JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1988, 49, 95-115 NUMBER 1 (JANUARY)

ESTABLISHING AUDITORY STIMULUS CONTROL OVER AN EIGHT-MEMBER EQUIVALENCE CLASS VIA CONDITIONAL DISCRIMINATION PROCEDURES

RICHARD R. SAUNDERS, JUDITH WACHTER, AND JOSEPH E. SPRADLIN

BUREAU OF CHILD RESEARCH, UNIVERSITY OF KANSAS

Two eight-member equivalence classes of visual stimuli were established during three phases of a training program. In Phase 1, two training arrangements were compared. In one, 3 subjects were taught on different trials to select from a single pair of comparison stimuli (A_1, A_2) in response to eight sample stimuli that were trained in pairs (B_1, B_2; C_1, C_2; D_1, D_2; E_1, E_2). In the second arrangement, subjects were taught to select from four pairs of comparisons (B_1, B_2; C_1, C_2; D_3, D_2; E_2, E_2) in response to two samples (A_1, A_2). Training with the single pair of comparison stimuli resulted in the development of equivalence relations (B_1C_1, B_2C_2, D_1B_1, D_2B_2, B_1E_1, B_2E_2, C_1D_1, C_2D_2, C_1E_1, C_2E_2, D_1E_1, D_2E_2, and their reciprocals) between the sample stimuli without direct training of these relations. In the other training arrangement, these relations among the comparison stimuli developed in the performance of 1 subject only. In Phase 2, three new pairs of stimuli (F_1, F_2; G_1, G_2; H_1, H_2) were substituted for three of the original pairs (B_1, B_2; C_1, C_2; D_1, D_2) and the training arrangements for the groups were reversed. Following training, the performances that showed equivalence relations on the probes in the first phase also showed equivalence relations in the second phase. If such relations did not develop in the first phase, they did not do so in the second phase. In Phase 3, relations between stimuli across the two previous phases (e.g., B_1F_1, B_2F_2, B_1G_1, B_2H_2, C_1F_1, etc.) were investigated. The 4 subjects whose performances showed the development of these relations were taught to select one stimulus from each class (E_1 and E_2) in response to a verbal label (I_1 and I_2) and then were tested to see if the verbal label controlled responding to the remaining members of the class (e.g., I_1A_1, I_2A_2, I_1B_1, I_2B_2, etc.). For 3 subjects, this generalized control occurred; for the 4th, generalization occurred only after verbal training with a second pair of visual stimuli (F_1 and F_2). In retests several months later, these auditory-visual relations were found to be intact or, if not, were recovered without direct training.

Key words: stimulus classes, stimulus equivalence, auditory control, stimulus control, matching-to-sample, discrete-trial procedures, remembering, conditional discrimination, retarded children and adults

---

The literature on the development of stimulus classes has grown steadily since Sidman's (1971) initial demonstration of the phenomenon, showing that stimulus class development is highly replicable under a variety of training conditions (Dixon, 1978; Dixon & Spradlin, 1976; Lazar, 1977; Lazar, Davis-Lang, & Sanchez, 1984; Mackay & Sidman, 1984; Sidman & Cresson, 1973; Sidman, Cresson, & Willson-Morris, 1974; Sidman, Kirk, & Willson-Morris, 1985; Sidman & Tailby, 1982; Sidman, Willson-Morris, & Kirk, 1986; Spradlin, Cotter, & Baxley, 1973; Spradlin & Dixon, 1976; Spradlin & Saunders, 1986; Stromer & Osborne, 1982; Wetherby, Karlan & Spradlin, 1983). There have also been several efforts to organize and conceptually unify this area of research (Baer, 1982; Fields, Verhave, & Fath, 1984; Sidman, 1986; Sidman & Tailby, 1982).

Sidman and Tailby (1982) proposed an analysis of stimulus class development in terms of the logical relations of equivalence—reflexivity, symmetry, and transitivity. According to this analysis, an equivalence class is present when performance shows (a) reflexivity, which is demonstrated by generalized identity matching (if a subject is trained to match A_1 to A_1, A_2 to A_2, B_1 to B_1, and B_2 to B_2, then C_1 will be matched with C_1 and C_2 with C_2 without additional training); (b) symmetry of sample and comparison stimuli (if during training A_1,

---

Preparation of this article was supported in part by Federal NICHD grants 5-3P0HD02528, and 1-P01HD-18955 to the Bureau of Child Research, University of Kansas.

Readers who are interested in a detailed description of the procedures and results obtained from Subjects TB, LE, and BB may obtain a complete report by writing to the first author.

Requests for reprints should be sent to Richard R. Saunders, Bureau of Child Research, P.O. Box 738, Parsons, Kansas 67357.

as the sample, controls selections of B_1 as the comparison stimuli, and A_2 controls selections of B_2, then B_1 as the sample should control responses to A_1 as the comparison stimulus and B_2 should control selection of A_2 without additional training); and (c) transitivity (if during training, A_1 as the sample controls selections of B_1 as the comparison stimulus and A_2 as the sample controls selections of B_2, and if B_1 as the sample controls selections of C_1 and B_2 as the sample controls selections of C_2, then A_1 should control selections of C_1 and A_2 should control selections of C_2 without further training).

Analogously, if the relations A_1B_1 and A_2B_2 and then A_1C_1 and A_2C_2 are taught, the tests of B_1C_1 and B_2C_2 are simultaneous tests for transitivity and symmetry—a direct test for equivalence relations (Sidman & Tailby, 1982). With either of the above procedures for training and testing, when all of the necessary relations are demonstrated, the two stimulus classes consisting of A_1, B_1, and C_1 and A_2, B_2, and C_2, respectively, are referred to as equivalence classes.

Fields et al. (1984) analyzed a number of the variables involved in stimulus class development, including class size and the number of nodes involved in training. The term "node" was used to refer to the stimulus that relates directly to two or more stimuli. For example, a class of five stimuli might be established by relating four stimuli (B, C, D, and E) to a single stimulus or node (A) or it might be established by relating A to B, B to C, C to D, and D to E in which there would be three nodes—B, C, and D. Sidman et al. (1985) have also distinguished among these different relations using a slightly different terminology. For Sidman et al., the one-node arrangements would produce "three-stage" equivalence relations between B and C, C and D, and D and E. The three-node arrangement would produce a "five-stage" equivalence relation between A and E, a "four-stage" relation between A and D, and so forth. In each case, the number of stages refers to the number of stimulus pairs required (including the derived relation) to produce the derived relation. When a relation of N stages is demonstrated, then relations of $N - 1$ or $N - 2$ stages that are subordinate to the N-stage relation are also demonstrated.

Although Fields et al. (1984) and Sidman and Tailby (1982) recognized the possibility that relations between stimuli within a class may depend on the number of stimuli in the class and the number of nodes between stimuli, they did not indicate that the direction of training should make any difference in the acquisition of a stimulus class or of equivalence classes in particular. That is, as far as these conceptual analyses go, it should make no difference whether a subject is trained to select a single comparison stimulus in the presence of four different sample stimuli or to select four different comparison stimuli in the presence of a single sample stimulus. In both cases, the size of the class is five and the number of nodes is one.

However, Spradlin and Saunders (1986) found that training with four samples and one comparison stimulus (the training node) led to performances that showed the development of equivalence relations among the samples. The performances of other subjects trained with one sample (the training node) and four comparison stimuli did not show equivalence relations among the comparison stimuli. The study reported here investigated whether these differences could be replicated. Additionally, those subjects whose performances showed equivalence relations under either training method were exposed to training and testing conditions designed to compare performance on one-node and two-node relations and to investigate the generalization of auditory stimulus control over members of the equivalence class.

METHOD

General Organization of Training Sequences

Subjects

Six retarded adolescents and young adults participated in the experiment. The chronological age, sex, intelligence quotient, and diagnosis of each are shown in Table 1. All test scores were current within 2 years and were obtained with a Wechsler test of intelligence. All subjects had participated in a two-choice simple-discrimination study designed to replicate Touchette's (1971) results with procedures involving progressive delays of prompts. Although the delayed-prompt study was held in the same building, it was conducted with a

Table 1
Subject characteristics.

Subject	Age	Length of institutionalization	Sex	Measured I.Q.	Diagnosis
TB	23 years, 8 months	14 years, 10 months	M	62	Mild mental retardation; seizures; visual, speech, motor, and behavior dysfunction
BB	11 years, 4 months	10 months	M	70	Mild mental retardation; metabolic epileptic disorder, aggressive behavior
RF	16 years, 9 months	4 years, 9 months	M	63	Mild mental retardation related to "oddities of movement" and speech abnormalities
BD	20 years, 3 months	3 years, 10 months	M	67	Moderate mental retardation, organic brain syndrome present since birth
LE	19 years, 6 months	6 years, 4 months	F	67	Mild mental retardation associated with cerebral defect, visual impairment, aphasic speech and convulsive disorder
TK	12 years, 3 months	10 months	M	70	Borderline to mildly delayed intellectual and socioadaptive functioning

different apparatus in a different room. All subjects exhibited functional speech.

Apparatus and General Procedure

Subjects were seated at a large table that supported a stimulus display box, shown in Figure 1, that housed the computer equipment used to control the stimulus displays, response consequences, and so on, for the experiment. The front wall of the box that separated the subject and the computer equipment was 120 cm wide by 61 cm tall. Three display windows measuring 5 cm by 5 cm each and spaced 2.5 cm apart were mounted in the wall 105 cm above the floor. During the experimental sessions, an Apple® monochrome monitor was positioned behind the display windows. Stimuli were displayed on the monitor screen, using Apple's high resolution graphics, centered in the display windows. Under the row of windows was a wooden panel within which were mounted three plastic-cased, spring-loaded buttons, one under each window. Each button had an exposed diameter of 2.5 cm. To the subject's lower left was a plastic box into which washers or tokens could be dispensed through an opening 5 cm by 5 cm in the front wall of the display box.

An Apple IIe® computer was used, equipped with supplemental hardware for memory expansion, timing, auditory feedback to the subjects, input-output functions, and data storage. A DSI® tray feeder, mounted behind the lower hole and plastic box, was used to dispense the washers. A summary of each session's data was printed at the end of that session. The number of responses, the latency for pressing the sample button and the latency for pressing the choice button, the stimulus configuration, and the programmed consequences for each trial were recorded on floppy diskette for each subject.

To begin each session, and on each subsequent trial of a session, the sample stimulus was displayed in the center window. A press of the button under this window by the subject resulted in the concurrent display of two comparison stimuli in the peripheral windows, one to the left of the sample and one to the right.

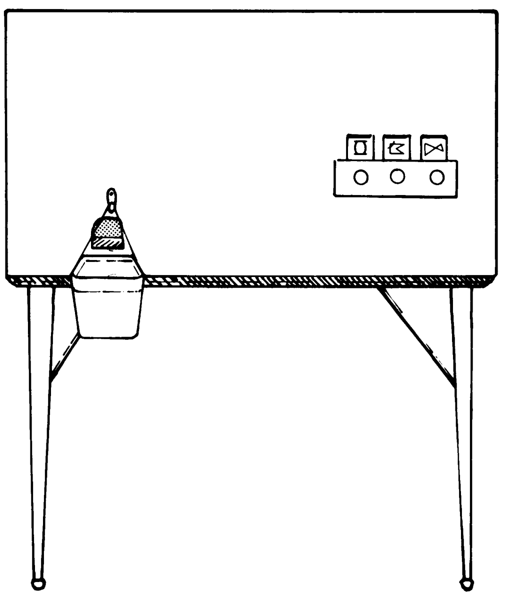

Fig. 1. A drawing of the apparatus, with stimulus-presentation windows, recessed buttons beneath the windows, and reinforcer aperture and cup to the left, as viewed by the subjects.

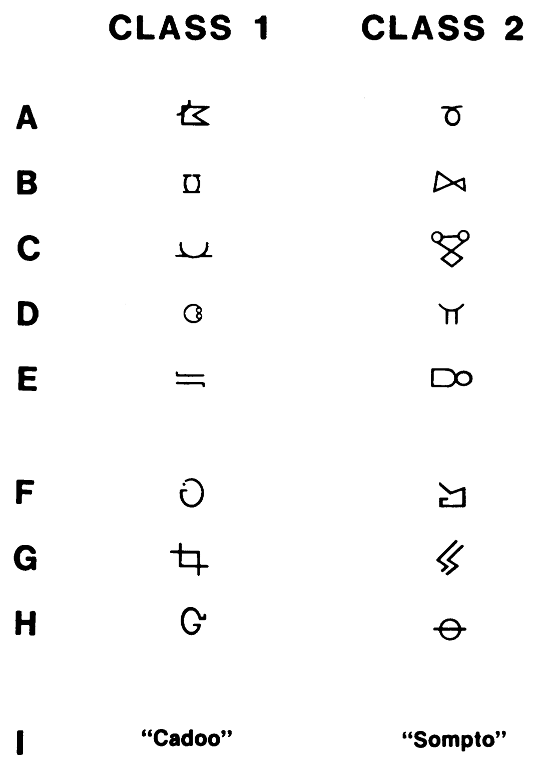

Fig. 2. The eight visual stimuli and one auditory stimulus in each of the two sets (potential classes) of stimuli used in the four phases of the experiment. The letters at the left indicate the order of introduction of each pair of stimuli, one stimulus from each set.

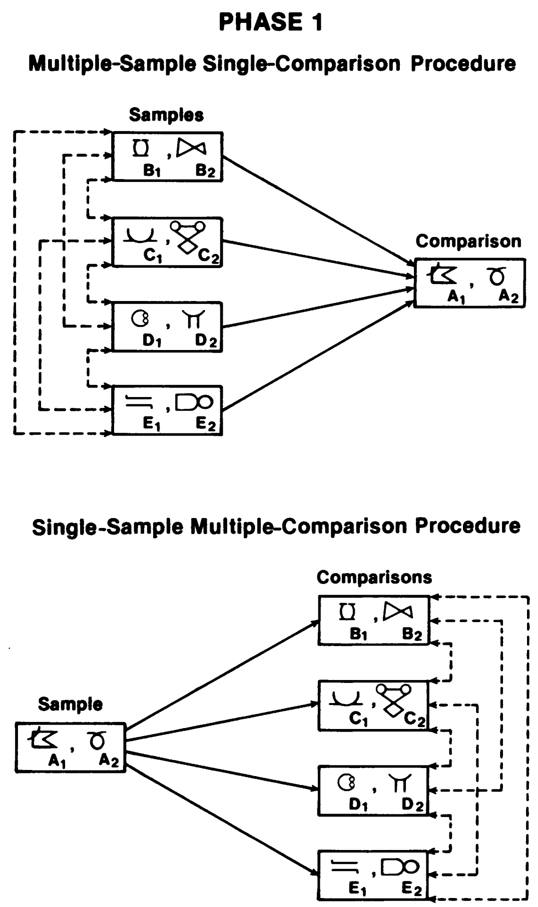

Fig. 3. A schematic of the training in Phase 1 of the experiment for those subjects trained with a multiple-sample single-comparison procedure and a schematic for those trained with the single-sample multiple-comparison procedure. Solid arrows indicate trained relations (e.g., B_1A_1, B_2A_2, and A_1B_1, A_2B_2) and dotted arrows indicate possible derived equivalence relations following training (e.g., B_1E_1, B_2E_2, E_1B_1, E_2B_2).

Further presses to the center button had no programmed consequences. A press of the button under either comparison window resulted in various programmed consequences, the temporary removal of all displays and the presentation of the next sample stimulus. During training, correct responses to the comparison stimuli resulted in a computer-generated auditory jingle and the delivery of a token as described below. These tokens were exchanged for pennies at the end of the session. Incorrect responses resulted in a brief buzzer sound. On probe trials (to be described later) and on any training trials with no programmed consequences, any comparison response resulted only in the removal of all displays and advancement to the next trial.

The experiment had four phases, involving a total of 16 visual stimuli and 2 auditory stimuli, as shown in Figure 2. The first phase consisted of two different training procedures to establish sets of conditional discriminations that would permit the tests for the formation of equivalence classes. In one procedure, a single pair of comparison stimuli was used with successive sets of two sample stimuli; in the other, a single pair of samples was used with successive sets of two comparison stimuli.

Figure 3 shows the training arrangements for the initial multiple-sample single-comparison and single-sample multiple-comparison procedures. As can be seen in the figures, there were two sets of stimuli, with five stimuli in each set (A_1, B_1, C_1, D_1, and E_1 in one set, and A_2, B_2, C_2, D_2, and E_2 in the second set). Although the training procedures involved a different direction of sample–comparison relations (e.g., BA vs. AB), the possible derived equivalence relations, marked by dashed lines, were identical for the tests of stimulus class development following the use of either procedure, except for the distinction between samples and comparisons.

Figure 3 shows that by training four conditional relations (BA, CA, DA, EA) or (AB, AC, AD, AE), the possibility of eight symmetry relations (A_1B_1, A_2B_2, A_1C_1, A_2C_2, A_1D_1, A_2D_2, A_1E_1, A_2E_2 or B_1A_1, B_2A_2, C_1A_1, C_2A_2, D_1A_1, D_2A_2, E_1A_1, E_2A_2) and 24 equivalence relations (B_1C_1, B_2C_2, B_1D_1, B_2D_2, B_1E_1, B_2E_2, C_1D_1, C_2D_2, C_1E_1, C_2E_2, D_1E_1, D_2E_2, and their reciprocals) occurs (the possible symmetry relations are not shown in the figure).

The second phase studied performance following training under the opposite condition from that used in original training. That is, a subject initially trained with the multiple-sample single-comparison procedure in Phase 1 was trained with the single-sample multiple-comparison procedure in Phase 2, and vice versa.

In Phase 2, three new stimuli in each set and two of the previous stimuli, A and E, were used in training (Figure 4). This additional training provided the possibility of six new symmetry relations and 24 new equivalence relations (AF, AG, AH, FG, FH, GH, and their reciprocals for each stimulus set). Phases 1 and 2 permit an evaluation of whether both multiple-sample single-comparison and single-sample multiple-comparison training would result in development of equivalence classes. However, if that development does occur in Phase 1, Phase 2 may be viewed not as establishing new classes but as simply adding new members to an existing class, because two of the members (A and E) of the initial classes are included in the set of five stimuli in each class during Phase 2.

Phase 3 involved combining Phase 1 and Phase 2 training trials in the same session. This combination provided a context for probes for relations between stimuli across the two phases (that is, tests between stimuli separated by two nodes). Possible derived relations in Phase 3 are shown in Figure 5.

Phase 1 and Phase 2 training via the common stimuli, A and E, provided the possibility of 36 new two-node equivalence class relations (BH, BG, BF, CH, CG, CF, DH, DG, and DF, and their reciprocals, HB, GB, FB, HC, GC, FC, HD, GD, and FD for each stimulus set). Thus, the combined procedures of Phase 1 and Phase 2 produced 14 relations trained directly (through seven conditional discriminations) with the possibility of 14 symmetry relations, 48 one-node equivalence relations, and 36 two-node equivalence relations emerging for each stimulus set. Tests in Phase 3 determined the degree to which the two-node relations had been established.

In Phase 4, subjects were trained to select the E_1 stimulus in response to I_1 (the auditory stimulus "cadoo") and the E_2 stimulus in response to the I_2 stimulus ("sompta"). The training relations and the seven possible un-

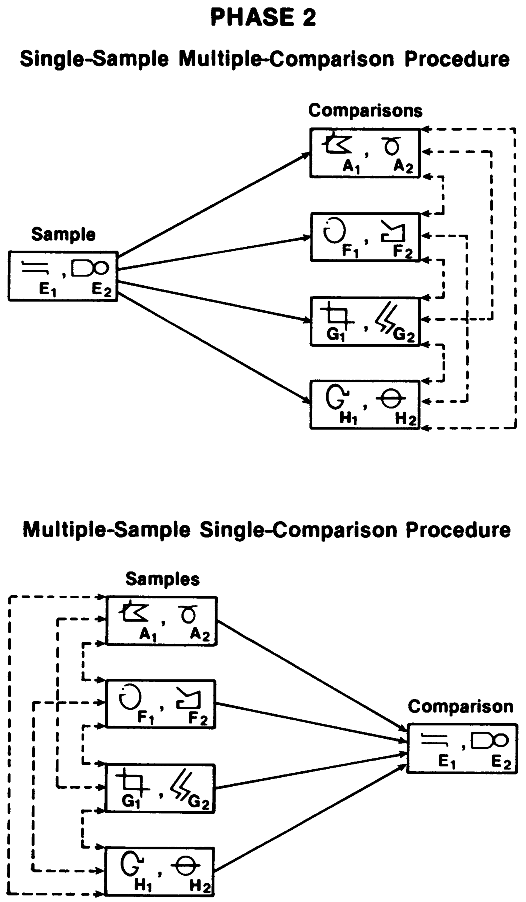

Fig. 4. A schematic of the training in Phase 2 of the experiment for those subjects trained with a single-sample multiple-comparison procedure and a schematic for those trained with a multiple-sample single-comparison procedure. Phase 2 training involved two pairs of stimuli used in Phase 1 and three new pairs of stimuli. Solid arrows indicate trained relations (e.g., E_1A_1, E_2A_2 and A_1E_1, A_2E_2) and dotted arrows indicate possible derived equivalence relations (e.g., A_1H_1, A_2H_2, H_1A_1, H_2A_2).

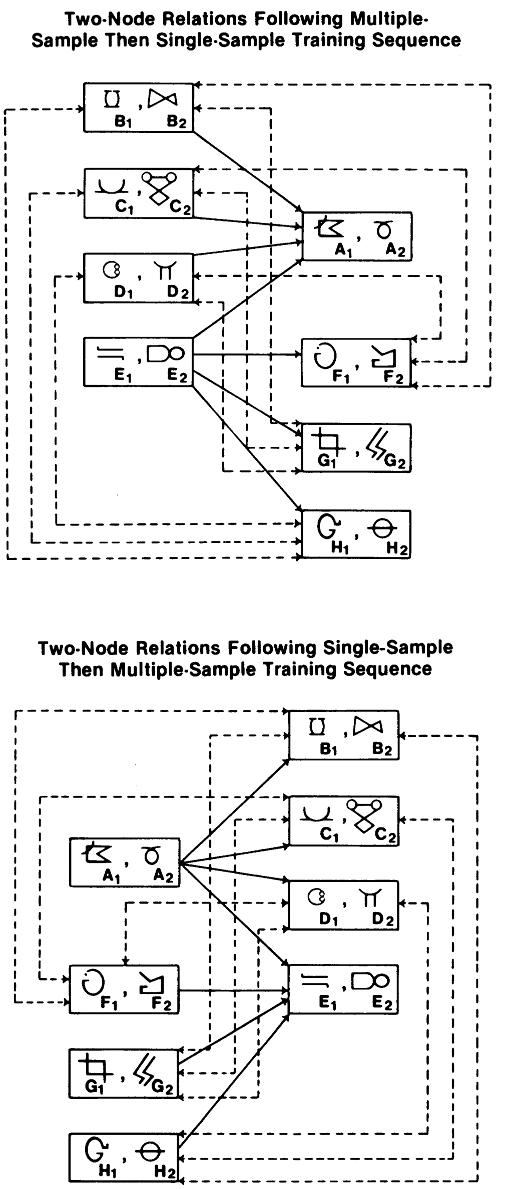

Fig. 5. A schematic showing all of the trained relations (solid arrows) among visual stimuli in Phase 3 for those subjects taught first with multiple samples and then with a single sample and a schematic for those trained with a single sample and then with multiple samples. The dotted arrows indicate the possible derived relations such as B_1H_1 and B_2H_2 related via two nodes, A_1 and E_1 or A_2 and E_2, respectively.

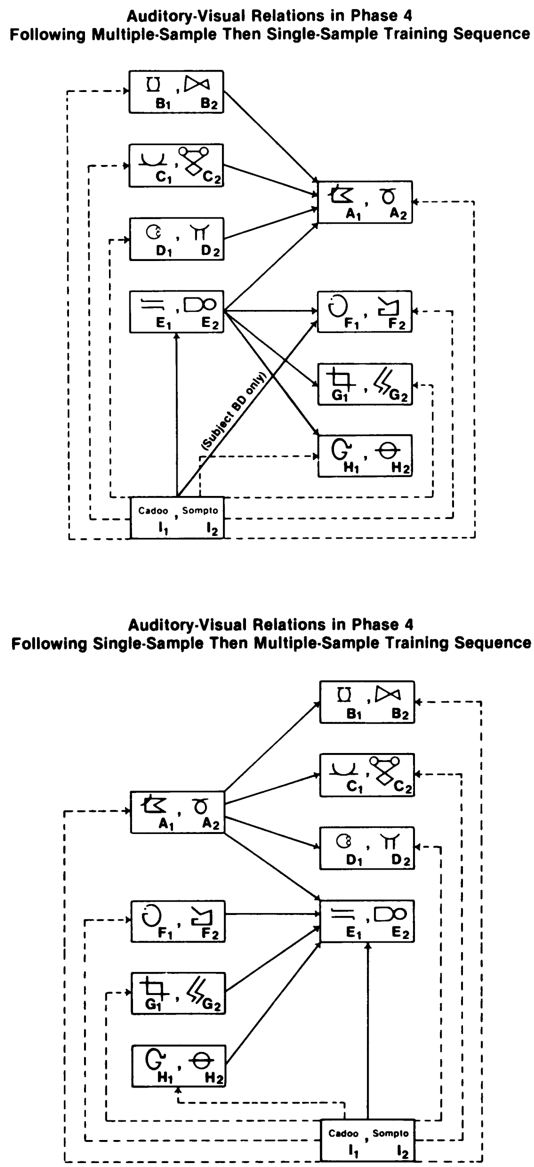

Fig. 6. A schematic showing all of the visual-visual and auditory-visual relations trained (solid arrows) through Phase 4 for those subjects first taught with multiple samples and then with a single sample prior to Phase 4 auditory-visual training and a schematic for those taught with a single sample and then with multiple samples prior to Phase 4. The dotted arrows indicate the possible derived auditory-visual relations (e.g., Cadoo—C_1 and Sompta—C_2) following Phase 4 training.

trained relations occurring as a function of training are shown in Figure 6. Four of the new untrained relations in each class are single-node relations (IA, IF, IG, IH), and three of the relations in each class are two-node relations (IB, IC, ID). Thus, a comparison of subjects' performance on one-node and two-node relations between stimuli within a class is possible.

In summary, the four-phase design of the experiment permitted the demonstration of an

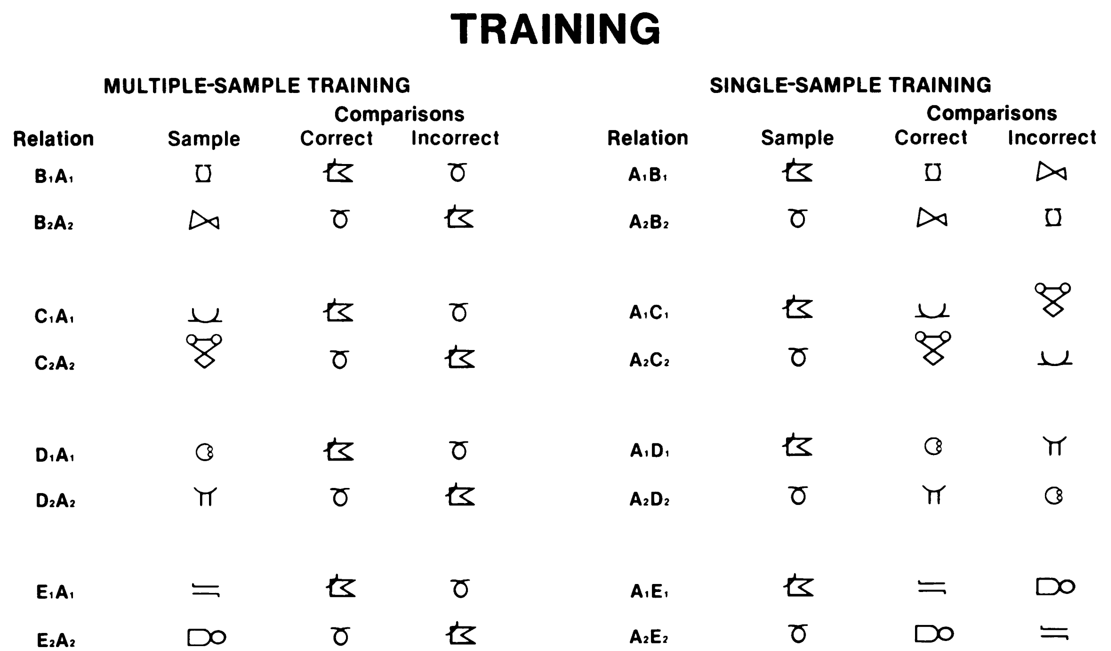

Fig. 7. The pairs of relations trained sequentially in Phase 1 under multiple-sample single-comparison procedures (e.g., B_1A_1 and B_2A_2) and single-sample multiple-comparison procedures (e.g., A_1B_1 and A_2B_2).

initial equivalence class among sample stimuli following multiple-sample single-comparison training or among comparison stimuli following single-sample multiple-comparison training, the evaluation of the development of two eight-member equivalence classes of visual stimuli based on both one-node and two-node relations, and the evaluation of the development of stimulus control over each eight-member equivalence class by an auditory stimulus.

Specific Procedures and Results

Pretraining. Prior to training the arbitrary conditional discriminations, training sessions of 32 trials of identity matching, with the visual stimuli shown in Figure 2, were conducted until a criterion of two consecutive sessions above 90% correct was met.

Phase 1 Procedure

The 6 subjects were unsystematically assigned to one of the two procedural groups described above and shown in Figure 3 (multiple-sample single-comparison or single-sample multiple-comparison multiple-comparison). By training the four conditional discriminations (BA, CA, DA, EA or AB, AC, AD, AE), eight symmetry relations and 24 equivalence (i.e., combined transitivity and symmetry) relations were possible within each of the two stimulus sets.

Baseline conditional discrimination training. Baseline training began with either the AB or BA conditional discrimination, depending on the group to which the subject was assigned. Subjects were instructed on the first two trials: "When the spool (flag) comes up, you push the button under the flag (spool). When the bow (apple) comes up you push the button under the apple (bow)." Instructions were repeated if subjects performance did not meet criterion after two sessions of training on the first conditional discrimination, but subjects received no further instructions after the criterion on this discrimination was met. Figure 7 shows the specific discriminations trained in both procedures.

The accuracy criterion for this first stage and each subsequent stage of training on the conditional discriminations was two consecutive 32-trial sessions above 90% correct or one session at 100% correct. After performance met criterion on the BA or AB discrimination training, those discriminations were discontinued and training on the CA or AC discrimination was introduced. Once performance on the CA or AC discrimination met criterion, training on the DA or AD discrimination only

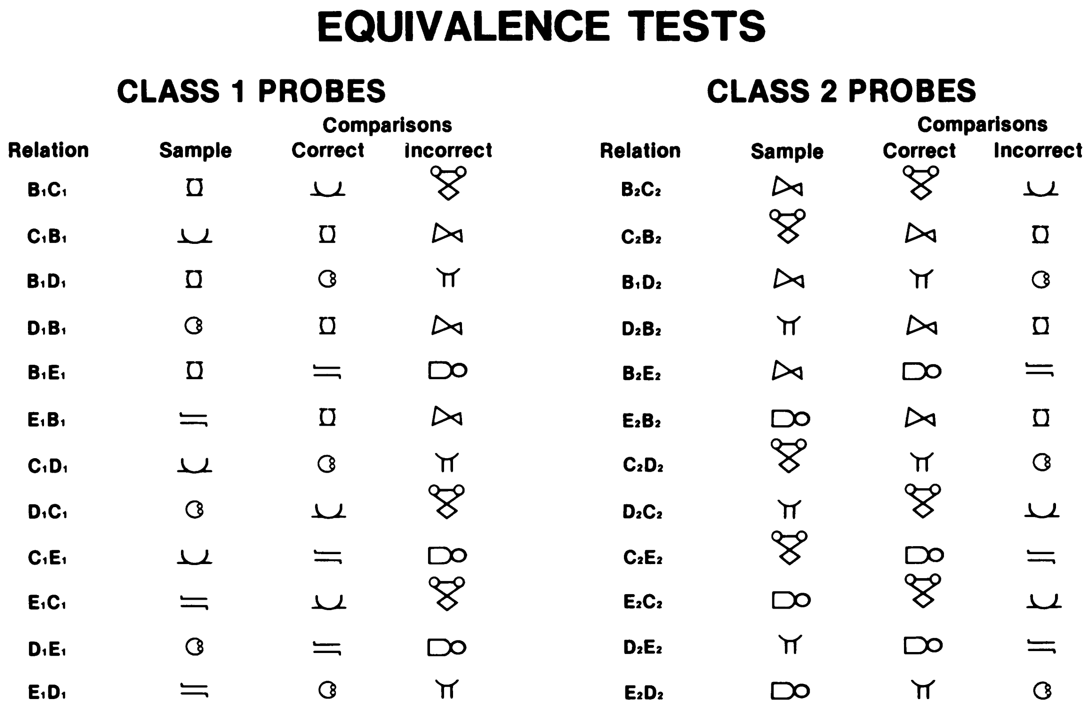

Fig. 8. The complete set of probes for equivalence for each of the two possible classes of visual stimuli following Phase 1 training with either the single-sample multiple-comparison or the multiple-sample single-comparison procedure.

was introduced and, similarly, training on EA and AE was conducted last. Next, training on the combined set of BA, CA, DA, EA or AB, AC, AD, AE discriminations was conducted. Throughout training, each discrimination always involved the correct comparison stimulus and the paired distractor from the other set of stimuli. Even on probe trials, comparison stimuli appeared as fixed pairs.

Sessions consisting of an equal number of trials of each discrimination were conducted until performance met criterion on the total set or until the performance showed obvious error patterns with one or more discrimination pairs within the set. If obvious error patterns were observed, additional training was conducted with only the discriminations on which errors were made, until performance again met criterion. Then, training on the total set of four discriminations was reintroduced. Alternating training between a particular discrimination(s) and the total set was continued until the subject's performance met criterion for the total, combined set. Next, training was continued, but with nonreinforcement (no sound or washer delivered) in 8 of the 32 trials. Training with intermittent reinforcement (variable ratio) continued until the subject's performance again met criterion.

Probes for equivalence class and symmetry. Following training, a series of sessions was conducted that probed for equivalence and symmetry relations. The probe trials for equivalence are shown in Figure 8. Each probe session for equivalence contained a sample of four probes from Class 1 and four from Class 2. The probe trials for symmetry are shown in Figure 9. Comparing Figure 7 with Figure 9 shows that stimulus relations for the symmetry trials for the multiple-sample single-comparison subjects are identical to the stimulus relations for the training trials for the single-sample multiple-comparison subjects, and vice versa.

Each probe session for symmetry (S) or equivalence (E) consisted of the 8 probe trials intermixed among 24 baseline training trials in which reinforcement occurred. The order and placement of probe trials within each session was unsystematic, except that probe trials were always separated by one or more training trials. For all subjects, probe sessions were alternated with baseline training sessions (T) in the following sequence: E, T, E, T, S, T,

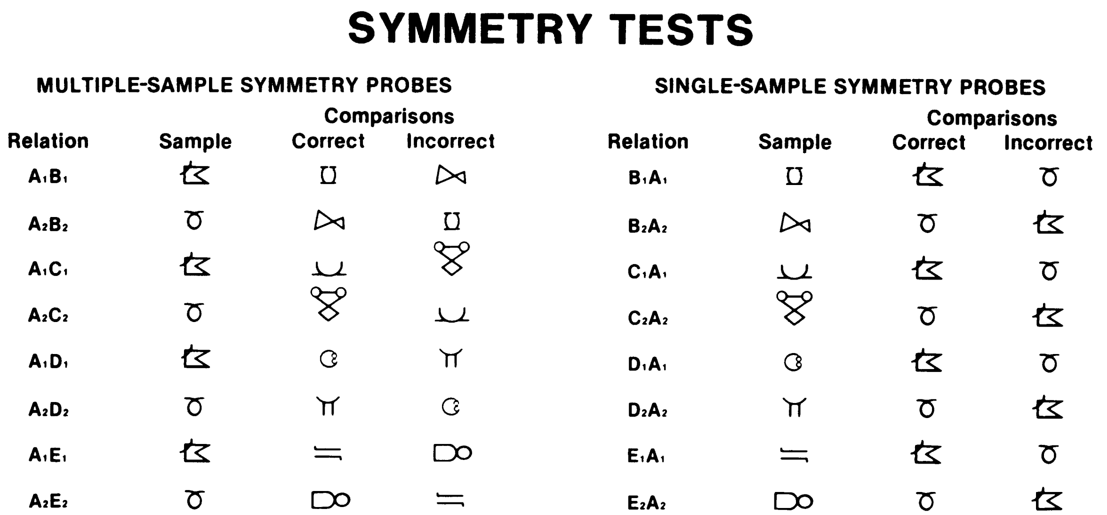

Fig. 9. The complete set of probes for symmetry of the trained visual-visual relations in Phase 1 with the multiple-sample single-comparison and with the single-sample multiple-comparison procedures.

T, T, T, E, S, E, S, E, S, E. If any subject performed on a baseline training session below the criterion of 90% correct, the sequence was interrupted and training continued until criterion performance was reestablished.

Phase 1 Results

All 6 subjects' performances met criterion on the identity-match pretraining within five sessions. Table 2 shows that between 21 and 69 sessions were required for the subjects' performances to meet the criterion for exposure to probe sessions. Across subjects, the total number of training sessions required was highly variable. The probe sequence did not disrupt baseline performances. During the probe sequence, only 1 subject's baseline performance fell below 90% correct, and that occurred in only one baseline training session.

Table 3 shows that 2 subjects' performances (BD and TK) in the multiple-sample single-comparison group showed equivalence rela-

Table 2

Number of sessions for each subject to reach criterion on each stage of training during Phase 1 and Phase 2.

		Multiple-sample single comparison				Single-sample multiple comparison			
			Subjects				Subjects		
Phase	Stage	Discriminations trained	BD	TK	LE	Discriminations trained	RF	BB	TB
1	1	BA	17	6	3	AB	6	3	6
	2	CA	2	2	3	AC	4	4	2
	3	DA	4	6	5	AD	5	4	2
	4	EA	3	2	3	AE	2	3	4
	5	[a] BA, CA, DA, EA (CRF)	4	9	24	[a] AB, AC, AD, AE (CRF)	2	16	27
	6	[a] BA, CA, DA, EA (VR)	39	2	2	[a] AB, AC, AD, AE (VR)	2	2	2
	Total		69	27	40		21	32	43
2	1	EA (Review)	1	1	1	AE (Review)	1	1	1
	2	EF	3	5	5	FE	5	6	3
	3	EG	2	3	4	GE	2	2	2
	4	EH	2	2	2	HE	2	2	2
	5	[a] EA, EF, EG, EH (CRF)	2	2	36	[a] AE, FE, GE, HE, (CRF)	2	8	8
	6	[a] EA, EF, EG, EH (VR)	3	2	2	[a] AE, FE, GE, HE, (VR)	2	2	2
	Total		13	15	50		14	21	18

[a] Refresher training on individual pairs to facilitate mixed-pair performance is included in these totals.

Table 3

Performance on probes for equivalence (E) and symmetry (S) expressed as number correct per session (of eight probe trials).

Phase	Training procedure	Subject	E	E	S	E	S	E	S	E	S	E	S	E
1	Multiple-sample	BD	6	8	8	4	8	6	7	8	8	7	8	7
	Single comparison	TK	6	1	3	2	8	8	8	7	7	8	7	7
		LE	5	5	7	6	8	7	8	7	8	7	8	6
1	Single-sample	RF	7	5	5	8	8	8	8	8	8	8	8	8
	Multiple comparison	BB	7	3	5	5	7	4	6	3	7	4	6	4
		TB	3	4	5	4	5	5	7	5	5	4	4	4
2	Mutiple-sample	RF	8	6	8	7	8	7	8	8				
	Single comparison	BB	3	7	6	4	7	4	5	6	6	4	5	6
		TB	4	4	4	4	5	4	4	4	4	3	4	4
2	Single-sample	BD	8	8	$\boxed{7}$	$\boxed{8}$	8	7	8	7				
	Multiple comparison	TK	7	8	8	8	8	8	8	8				
		LE	4	5	4	8	6	7	8	6				

Number correct (of six probe trials) on probes for inclusion of Phase 2 stimuli in Phase 1 equivalence classes.

Relation trained:		EF	EG	EH
Relations tested:		BF, CF, DF	BG, CG, DG	BH, CH, DH
Number correct for	BD:	6	5	5
each subject	TK:	3	2	3
	LE:	2	4	4
		FE	GE	HE
		FB, FC, FD	GB, GC, GD	HG, HC, HD
	RF:	6	5	6
	TB:	4	3	2
	BB:	4	4	3

Number correct on combined Phase 1/Phase 2 probes (of 16 probe trials) and two-node equivalence probes (of 18 probe trials) across the final sequence of sessions in Phase 3.

Subjects	Final combined Phase 1 and 2 (one-node) probes	Final two-node equivalence probe sessions				
TB	15	18	15	16	17	
RF	15	18	17	17		
TK	15	14	16	16	18	
BD	15	17	18	17	18	
LE	16	14	15	16	15	13

Number correct on auditory probes after auditory pair training and after time lapse.

Subjects	After first pair trained (14 probe trials)						After second pair trained (12 probe trials)						After time lapse (12 probe trials for BD, 14 probe trials for TD, RF, and TB)					
BD	7	8	6	6	5	6	12	11	12	12	10	12	11	10	11	12	11	11
TD	14	12	14	14									14	12	12	12		
RF	14	14	14	14									13	14				
TB	13	13	14	14	14								7	9	14	12	11	14

□ = These two probe sessions erroneously conducted in opposite order from that indicated in the table.

tions on all eight equivalence probe trials in one or more equivalence probe sessions. The other subject's performance (LE) showed equivalence relations on seven of eight trials in three sessions. All of the subjects' performances showed equivalence relations on at least six trials in each of the last four equivalence probe sessions.

One subject's performance (RF) in the single-sample multiple-comparison group showed equivalence relations on every probe trial in each of his last five equivalence probe sessions. Subject BB's performance showed equivalence relations on seven probe trials during the first probe session and then showed approximately chance levels thereafter. Subject TB's performance closely approximated the chance level in every equivalence probe session.

The multiple-sample single-comparison subjects' performances showed symmetry relations on seven or more of the eight probe trials during each of the last four symmetry probe sessions. In contrast, only the performance of Subject RF in the single-sample multiple-comparison group demonstrated symmetry relations consistently, but Subject BB's performance showed symmetry relations on six and seven of eight probe trials in each of the last four symmetry probe sessions. Subject TB's performance in symmetry probe sessions showed near-chance responding in all but one probe session.

Analyses of performances on equivalence and symmetry probes were made to determine if key positions, particular stimuli, or particular stimulus pairs were related to or controlled performance. In Subject TB's performance, seven responses occurred on the right key during the first equivalence probe session, then responses were distributed equally across the next equivalence and symmetry probe sessions, and finally nearly all responses occurred on the right key in the remaining five probe sessions. Eight left-key responses occurred in the third equivalence probe session for Subject BD, and six of eight responses occurred on the left key in his fourth equivalence probe session. Responses were distributed equally to both keys on the remainder of the symmetry and equivalence probes. No other subject exhibited position bias.

Subject BB's performance was the only performance that showed selection of a particular member of a stimulus comparison pair regardless of the sample. On 13 of 14 trials in which C_1 and C_2 were presented as comparisons, C_2 was selected (see Figure 3). On 10 of 13 trials in which B_1 and B_2 occurred as comparisons, B_2 was selected. When D_1 and D_2 were presented as comparisons, D_1 was selected on 10 of 14 occurrences. Selections of E_1 and E_2 were distributed approximately equally. No other subject's performance showed any consistent responses to particular comparison stimuli.

Subject LE's performance on equivalence probes did not show equivalence relations on 13 of the 56 equivalence probe trials. Of these, 10 trials involved stimuli C_1 and C_2. When symmetry was not shown, the trials also involved C_1 and C_2. No other subject's performance showed systematic responding to a particular stimulus pair while maintaining correct performance on items not involving that pair.

Figure 10 shows the median latency for each subject's responses during the final (criterion) session of pretraining and each stage of training. For all of the subjects in the multiple-sample single-comparison procedure response latencies for selections of the sample stimuli in the last session of each training stage were roughly equivalent to the latencies in the last sessions of every other training stage, regardless of the number of sessions required for performance to meet criterion in each stage. In contrast, median latencies for all subjects in the single-sample multiple-comparison procedure were 1 to 2 s longer in most stages subsequent to pretraining. Although latencies on specific training pairs varied cross and within subjects, latencies in the mixed pairs and variable-ratio training stages were consistently longer for these subjects than for those in the multiple-sample single-comparison procedures. Latencies on training trials were not disrupted for any subjects upon introduction of probe trials, although latencies on the first probe trial tended to be several seconds longer than were training trial latencies. Median latencies on probe trials thereafter were slightly longer than on training trials. The range of latencies on both probe and training trials was generally consistent and deviated little from the median.

Phase 2 Procedures

Phase 2 was designed to study the performances on training procedures opposite to those used in original training. That is, if a subject were initially trained with the multiple-sample single-comparison procedure in Phase 1, he or she was trained with the single-sample multiple-comparison procedure in Phase 2, and vice versa.

The same experimental apparatus and procedures were used to train the subjects in Phase 2 as were used in Phase 1. However, the Phase 2 stimulus pairs F, G, and H (as shown in Figure 4) replaced stimulus pairs B, C, and D.

As stated earlier, Phases 1 and 2 permitted

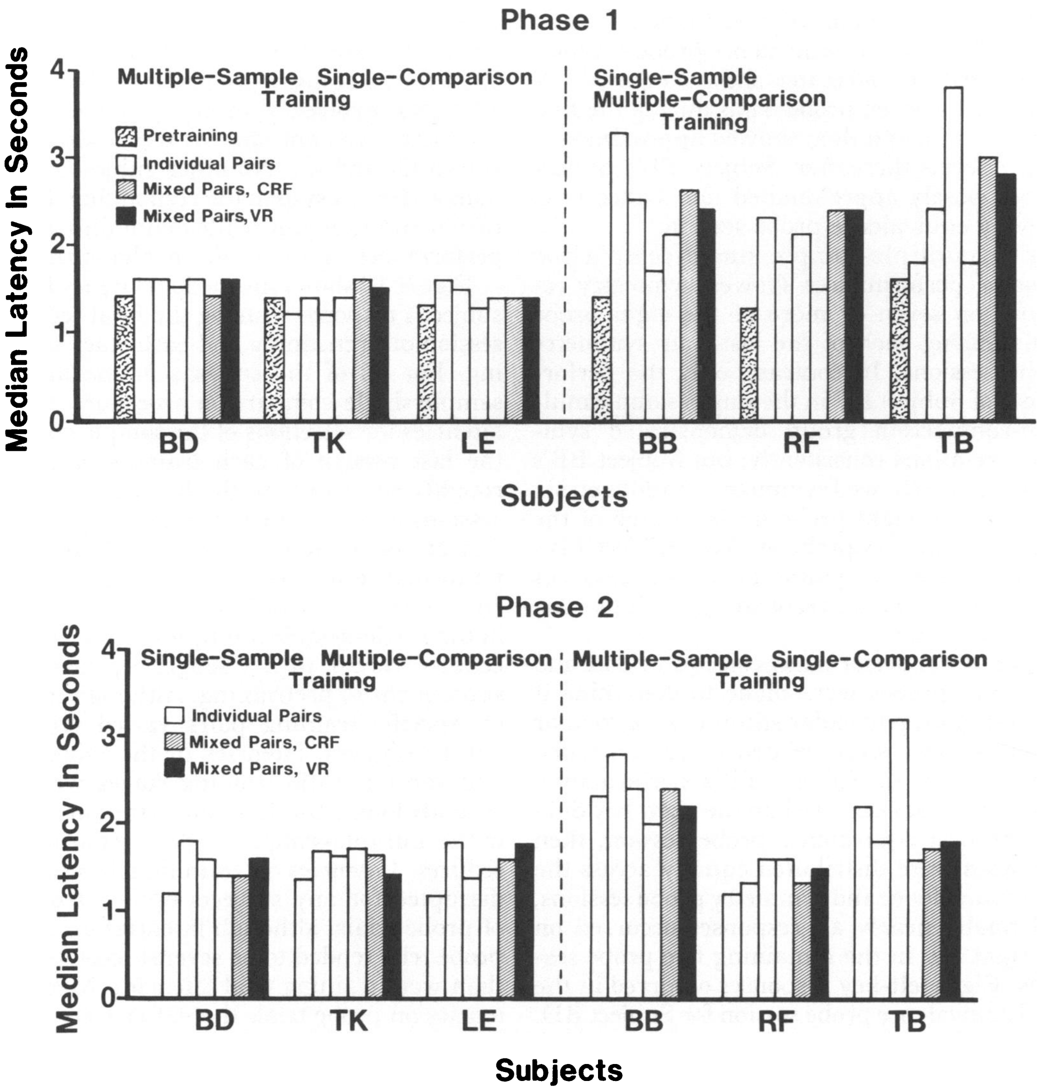

Fig. 10. The median latencies for each subject during Phase 1, for the final (criterion) session in pretraining and in each stage of training, and the median latencies for each subject during Phase 2 for the final (criterion) session in refresher training on the AE or EA discrimination, and in each stage of training with the Phase 2 stimuli.

evaluation of whether multiple-sample single-comparison and single-sample multiple-comparison training both result in equivalence class development when taught in sequence.

Single-pair training and probing. Prior to beginning training on the EF conditional discrimination, subjects in the multiple-sample single-comparison procedure during Phase 1 were given 32 review trials on the EA discrimination. Subsequently, the subjects received training on the EF discrimination (as shown in Figure 4). When the subject's performance met criterion on the EF discrimination, a probe session was conducted in which six BF, CF, and DF probe trials without reinforcement were intermixed with 26 EF training trials. Reinforcement was also omitted in two of the 26 training trials, to maintain a ratio of 8 trials without to 24 trials with reinforcement. After this probe session, the subject received EG training until performance met criterion; then a probe session was conducted in which six BG, CG, and DG probe trials were intermixed with 26 EG training trials (two without reinforcement). Then EH training sessions were introduced and continued until performance met criterion. Next, a probe session was conducted in which six BH, CH, and DH trials without reinforcement were intermixed with 26 EH training trials (two without reinforcement).

Subjects who received single-sample multiple-comparison training during Phase 1 were given 32 review trials on the AE discrimination prior to training on the FE discrimination (see Figure 4). After the subject's performance met criterion on the FE discrimination, one session was conducted in which six FB, FC, and FD probes were intermixed with 26 FE training trials. As was the case for all probe sessions described above, reinforcement was omitted in 2 of the 26 training trials. After this probe session, GE discriminations were trained to criterion. Then GB, GC, and GD probes were introduced for one session. Finally, HE training was completed, and HB, HC, and HD probes were introduced for one session.

These single-session probes were designed to reveal whether the previously established classes (for Subjects BD, TK, LE, and RF) were being expanded as new relations were trained with members (E_1 and E_2) of the Phase 1 classes. These probes were also two-node probes (where A and E are the training nodes). Correct performance on these probes would indicate that the possible one-node relations from each step of Phase 2 training had developed, but developed without testing for them directly. Performance on these probes, however, did not affect the established training sequence.

Combined discrimination training and probing. Following training on the individual discriminations (EF, EG, EH or FE, GE, HE), subjects were trained on all of the discriminations (including AE or EA) intermixed in sessions of 32 trials. These sessions were continued until the subject's performance met criterion.

When a subject's performance met criterion on the combined set, training was continued, but with reinforcement omitted in 8 of the 32 trials, until performance again met criterion. Then, two probe sessions for equivalence were introduced. The equivalence probe sessions included a sample of eight probe trials drawn from the AF, FA, GA, AG, HA, AH, FG, GF, FH, HF, GH, HG discriminations. No figure is provided to depict these trials, but it would be identical to Figure 8 with the F, G, and H symbols substituted for the B, C, and D symbols. These probes were introduced into a baseline of training trials used in the combined discrimination training. Thus, although probe items were sampled from the total set for both of the groups, the baseline trials into which they were intermixed differed according to the combined training given the groups. After probe sessions for equivalence class were conducted, a probe session to test for symmetry followed; then, probe sessions for equivalence and symmetry were alternated until at least five equivalence probe sessions and three symmetry probe sessions had been completed.

Phase 2 Results

Table 2 shows the number of sessions required for the performance of each subject to meet criterion in each stage of training. Table 3 shows the subjects' performances on the probes after training with individual pairs in Phase 2. After EF training, only Subject BD's performance showed that samples B_1, C_1, and D_1 controlled selections of F_1 whereas samples B_2, C_2, and D_2 controlled selections of F_2. That is, Subject BD showed generalization from EF training to BF, CF, and DF probes. The performance of the other 2 subjects (TK and LE) trained in Phase 2 with single-sample multiple-comparison procedures did not show this control by samples not directly trained with the F comparisons. Subject BD's performance on subsequent probes with the B, C, and D samples and the G and H comparisons also showed the generalization of sample control, following the EG and EH training, respectively. The performance of the other 2 subjects (TK and LE) remained at approximately chance level on the probes involving the G and H stimuli.

After FE training, only Subject RF's performance showed that the F_1 stimulus as sample controlled selections of B_1, C_1, and D_1 as comparison stimuli, whereas the F_2 sample controlled selections of B_2, C_2, and D_2 as comparison stimuli. That is, after Subject RF had received FE training he generalized to FB, FC, and FD discriminations. After training on the GE relations and the HE relations, this subject's performance showed control by the G and H samples, respectively, over the B, C, and D comparisons as well. The performances of the other 2 subjects (TB and BB) were at approximately chance level across all three probes, but these subjects' performances also had not shown equivalence relations after Phase 1 training. Analysis of responding on probe trials indicated that Subject TB's performance was again comprised predominantly of responses on the right-hand key (17 of 18 probe trials). Subject BB's performance also again showed strong control by stimulus B_1 on six of six trials, stimulus C_2 on five of six trials, and stimulus D_2 on five of six trials. None of the other subjects showed consistent control by position or specific stimuli during these probes.

Table 3 shows results from probes for the development of equivalence classes during Phase 2. Performances of subjects BD and TK, who received the single-sample multiple-comparison training during Phase 2, showed equivalence and symmetry relations throughout the equivalence and symmetry probes of Phase 2. Subject LE's performance was at approximately chance level on both the symmetry and equivalence probes until the third probe session, when her performance showed equivalence relations on eight of eight equivalence probe trials. Thereafter, her performance showed equivalence relations on 75% or better in all symmetry and equivalence probes.

Subject RF's performance showed symmetry relations on all symmetry probes and equivalence on at least 75% of all equivalence probes. The performances of Subjects BB and TB were not consistently above chance on equivalence probes. However, Subject BB's performance was more variable than that of TB, showing symmetry and equivalence relations somewhat above chance in some sessions.

Analysis of performance on symmetry and class probes in these probe sessions again indicated that Subject TB made a high proportion of his responses on the right key (86 of 96 responses). Subject BB's performance also showed strong stimulus control by specific stimuli—selecting F_1 on 10 of 10 trials when it occurred as a comparison stimulus and selecting H_2 on 8 of 10 trials when it occurred as a comparison stimulus.

Figure 10 shows median latencies across the stages of training. All subjects in the single-sample multiple-comparison procedure produced latencies nearly identical in overall pattern within and across the training stages to those produced in Phase 1 (multiple-sample single-comparison). Of the 3 subjects in the multiple-sample single-comparison procedure in Phase 2, only RF responded with latencies indicative of the pattern shown by the original multiple-sample single-comparison group in Phase 1. Subjects BB and TB again responded with longer and more variable patterns, but with TB's latencies in the mixed-pair and variable-ratio stages more similar to RF than to BB. Again, the introduction of probe stimuli did not affect latencies on training trials, but did temporarily produce longer latencies on the first few probes in the initial probe sessions of either type. These longer latencies on probes were generally short lived.

Phase 3 Procedures

To be exposed to the probes in Phase 3, which were designed to evaluate two-node generalization across Phase 1 and Phase 2 stimuli, a subject's performance was required to indicate equivalence consistently on both Phase 1 and Phase 2 equivalence probes (near eight of eight per session).

For Subjects BD, RF, and TK, this demonstration was accomplished with the sequence of conditions shown in Table 4. The sequence began with reintroducing Phase 1 training sessions until the subject's performance met the original criterion, then reintroducing Phase 1 equivalence probe sessions. When the subject's performance showed eight of eight correct on Phase 1 equivalence probes, Phase 2 training sessions were introduced until the subject's performance showed eight of eight correct on Phase 2 equivalence probes. These Phase 1 training and equivalence probe sessions and Phase 2 training and equivalence probe sessions were alternated until the subject's performance showed eight of eight correct on two consecutive Phase 1 and 2 equivalence probes sessions. Then, sessions were conducted with a combined set of 16 Phase 1 training trials and 16 Phase 2 training trials.

Once performance met criterion on the combined Phase 1 and Phase 2 sessions under reinforcement conditions, combined Phase 1 and Phase 2 sessions were presented in which no token or feedback was given in any trial. The subject was informed at the beginning of the session that the chime and token would not be presented when he was correct and that the buzzer would not buzz when he was incorrect. The subject was paid one penny for each correct response on a training trial and one penny for each probe trial, regardless of his performance, at the end of each 32-trial session. Subjects were not informed regarding the specific way in which the pennies were earned. Once the subject's performance was between 90% and 100% correct for these combined Phase 1 and Phase 2 training sessions, without differential reinforcement, one or more sessions were conducted in which eight Phase 1 equivalence probes and eight Phase 2 equivalence probes were intermixed with eight Phase 1 training trials and eight Phase 2 training trials. No feedback or reinforcer was delivered in either the training or the probe trials. Once the subject's performance was at 15 or 16 correct on the combined Phase 1 and Phase 2 equivalence probes, sessions were conducted in which 18 two-node equivalence class probes (BF, FB, BG, GB, BH, HB, CF, FC, CG, GC, CH, HC, DF, FD, DG, GD, DH, HD—as shown in Figure 5) were intermixed with seven Phase 1 and seven Phase 2 training trials. During these sessions no feedback was given in training or probe trials.

Because the performances of Subjects TB, LE, and BB were not near eight of eight correct on equivalence probe trials during Phase 2, additional procedures were introduced in an attempt to improve performance. Initially for Subject TB these procedures consisted of presenting training pairs (AE and FE) and reinforcing selections as before, but with interspersed equivalence probes (FA and AF) without reinforcement. This procedure did not result in generalization to FA and AF probes. However, when all trials (EF, FE, AF, FA) were presented without reinforcement or feedback, generalization did occur to the AF and FA pairs. Subsequent sessions replicated the differences in generalization between the baselines, with and without reinforcement. Eventually, when presented Phase 1 and Phase 2 equivalence probes in the appropriate baselines without reinforcement, Subject TB's performance began to show generalization in both Phase 1 equivalence probe sessions and Phase 2 equivalence probe sessions. From that point on, TB was presented with both combined Phase 1 and Phase 2 equivalence probe sessions and two-node equivalence probe sessions exactly as were BD, TK, and RF (probes among training trial, both without reinforcement).

The procedures used with subject LE consisted primarily of presenting a large number of additional Phase 1 and Phase 2 training and equivalence probe sessions under baseline conditions with and without reinforcement. Subject LE's performance never consistently showed eight of eight correct on Phase 1 or Phase 2 equivalence probes, but consistently showed six of eight or seven of eight correct.

Finally, even though Subject LE's performance never reached eight of eight correct on the Phase 1 or Phase 2 equivalence probes individually, combined Phase 1 and Phase 2 training and probe sessions were introduced to LE just as they had been introduced for Subject BD, TK, RF, and TB. Strangely, LE's performance was 16 of 16 correct on the combined Phase 1 and Phase 2 probes, and she was therefore introduced to the two-node equivalence probes.

The experiment was terminated for subject BB because he stated he no longer wanted to participate in the study.

Table 4

Sequence of training and probe conditions in Phase 3 for Subjects BD, RF, and TK.

Phase 1 training with variable ratio reinforcement.
Phase 1 probes for equivalence classes.
Phase 2 training with variable ratio reinforcement.
Phase 2 probes for equivalence classes.
Alternating sessions of Phase 1 probes for equivalence and Phase 2 probes for equivalence.
Combined Phase 1 and Phase 2 training with continuous reinforcement.
Combined Phase 1 and Phase 2 training with no programmed consequences.
Combined Phase 1 and Phase 2 probes for equivalence intermixed with Phase 1 and 2 training trials with no programmed consequences.
Phase 3 two-node equivalence class probes intermixed with Phase 1 and 2 training trials with no programmed consequences.

Phase 3 Results

Results for Phase 3 are shown in Table 3. All subjects' performances were 15 or 16 of 16 correct on the final exposure to the combined Phase 1 and Phase 2 probes. The performances of Subjects BD, RF, and TB were 18 of 18 correct at least once within the last three or four two-node equivalence probe sessions. Two-node equivalence probe performances for BD and RF were at or above 17 of 18 correct for all two-node equivalence probe sessions.

These performances for BD and RF were anticipated based on their performance on the two-node probes during Phase 2 training. Subject TB's performance on the second and third of these probe sessions dropped below 17 of 18 correct, but then was at 17 of 18 correct during the last probe session. Subject TK's performance during these two-node equivalence probe sessions ranged between 14 of 18 correct and 17 of 18 correct until the last probe session when it was 18 of 18 correct.

Subject LE's performance on earlier two-node equivalence probe sessions in Phase 3 ranged from 11 of 18 to 15 of 18 correct (not shown in Table 3). Because her performance on the combined Phase 1 and Phase 2 baseline trials deteriorated during these four two-node equivalence probe sessions, two combined Phase 1 and Phase 2 training sessions with reinforcement and two combined Phase 1 and Phase 2 training sessions with no reinforcement or feedback were conducted. Then four more two-node equivalence probe sessions were conducted. Her performance on probe trials ranged from 13 of 18 to 15 of 18 correct (not shown in Table 3). Two sessions of Phase 1 and Phase 2 training trials with reinforcement followed by two sessions of Phase 1 and Phase 2 training trials without reinforcement were conducted again. At this point, a session was conducted in which combined Phase 1 and Phase 2 equivalence probes were presented. Her performances on this combined Phase 1 and Phase 2 probe session was 16 of 16 correct. From this point on, two-node equivalence probe sessions and combined Phase 1 and Phase 2 equivalence probe sessions were presented in an unsystematic order for 26 more sessions. Her performance on the combined Phase 1 and Phase 2 probe sessions ranged between 14 of 16 and 16 of 16 correct. Her performance on the final two-node equivalence probes following a 16 of 16 combined Phase 1 and Phase 2 probe session was consistently lower than her combined Phase 1 and Phase 2 one-node equivalence probes, remaining between 13 of 18 and 15 of 18 correct (shown in Table 3).

Phase 4 Procedures

Phase 4 included the 4 subjects (BD, TK, RF, TB) whose performances demonstrated equivalence class development of both the five-member Phase 1 and Phase 2 classes and the eight-member class of all sample and comparison stimuli. This phase was designed to assess whether a subject who was taught to select one stimulus from the class in response to an auditory label would select the remaining seven stimuli in that class in response to that same label and whether there were differences in accuracy on one-node versus two-node relations.

The basic training in Phase 4 involved teaching the subject to select E_1 when the auditory stimulus "cadoo" was presented and to select E_2 when "sompta" was presented. Figure 5 shows all training that Subjects BD and TK and Subjects RF and TB, respectively, had received by the time Phase 4 training was completed.

Phase 4 training began with presentation of sessions in which the auditory stimulus "cadoo" and the auditory stimulus "sompta" were presented 16 times each in an unsystematic order. Each trial began with the presentation of a green square in the center display window. Nothing further occurred until the subject pressed the key below the center window; then the E_1 and E_2 stimuli appeared in the two outside windows and either the auditory stimulus "cadoo" or "sompta" was presented. If the subject responded within 3 s, no further auditory stimulus was presented; if he did not respond, the auditory stimulus was repeated at intervals of 3 s until a response occurred. If the response was correct, a token was delivered and a chime-like noise sounded; if it was incorrect, a buzzer-like noise sounded and no token was delivered. This training continued until criterion was reached or until it became apparent that the specified performances were not going to occur. The performance of 3 subjects (BD, TB, and RF) met criterion for the auditory stimulus training. Then two sessions were conducted of 32 trials each, in which E_1 and E_2 were presented as comparison stimuli with "cadoo" and "sompta" as samples, but no reinforcement or feedback was given during the session for either correct or incorrect responses. The subject was given up to 32 pennies after each session, calculated as for Phase 3. Then, probe sessions in which each of the seven remaining pairs of visual stimuli (A, B, C, D, F, G, and H) were presented—once each with "cadoo" and with "sompta" as the samples for a total of 14 probe trials interspersed with 18 auditory training trials. No reinforcement or feedback was given for any trial, but the subject was paid at the end of each completed session, calculated as above.

Because the procedure described above was not adequate to establish criterion performance on the training trials for TK, the following supplementary procedures were implemented: This subject's performance remained at near chance for nine training sessions, so additional training conditions were introduced. Initially, in this additional training, only the stimulus "cadoo" was presented during two 32-trial sessions and then only the stimulus "sompta" was presented during one 32-trial session. Although the subject performed at high levels when each auditory stimulus was presented alone, performance fell to near chance when the two auditory stimuli were intermixed during three subsequent sessions. At this point, a series of sessions was presented in which 16 trials with "cadoo" as sample was followed by a series of 16 trials with "sompta" as sample. Performance was above 90%, so a series of sessions was conducted with eight consecutive trials with "cadoo" alternating with eight trials with "sompta." Later series involving four consecutive trials of "cadoo" alternating with four trials of "sompta" were presented until the subject reached criteria. Then regular, randomly sequenced auditory training trials were introduced as described above. Because it had been many sessions and 20 days since TK had been exposed to the visual conditional discriminations that formed the basis for classes, he was given three sessions in which all of the relevant visual training trials were presented under extinction conditions. Performances on the last two sessions were 100% correct, so he was then given additional sessions of auditory stimulus training followed by the auditory probe sessions.

Phase 4 Results

As shown in Table 3, the performances of Subjects TK, RF, and TB were at the level of at least 12 of 14 or correct during all auditory probe sessions. Subject BD's performance did not show generalized control from the verbal stimuli (cadoo and sompta) to the remaining seven pairs of visual stimuli after training with only one pair. Additional sessions were conducted with Subject BD in which the prior combined Phase 1 and Phase 2 baseline training items were reviewed and two-node equivalence probes were presented. Then the verbal training trials and verbal probes were introduced again. Because generalization of verbal control did not occur, sessions were conducted in which verbal training items and the combined Phase 1 and Phase 2 baseline training items were intermixed. Generalization on the verbal probes still did not occur. At this point verbal training with a second pair (F) was conducted. After his performance met criterion with the second pair, sessions were conducted in which the verbal training trials with the two pairs were intermixed. After criterion with reinforcement was reached, two sessions were conducted in which the two verbal training pairs were presented without reinforcement or feedback. At this point, sessions that intermixed 12 probes for generalization to the six pairs of stimuli were intermixed with the 20 verbal training trials. After training on the second pair was completed, his performance on probe items ranged from 10 of 12 correct to 12 of 12 correct on probe trials.

None of the subjects' performances indicated differential generalization as a function of one-node versus two-node relations. At this point, the experiment was suspended for all 4 subjects for a period of 2 to 5 months. The experiment was then resumed for follow-up probes after this time lapse. As before, no reinforcement or feedback was given on any probe or training trials, but payment was made at the end of the session in the same manner as in Phases 3 and 4. As can be seen in Table 3, Subjects TK and RF's performances on probe trials were always at or above 12 of 14 correct. Subject BD's performance was 10 of 12 correct once and at least 11 of 12 correct otherwise. Initially, during follow-up, Subject TB performed at approximately chance level on both training trials and probe trials. However, despite no feedback on any training or probe trial, his performance improved so that by the third follow-up session his performance on both training and probe trials was 100% correct.

DISCUSSION

After completing the four phases of this experiment, 4 of the original 6 subjects' performances indicated the development of two eight-member visual classes. These eight-member classes included both one-node relations and two-node relations. Additionally, after one or two members of the eight-member class were brought under the control of an auditory stimulus, that control was generalized to the other six or seven members of the class without additional direct training. Follow-up tests indicated that these eight-member classes and the auditory control over them remained intact, without any opportunity for rehearsal, for periods of 2 to 5 months. These follow-up tests for auditory control were made without reinforcement of feedback on either the auditory training trials or the auditory probe trials.

Both the models described by Sidman and Tailby (1982) and by Fields et al. (1984) predict that equivalence classes should be as readily established if subjects are taught to select a single comparison stimulus in response to any of several samples as if subjects were taught to select any of several comparison stimuli in relation to a single sample. Phase 1 and Phase 2 were designed to investigate this issue. During Phase 1, equivalence relations were evident in the performances of all 3 subjects who were taught to select a single comparison stimulus in response to four sample stimuli and of 1 subject in the group with several comparison stimuli. Hence, the formation of equivalence classes is not totally dependent on the direction of training. However, the fact that only 1 of 3 subjects' performances in the single-sample multiple-comparison procedure showed the development of an equivalence class, coupled with previous results of Spradlin and Saunders (1986), suggests that the two types of procedure may not be equally effective in creating equivalence classes.

The latencies recorded during Phase 1 are not definitive. Although all 3 single-sample subjects exhibited higher overall latencies than did the multiple-sample subjects, Subject RF's performance, nevertheless, showed equivalence class relations. Shorter latencies on criterion performance may often accompany equivalence class development, but it is not a requirement.

The results of the second phase, in which training procedures were reversed, casts some doubt on the interpretation that multiple-sample single-comparison training is more likely to produce equivalence classes. In Phase 2, the same subjects whose performances showed equivalence class development in Phase 1 also had performances that showed class development, and for two of these, this occurred by the enlargement of the Phase 1 class. This finding suggests either that the differences obtained in Phase 1 probably related to individual differences among subjects or that if subjects learn to form classes in Phase 1, they are also more likely to form classes in Phase 2, either by class expansion or by new class formation. For the 2 subjects whose performances showed class expansion, their performance also showed that two-node relations (FB, GC, etc.) could occur without overt performance of the new one-node relations (AF, AG, etc.) intermediate to the two-node relations. Moreover, this indicates that these one-node equivalence relations were probably present prior to overt tests for them.

In Phase 3, the development of two-node relations occurred for 2 additional subjects (TK and TB); there seemed to be little difference between performances based on two-node relations and those based on one-node relations, although these performances were never compared within the same probe session. However, another subject (LE) was given a series of combined Phase 1 and Phase 2 equivalence probe session (one-node relations) intermixed with two-node equivalence sessions (two-node relations). This subject consistently scored lower in probe sessions for two-node relations than in probe sessions for one-node relations. Although Subject TK's performance showed no differences between one-node and two-node probes in Phase 3, it is interesting that correct performance on two-node equivalence probes (the single-session generalization probes during Phase 2 training) did not occur until after the performance on one-node probes in Phase 2 showed equivalence relations. (Subject TB's performance did not show equivalence relations in Phase 1 or 2.) These findings are consistent with earlier results by Sidman and his colleagues (Sidman et al., 1985, 1986) and the predictions based on the analyses proposed by Fields et al. (1984).

During Phase 4, auditory control over the classes of visual stimuli was demonstrated with all 4 subjects, and 3 of the 4 subjects maintained their accurate performances on verbal probes (performances that involved over 100 untrained relations) on the time-lapse test. The fourth subject (TB) initially performed on the final test at near chance on both baseline and probe trials and then gradually improved to 100% correct without programmed consequences.

The fact that Subject TB's performance improved across unreinforced trials raises some interesting questions concerning human remembering and the effects of erroneous performance on subsequent performance. Under conditions of no reinforcement or feedback, one might argue that advancement to the next trial would reinforce whatever response was made and thus should strengthen not only correct but incorrect responses as well. If this were the case, one would expect the error pattern to stabilize. Clearly this did not occur. Why not? It is possible that a network of trained and untrained relations under interdependent stimulus control had been established and that, during the posttest interval, stimulus control over some relations, including the initial baseline relations, was lost. The stimulus control of the intact relations may have served as the basis for ultimately recovering all of the relations involved in the stimulus class.

Fields et al. (1984) state, "As stimulus class size increases, the number of transitive relations increases at an ever expanding rate. The theoretical relationship could be considered metaphorically as the behavioral analogue of a cognitive breeder reactor because the number of transitive relations created far exceeds the number of relations that must be established by training. Does this theoretical relationship describe the empirical relationships induced in subjects exposed to appropriate training?" (p. 148).

The current study extends the stimulus class size demonstrated in retarded subjects to nine (eight visual stimuli and one auditory stimulus). By the completion of Phase 4, Subjects RF, TK, and BD had been trained on eight or nine specific conditional relations and performed correctly on 112 (110 for DB) derived relations. Clearly, with this number of stimuli, the theoretical relationship does describe the empirical relations induced by the training and probing. Moreover, the induced relations remained stable over time, even up to 5 months in the absence of practice.

One might assume that, as more and more stimuli are added in training, an upper limit would be reached in which either the total class would begin to break down or it would be very difficult to add new members to the class. With nine stimuli this does not seem to be the case. Conversely, it is possible that as more members are added to a class, the relations within the class become more stable and new members may be added more easily and be more likely to be maintained. If a stimulus class consists of only two members, A and B, there can be only linking relations involving A and B. If any variable reduces the strength of those AB relations, the class is essentially destroyed. In a class of four members, A, B, C, and D, there is a network of linking relationships AB, AC, AD, BC, BD, CD, and their reciprocals. The maintenance of the AB relationship is now multiply determined and less susceptible to disturbance.

If this is the way stimulus classes work, then new members to classes should be added more easily to a large class than to a small class. This possibility is consistent with the observation that repeated testing is necessary for the performances of some subjects to show equivalence relations. The formation of the class can commence with a few, well-established relations. Temporally contiguous exposure to these and some less well-established relations may add the latter to the array of established relations (in much the same way as occurred in the performance of Subject TB during the posttest). With sufficient exposure or opportunities to respond, the entire equivalence class is formed and exhibited. Then, by training a relation between some new stimulus and one of the equivalence class members, multiple new equivalence relations are established, and the class expands accordingly.

The finding that equivalence classes often are demonstrated only after several probes have been presented has led Sidman et al. (1985) to propose that the baseline training conditions in themselves are insufficient for the development of equivalence classes and that the classes come into existence only through their testing. That is, the tests provide the final conditions which are required to bring about the development of stimulus classes. This suggests that the classes could not have been formed prior to the initial test. There are two problems with this argument. First, there are some subjects who perform at the level of 100% correct on the their first test or, like BD and RF, perform correctly on two-node tests without ever performing some of the intermediate relations. In these cases, the baseline training conditions are sufficient for the formation of the classes. Second, there is the fact that the nature of the probe sessions themselves influence whether equivalence relations will be demonstrated. Sometimes when probes without reinforcement are intermixed with baseline trials in which reinforcement occurs, equivalences are not demonstrated; however, when these same probes are intermixed among baseline trials in which reinforcement is omitted, equivalences are demonstrated (Lazar et al., 1984; Sidman et al., 1985). This finding suggests that some very subtle factors may determine whether or not equivalences are demonstrated on probe trials.

In the present study, Subject TB failed to demonstrate either symmetry or equivalence class development during Phase 1 and Phase 2 probes. On probe trials he rather consistently selected the stimulus in the right-hand position. This performance was manifest consistently when probes were presented in a training baseline that included reinforcement. However, when probes were presented in a training baseline that did not include reinforcement, he began to select both symmetry and equivalence class probe items in much the same fashion as Subjects BD, TK, and RF. Subject TB's performance suggests that, even before he performed correctly on equivalence class probes, these derived relations already may have been developed, but they were masked by another form of stimulus control—that of position.

Regardless of whether one views the tests as the final necessary step required for the formation of equivalence classes or as demonstrating the classes already formed during baseline training, there can be no doubt that performance on probe items improves with repeated testing. The decrease in errors with repeated testing has been repeatedly observed previously (Sidman et al., 1974, 1985, 1986; Spradlin et al., 1973; Spradlin & Saunders, 1986) as well as here, both in the initial testing and on retesting several months later. The improvement or reestablishment of performance under conditions of no feedback is an important area for future research.

REFERENCES

Baer, D. M. (1982). Applied behavior analysis. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary behavior therapy* (pp. 276–309). New York: Guilford Press.

Dixon, M. H. (1978). Teaching conceptual classes with receptive label training. *Acta Symbolica,* **9,** 17–35.

Dixon, M. H., & Spradlin, J. E. (1976). Establishing stimulus equivalence among retarded adolescents. *Journal of Experimental Child Psychology,* **21,** 144–164.

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis of Behavior,* **42,** 143–157.

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

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

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

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

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Erlbaum.

Sidman, M., & Cresson, O., Jr. (1973). Reading and crossmodal 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., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior,* **43,** 21–42.

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.

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities,* **6,** 1–19.

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

Spradlin, J. E., & Saunders, R. R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classifications vs comparison classification. *Analysis and Intervention in Developmental Disabilities,* **6,** 41–58.

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.

Touchette, P. E. (1971). Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior,* **15,** 347–354.
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,* **40,** 69–78.

Received January 20, 1987
Final acceptance September 19, 1987