA.
3	2D	S: Hmmmmmmmmmmmm. I think that might be XI.
4	3C	S: That's GAMMA.
5	2B	S: That's XI.
6	3B	S: That's GAMMA.
7	1D	S: Hmmmmmmmmmmmm. – I think that might be LAMBDA. Or somethin' like it.
8	1B	S: That's LAMBDA.
9	2C	S: THAT'S XI.
10	1C	S: That's LAMBDA.
11	2B	S: That's XI.
12	3C	S: That's—GAMMA.
13	3B	S: That's GAMMA.
14	2D	S: I—I have a slight feeling that might be XI.
15	2C	S: I think that's XI.
16	3D	S: It might be GAMMA, but I'm not too sure.
17	1D	S: I—I thin'—It might be LAMBDA, but I'm not sure.
18	1B	S: That's LAMBDA.
19	1D	S: I like the LAMBDA, but I'm not too sure.
20	1C	S: That's LAMBDA.
21	2D	S: That might be XI.
22	2B	S: That's XI.
23	1B	S: That's LAMBDA.
24	3B	S: That's GAMMA.
25	2C	S: That's XI.
26	3C	S: That's GAMMA.
27	3D	S: Mmmmmm. I have a sl- slight feeling that just might be GAMMA.
28	2C	S: XI.
29	1D	S: It might be LAMBDA, but I'm not sure.
30	2B	S: XI.
31	1C	S: LAMBDA.
32	3C	S: GAMMA.
33	1B	S: LAMBDA.
34	3B	S: GAMMA.
35	2D	S: XI. ...... I think.
36	3D	S: Mmmmmm, I think that might be GAMMA, but I'm not too sure.
37	1B	S: LAMBDA.
38	2D	S: Maybe XI.
39	1D	S: Like LAMBDA, but I'm not too sure.
40	3D	S: That might be GAMMA.
41	3C	S: That *is* GAMMA.
42	2C	S: That's XI.
43	1C	S: LAMBDA.
44	3B	S: GAMMA.
45	2B	S: XI.

be tested. All six children who were able to do DB performed well in their CB tests.

The CB relations, in which the children matched Set-C samples to Set-B comparisons, could have emerged through the transitivity paradigm:

(TR2) If CA and AB, then CB.

Any child able to do CB must also have been capable of CA and AB. Again, the baselines verified the AB relations. The CA relations could have emerged through symmetry paradigm SY1:

(SY1) If AC, then CA.

The current procedures precluded direct tests of symmetry in the auditory-visual relations, AB, AC, and AD, because the children would have been required to match visual samples to auditory comparisons; presenting several auditory comparisons simultaneously was not feasible. The CA relations, indicative of AC symmetry, were therefore not tested.

The other four-stage test, BD, called for different prerequisites. Matching the Set-B samples to Set-D comparisons could have emerged by transitivity through:

(TR3) If BC and CD, then BD.

Any child able to do BD must also have been capable of BC and CD, although neither had been explicitly taught. All six children who were able to do BD performed well also in their BC and CD tests. The BC relations could themselves have emerged by transitivity:

(TR4) If BA and AC, then BC.

Here, the baselines verified the AC relations. BA would have been indicative of AB symmetry:

(SY2) If AB, then BA.

Again, however, auditory-visual symmetry could not be tested directly.

Emergence of the second BD prerequisite, the CD relations (see TR3), proved the visual-visual DC relations to be symmetric:

(SY3) If DC, then CD.

CD relations, indicative of DC symmetry, were prerequisite not only for the four-stage BD relations (by TR3), but for the three-stage AD relations also. The simplest route for the emergence of the AD relations was:

(TR5) If AC and CD, then AD.

Baseline performances verified the AC relations, and children who proved capable of both AD and BD did well also in the CD test (two did not have the CD test; see Figure 10).

All children, therefore, whose DB and BD performances signified the establishment of four-member stimulus classes also provided the necessary internal consistences in their other data. The prerequisite lower-level relations were intact.

Subjects J.L. and J.O. failed to replicate the other six children's results, their BD and DB performances revealing the absence of the transitive properties (TR3 and TR1) required by four-stage equivalence relations. Might an absence of lower-level symmetry or transitivity have been responsible for the failure of these higher-level relations to emerge? Additional tests given to the two children who did not exhibit four-member stimulus classes are summarized sequentially in Table 5.

Subject J.L., whose failure to confirm four-stage equivalence occurred before the desirability of systematic followup had become evident, did not receive all of the tests needed for a complete evaluation of the relations between higher- and lower-level performances. The upper section of Table 5 shows that he scored only 47% in his first four-stage DB test, and 40% when tested for the three-stage AD relations (Tests 1 and 2). Then, he achieved 97 and 93% in the other three-stage probes, BC and CB (Test 3). Repetition of the four-stage DB probes (Test 4) yielded a score of only 17%. Subject J.L. failed to demonstrate four-member classes or three-member ACD classes (denoted by the lower triangle in Figure 2) but did substantiate the formation of three-member ABC classes (the upper triangle in Figure 2). Strong DC relations in the baseline (righthand column of Table 5) and solid CB performances (Test 3) showed the prerequisites for four-stage DB equivalence (see TR1) to be intact. Subject J.L.'s followup tests, therefore, did not supply any obvious explanation for the failure of DB to emerge. Unfortunately, the other four-stage relations, BD, were not evaluated, nor were the CD relations, absence of which would have clarified the failure of AD (see TR5) to emerge.

The other subject, J.O., had a more adequate series of followup tests (lower section of Table 5). His first four-stage test yielded 83% correct; he gave no more than two incorrect responses in 10 opportunities with any Set-D sample. When BD was probed, however, he scored only 50% (Test 2). Then, in two more DB tests, his score dropped to 57 and 60% from its initially high level. Whether the BD test had somehow degraded the DB performance, or whether DB would have deteriorated anyway cannot be determined. Nevertheless, Subject J.O.'s first four tests provided no convincing evidence of four-member stimulus classes. A series of tests then undertook to determine whether these failures to document four-stage equivalence were accompanied by the absence of one or more of the prerequisite lower-level relations.

AD probes came next (Test 5), and the low score of 60% indicated that the three-member ACD classes (lower triangle in Figure 2) had not formed. Because the auditory-visual AD probes might have helped to reinstate the DB relations, DB was tested once again; the out-

Table 5

Subjects J.L. and J.O. Scores (percentage correct) during successive tests. Column headings refer to relations in Figure 2. 'NAM' denotes oral naming tests with the Set-B, C, and D stimuli. Starting with Subject J.O.'s Test 8, baseline and probe trials were unreinforced.

		Equivalence					*Sym*	*Nam*	*Baseline*		
		Four-stage		*Three-stage*							
Subject	*Test Number*	*BD*	*DB*	*AD*	*BC*	*CB*	*CD*	*B,C,D*	*AB*	*AC*	*DC*
J.L.	1	—	47	—	—	—	—	—	100	100	97
"	2	—	—	40	—	—	—	—	—	100	97
"	3	—	—	—	97	93	—	—	100	100	—
"	4	—	17	—	—	—	—	—	97	100	100
J.O.	1	—	83	—	—	—	—	—	100	97	100
"	2	50	—	—	—	—	—	—	100	87	97
"	3	—	57	—	—	—	—	—	100	100	90
"	4	—	60	—	—	—	—	—	100	97	100
"	5	—	—	60	—	—	—	—	—	100	100
"	6	—	43	—	—	—	—	—	100	100	97
"	7	—	—	50	—	—	—	—	—	97	100
"	*8*	—	50	—	—	—	—	—	97	93	100
"	9	—	—	—	57	30	—	—	97	90	—
"	10	—	—	—	—	—	37	—	—	—	67
"	11	—	—	—	—	—	—	90	—	—	—
"	12	—	10	—	—	—	—	—	83	97	93
"	13	—	—	—	—	—	—	97	—	—	—

come this time was only 43% (Test 6). Repetition of the AD probe (Test 7) then yielded another low score (50%).

Since reinforcement occasionally followed baseline trials, but never probe trials, the deterioration of Subject J.O.'s performances after his first test might have come about because he had discriminated these contingencies. Subsequent tests therefore omitted reinforcement after all trials, probe and baseline. The child was told, "No pennies this time—we'll do colors with pennies later." To accustom him to the extinction procedure, he was given sets of unreinforced baseline trials, with no probes, until he maintained the requisite accuracy. Then, with baseline accuracies remaining high during his first extinction test (Test 8), the DB probes again gave a score of only 50%.

BC and CB, the next probes (Test 9), yielded 57 and 30% correct, indicating that three-member ABC classes (upper triangle in Figure 2) also had not formed. The child's low CB score could have accounted for the failure of the four-stage DB relations to emerge (see TR1). His low BC score revealed the absence of prerequisites for BD, the other four-stage relations (see TR3). Then, a score of only 37% in Test 10's CD probes revealed that the DC relations were not symmetric. (CD probe trials actually disrupted the child's DC baseline—Table 5's righthand column shows no other test in which the DC baseline fell below 90%.) The absence of DC symmetry (see SY3) could by itself have accounted for the failures both of the four-stage BD and the three-stage AD relations to emerge (transitivity paradigms TR3 and TR5, respectively).

Subject J.O.'s poor performances on the four-stage DB and BD tests and on the three-stage AD test had therefore proven consistent with the absence of one or more prerequisites. Indeed, none of the testable prerequisites had emerged. When tested for naming, however, he scored 90% (Test 11). Since naming the letters might have helped establish the equivalence relations, he was given the DB test once again; this time, he achieved only 10% (Test 12). A repetition of naming (Test 13) then yielded another high score (97%).

DISCUSSION

Having learned three sets of conditional discriminations, AB, AC, and DC (solid arrows in Figure 2), six of the eight children proved capable of six new sets of conditional discriminations they had not been explicitly taught: DB, BD, AD, BC, CB, and CD (broken arrows in Figure 2). Their BD and DB performances documented the emergence of three four-member stimulus classes, and all six children demonstrated the necessary lower-level transitive and symmetric properties in the relations among class members. Without such consistency in the prerequisites, the stimulus-class formulation would need major modification if it was to remain useful. When asked what the letters were, the children called them "lambda," "xi," and "gamma," in accord with class memberships the procedures had established. Each "if . . . then" relation was also an equivalence relation; the conditional-discrimination procedures had generated matching-to-sample performances.

The efficiency of the four-stage equivalence paradigm in generating new performances endows it with considerable practicality; the very process of equivalence testing yields a remarkable teaching spinoff. After the children were explicitly taught nine different sample-comparison relations (three represented by each solid arrow in Figure 2), 18 new stimulus relations (three represented by each broken arrow in Figure 2), and nine new oral naming relations emerged. The ratio of emergent to directly taught performances was 27/9.

Teaching efficiency increased as the stimulus classes grew in size. For example, the three-stage paradigm denoted by ABC in Figure 2 depicts six directly established stimulus relations, three represented by AB and three by AC. Six more stimulus relations (three in BC and three in CB) and six oral naming relations (B- and C-naming) emerged from this teaching; emergent performances exceeded those directly taught by a factor of two. Then, building the four-stage paradigm by teaching the children to match three samples from Set D to comparisons in Set C added 12 more stimulus relations (three each in AD, CD, DB, and BD) and three additional oral naming relations (D-naming) to their repertoires; three directly taught performances now generated 15 emergent performances, a factor of five. This 2.5-fold increase in teaching efficiency with the addition of a single member to each stimulus class hints at the potentially explosive nature of the process.

The performances involved here are in principle far from trivial. Matching auditory to visual stimuli (AB, AC, and AD) can represent simple auditory comprehension—understanding spoken words in reference to text; matching visual stimuli to each other (BD, DB, BC, CB, CD, and DC) can constitute simple reading comprehension—understanding text in reference to other objects; naming textual stimuli aloud can be simple oral reading. Most texts designed to evaluate aspects of reading include similar tasks. Nevertheless, formal resemblances between conditional discrimination and reading do not prove one relevant to the other. The establishment of stimulus classes does prove this relevance. Pointing to a picture in response to a printed word denotes reading comprehension only if the word and picture are related by equivalence and not merely by conditionality. Stimulus classes formed by a network of equivalence relations establish a basis for referential meaning. The equivalence paradigm provides exactly the test that is needed to determine whether or not a particular conditional discrimination involves semantic relations.

Linguistic analysis has challenged functional behavioral analysis to account for new behavior that has no apparent reinforcement history (e.g., Chomsky, 1965; Fodor, Bever, & Garrett, 1974). The equivalence paradigm takes a short step in this direction by specifying procedures for generating new and seemingly unreinforced matching to sample and oral naming. In revealing a class whose members are related by equivalence, the paradigm also exposes a source of reinforcement for the new behavior. By definition, the existence of a class of equivalent stimuli permits any variable that affects one member of the class to affect all members. Even when stimuli bear no physical resemblance to each other, their inclusion within a class provides a route for extending the influence of reinforcement and other variables. Direct reinforcement of the AB, AC, and DC relations (Figure 2) extends also to all of the other possible relations within each four-member class. It is therefore not correct to assume that the new matching and naming performances emerged without a reinforcement history.

The children's ability to name the letters in Sets B and C confirmed earlier experiments that used other kinds of stimuli (Sidman, 1971; Sidman & Cresson, 1973; Sidman, Cresson, & Willson-Morris, 1974). Their consistent naming of the Set-D letters was of special interest since they had never been taught explicitly to select those letters conditionally upon dictated-name samples. Two subjects (E.M. and I.C., Table 3) even did the naming test before their ability to match Set-D stimuli to dictated names was tested. The production of Set-D names that were consistent with classes revealed in the matching-to-sample tests raised the possibility that naming might have been needed to mediate the emergent conditional relations. An earlier study, in which a few new conditional relations emerged before subjects were able to name the stimuli, suggested that stimulus equivalence was independent of naming (Sidman, Cresson, & Willson-Morris, 1974). Here, Subject E.W. gave all the stimuli names that were consistent with their class membership, but his hesitations and expressions of doubt (Table 4) indicated strongly that although he was capable of naming the Set-D letters, he had never done so until the naming test. The new conditional discriminations involving the D-stimuli emerged before he had ever applied names to those letters.

Subject E.W. was the only one of the eight who yielded such a finding, but his demonstration that the stimulus classes could form in the absence of naming cannot be dismissed. Although naming, when it occurs, may indeed facilitate equivalence relations (Goldiamond, 1962), Subject E.W.'s responses argue strongly against the necessity of such a role. Given that naming is not necessary for class formation, the likelihood and nature of a facilitative role remain matters for experimental study.

Table 4's naming transcription indicated that stimulus classes could emerge via equivalence relations even before a subject had applied a consistent name to each member of a class; naming was not necessary. Subject J.O.'s naming tests (Table 5) showed that consistent naming was not sufficient to establish stimulus classes via equivalence relations. He applied "lambda," "xi," and "gamma" to the letters of Sets B, C, and D in accord with the very four-member classes that the equivalence tests had failed to substantiate. He named the Set-D letters consistently in spite of his inability to match those letters to their dictated names. A response that is common to several stimuli may define a class but does not by itself establish equivalence relations among the class members. Upon reflection, the dichotomy shown here between classes defined by naming and classes defined by equivalence is not surprising; the relation, "is the name of," does not possess the defining properties of an equivalence relation (Table 1).

It might have been tempting to view naming as an indirect test of symmetry in the auditory-visual relations, AB, AC, and AD. Instead of presenting dictated-name comparisons, oral naming tests permitted the children to produce the names themselves in response to the printed letters. It seems reasonable to presume that children who could name printed letters aloud would also be able to select those names when they heard them spoken by someone else. Subject J.O., however, proved this reasonable presumption to be incorrect. His accurate production of the Set-D letter names, after showing himself unable to match those letters to their dictated names, demonstrated that emergent naming did not constitute a valid test of auditory-visual symmetry. Again, a possible relation between naming and equivalence proved illusory.

Formal resemblances exist between the accounts of stimulus equivalence arising from conditional-discrimination and paired-associate procedures. The paired-associate literature refers to symmetry as "backward association" (e.g., Ekstrand, 1966), to transitivity as "chaining," and to equivalence paradigms A and B (Table 1) as "response equivalence" and "stimulus equivalence," respectively (e.g., Jenkins, 1963). A thorough comparison of the formulation derived from the paired-associate tradition and the one proposed here must await a more appropriate vehicle, but at least one fundamental distinction that the terminological similarities might mask is relevant here. Paired-associate methods, which often require a subject to respond differentially to each stimulus by producing its "name," have led to widespread acceptance of response mediation as the mechanism responsible for the establishment of equivalence relations (e.g., Jenkins, 1963, 1965; Jenkins & Palermo, 1964). Although it is clear that differential responses can mediate the emergence of new stimulus relations, the successful use of conditional discriminations to generate equivalences raises considerable doubt about the necessity for postulating the existence of mediating responses. Conditional discrimination requires no differential responses to individual stimuli; the only necessary overt response is pointing, or touching, which is the same for all sample and comparison stimuli. It is sometimes said that the subject responds differentially to each sample by choosing a particular comparison, but this "choosing" response can be defined only with reference to a stimulus (Sidman, 1978). Conditional discrimination therefore involves relations between stimuli. It is not possible to obviate this conclusion, as some have attempted, by reserving the term, stimulus, for sample stimuli, and applying the term, response, to comparison stimuli. Thus, in addition to evidence cited above that shows naming to be neither necessary nor sufficient for generating stimulus equivalences, the very logic of the conditional-discrimination procedure suggests also that no other kind of mediating response need be postulated.

REFERENCES

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

Ekstrand, B. R. Backward associations. *Psychological Bulletin*, 1966, **65**, 50-64.

Fletcher, F. G., Stoddard, L. T., & Sidman, M. A system for unlimited repetitive presentation of auditory stimuli. *Journal of Experimental Child Psychology*, 1971, **11**, 165-169.

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

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

Jenkins, J. J. Mediated associations: Paradigms and situations. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963.

Jenkins, J. J. Mediation theory and grammatical behavior. In S. Rosenberg (Ed.), *Directions in psycholinguistics*. New York: Macmillan, 1965.

Jenkins, J. J., & Palermo, D. S. Mediation processes and the acquisition of linguistic structure. In U. Bellugi & R. Brown (Eds.), The Acquisition of Language. *Monographs of the Society for Research in Child Development*, 1964, **29** (Serial No. 92).

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

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

Sidman, M. Remarks. *Behaviorism*, 1978, **6**, 265-268.

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

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

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.

Received April 17, 1981
Final acceptance August 6, 1981