JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1997, **68,** 67–91 NUMBER 1 (JULY)

PREDICTING THE EXTENSION OF EQUIVALENCE CLASSES FROM PRIMARY GENERALIZATION GRADIENTS: THE MERGER OF EQUIVALENCE CLASSES AND PERCEPTUAL CLASSES

LANNY FIELDS, KENNETH F. REEVE, BARBARA J. ADAMS, JOHN L. BROWN, AND THOM VERHAVE

QUEENS COLLEGE AND THE GRADUATE SCHOOL AND UNIVERSITY CENTER/CUNY

In Experiment 1, 6 college students were given generalization tests using 25 line lengths as samples with a long line, a short line, and a "neither" option as comparisons. The neither option was to be used if a sample did not go with the other comparisons. Then, four-member equivalence classes were formed. Class 1 included three nonsense words and the short line. Class 2 included three other nonsense words and the long line. After repeating the generalization test for line length, additional tests were conducted using members of the equivalence classes (i.e., nonsense words and lines) as comparisons and intermediate-length lines as samples. All Class 2 comparisons were selected in the presence of the test lines that also evoked the selection of the long line in the generalization test that had been given before equivalence class formation. Class 1 yielded complementary findings. Thus, the preclass primary generalization gradient predicted which test lines acted as members of each equivalence class. Regardless of using comparisons that were nonsense words or lines, the post-class-formation gradients overlapped, showing the substitutability of class members. Experiment 2 assessed the discriminability of the intermediate-length test lines from the Class 1 (shortest) and Class 2 (longest) lines. The test lines that functioned as members of an equivalence class were discriminable from the line that was a member of the same class by training. Thus, these test lines also acted as members of a dimensionally defined class of "long" or "short" lines. Extension of an equivalence class, then, involved its merger with a dimensionally defined class, which converted a close-ended class to an open-ended class. These data suggest a means of predicting class membership in naturally occurring categories.

Key words: equivalence classes, generalization of emergent relations, perceptual classes, class merger, discriminability, key press, human adult

An equivalence class is established by first training $n - 1$ conditional relations between n stimuli that do not obviously resemble each other. The remaining $n2 - (n - 1)$ untrained pairwise relations are then presented as probes to assess class formation. Class-consistent responding in the presence of the probe configurations demonstrates the emergence of the equivalence class, that is, the properties of reflexivity, symmetry, transitivity, and the combined properties of symmetry and transitivity among the stimuli in the trained relations (Bush, Sidman, & de Rose, 1989; Fields, Adams, Newman, & Verhave, 1992; Fields & Verhave, 1987; Fields, Verhave, & Fath, 1984; Sidman, 1990, 1994; Sidman & Tailby, 1982).

Although an equivalence class contains a finite number of members, stimuli that resemble one of those members can also come to function as members of the class (Barnes & Keenan, 1993; Bush, 1993; Cowley, Green, & Braunling-McMorrow, 1992; DeGrandpre, Bickel, & Higgins, 1992; Fields, Adams, Brown, & Verhave, 1993; Fields, Adams, Buffington, Yang, & Verhave, 1996; Fields, Reeve, Adams, & Verhave, 1991; Haring, Breen, & Laitinen, 1989; Stromer, Mackay, Howell, & McVay, 1996). When that occurs, the number of stimuli in the class is no longer bounded (Adams, Fields, & Verhave, 1993b).

Bush (1993), DeGrandpre et al. (1992), and Barnes and Keenan (1993) demonstrated the extension of an equivalence class to only one new stimulus that was a variant of a

This research was supported partly by ARI Contract MDA903-90-C-0132, NICHHD Grant RO1-HD21110, and PSC/CUNY Grant 661287 to Lanny Fields. Our thanks to Nursel Kahya for her constructive criticisms in the preparation of this manuscript, to Wei Yang for his assistance in the preparation of the figures, and to Dawn Buffington for her assistance in running subjects and analyzing data. Wei Yang and Dawn Buffington were supported in part by NYS-OMRDD fellowships.

Reprints can be obtained from Lanny Fields at Queens College/CUNY, 65-30 Kissena Boulevard, Flushing, New York, 11367.

class member. Haring et al. (1989), Cowley et al. (1992), and Stromer, Mackay, Howell, and McVay (1996) demonstrated extension of an equivalence class to a set of stimuli that were variants of a class member. In all of these studies, however, the variants were not ranked in terms of physical or perceived similarity to the class member.

In contrast, Fields et al. (1991) demonstrated the extension of equivalence classes to stimuli that varied quantitatively from a particular member of an equivalence class. They established two three-member equivalence classes that consisted of stimuli represented here by the letters A, B, and C, by training AB and BC relations. The A and B stimuli in both classes were nonsense words; the C stimuli were long lines in Class 2 and short lines in Class 1. The generalization of equivalence relations was tested by the presentation of intermediate line lengths as samples with the A1 and A2 nonsense words as the comparisons. The test lines most similar in length to the Class 2 lines always occasioned selection of the Class 2 nonsense word. As the length of the test lines decreased, the likelihood of selecting the Class 2 comparison word decreased systematically to zero. Complementary results were obtained for Class 1.

In the studies conducted by Barnes and Keenan (1993), Bush (1993), DeGrandpre et al. (1992), Fields et al. (1991), Cowley et al. (1992), Haring et al. (1989), and Stromer, Mackay, Howell, and McVay (1996), generalization tests were not conducted for all emergent relations. Therefore, none of these studies provide evidence that the test stimuli had become related to all members of the equivalence class used in training. This limitation was remedied by Fields, Adams, Brown, and Verhave (1993) and more recently by Fields et al. (1996), who measured, across trials, the selection of each member of an equivalence class in the presence of dimensional variants of a particular member of the class. A variant is a stimulus that differs from another stimulus along some quantitatively definable dimension (Fields et al., 1991). Two four-member equivalence classes were established using three nonsense words as the A, B, and C stimuli and a line as the D stimulus. The D stimuli in Classes 1 and 2 were a short and a long line, respectively. Classes were formed by training AB, BC, and CD (three word-word relations and one word-line relation). After class formation, subjects were presented with primary generalization tests (Balsam, 1988; D. Blough, 1983; Galizio & Baron, 1976; Guttman & Kalish, 1956; Hanson, 1959; Honig, 1969; Honig, Boneau, Burstein, & Pennypacker, 1963; Rilling, 1977; Thomas, 1993) in which many different intermediate-length lines (dimensional variants) were presented as samples with the D1 and D2 lines as the comparisons. Subjects were also presented with analogous generalization tests of symmetry in which the C1 and C2 words were the comparisons and generalization tests of equivalence in which the comparisons were either the B1 and B2 or A1 and A2 words. All tests were conducted on a concurrent basis.

For each test line presented, there was a high correlation between the selection of the Class 2 words and the Class 2 line in the generalization tests of emergent relations. Increasing the disparity between the length of the test line and the Class 2 line resulted in a systematic decline in the selection of the Class 2 comparisons. The same range of intermediate-length test lines occasioned the exclusive selection of the Class 2 line and the Class 2 words. These test lines, then, acted as members of Equivalence Class 2. Complementary results were obtained for Class 1. Although the pattern of results was consistent across subjects, the particular ranges of test lines that functioned as members of Classes 1 or 2 varied across subjects.

In both the Fields, Adams, Brown, and Verhave (1993) and Fields et al. (1996) studies, the overlapping generalization gradients for the different kinds of probes identified the intermediate-length test lines that functioned as new members of an equivalence class. Because the tests were conducted concurrently, however, the performance on one type of probe test could not be used to predict the future performances occasioned by the same line samples in the other types of probe tests. Indeed, the studies of generalization of equivalence classes conducted by Bush (1993), Barnes and Keenan (1993), Cowley et al. (1992), DeGrandpre et al. (1992), Haring et al. (1989), and Stromer, Mackay, Howell, and McVay (1996) also did not include procedures that allowed for the prediction of the variants that would function as members of an equivalence class. Such a prediction might be made by conducting a primary generalization test prior to equivalence class formation. For example, the procedure used by Fields, Adams, Brown, and Verhave (1993) could be replicated with the addition of a generalization test of line length prior to equivalence class formation. The predictive value of performances on such a test could be determined by comparing them with the performances occasioned by the generalization tests of emergent relations conducted after equivalence class formation. There are two plausible outcomes; the generalization gradients obtained prior to class formation could either (a) overestimate the degree of generalization or (b) overlap the generalization gradients obtained after equivalence class formation. In the former case, the pre-class-formation gradient would overestimate the range of variants that would come to function as members of the equivalence class. In the latter case, the pre-class-formation gradients would precisely predict the range of variants that would come to function as members of the equivalence class.

Overestimation. Three sources of evidence drawn from studies of primary generalization suggest that the generalization gradient obtained prior to equivalence class formation would overestimate the range of stimuli that would come to function as members of an equivalence class. First, several studies have shown that primary generalization gradients become sharper with repeated testing (P. Blough, 1971, 1972; Friedman & Guttman, 1965; Mishkin & Weiskrantz, 1959; Thomas & Barker, 1964). In the experiment reported in this paper, the addition of primary generalization tests prior to equivalence class formation would add repeated tests of primary generalization. Second, generalization gradients are sharper for subjects who learn a label for a positive stimulus (S+) prior to a generalization test than for subjects who do not learn S+ labels (Galizio & Baron, 1976). In the present experiment, after obtaining the primary generalization gradient, one line becomes a member of an equivalence class by linking it with a nonsense word. This process may be akin to learning a label for a stimulus that is an S+ in a generalization test (Dickins, Bentall, & Smith, 1993; Spradlin & Dixon, 1976). Third, the generalization gradients obtained following intradimensional training (Balsam, 1988; Hanson, 1959; Honig & Urcuioli, 1981; Rilling, 1977) or extradimensional training (Balsam, 1988; Honig, 1969; Honig & Urcuioli, 1981; Rilling, 1977; Thomas, Freeman, Svinicki, Burr, & Lyons, 1970) are sharper than those obtained following comparable control or pseudodiscrimination training. In the experiment reported in this paper, word-word and word-line conditional discriminations were trained as the prerequisites to equivalence classes. The word-word relations can be viewed as instances of extradimensional discrimination training relative to the measurement of generalization along the dimension of line length. When the word-line conditional discriminations are established, a word is presented as a sample and two different lines are presented as comparisons. The word-line trials, then, include intradimensional discrimination training between comparison stimuli. Assuming the generality of the effects of repeated testing, stimulus labeling, and interposed discrimination training across subject populations and training and testing format, the analysis just offered implies that generalization gradients obtained before the establishment of equivalence classes would likely overestimate the range of variants that would come to function as members of an equivalence class.

Overlap. There are grounds, however, for expecting overlap instead of overestimation. Two sources of evidence suggest that a generalization gradient obtained prior to equivalence class formation would precisely predict the range of stimuli that would come to function as members of an equivalence class. First, Fields, Adams, Brown, and Verhave (1993) found that generalization gradients of emergent relations after equivalence class formation remained stable across more than 2,000 test trials. Second, Fields et al. (1996) established equivalence classes by training AB, BC, and CD, where the A, B, and C stimuli were nonsense words and the D stimuli were long or short lines. Then, generalization tests of emergent relations identified the test lines that functioned as members of each equivalence class. Thereafter, a different response was trained to a stimulus in each class. A transfer test showed that each response transferred to the test lines that functioned as members of the corresponding class. The degree of response transfer to the test lines was predicted from the performances occasioned by the same test lines when they were previously presented during the generalization tests of the emergent relations. These results suggest that a primary generalization gradient obtained prior to equivalence class formation would predict rather directly the range of stimuli that would come to function as members of an equivalence class.

EXPERIMENT 1

The purpose of Experiment 1 was to determine whether generalization test performances obtained prior to equivalence class formation overestimate or precisely predict the range of variants that come to function as members of an equivalence class. After conducting a primary generalization test of line length, two equivalence classes were established, each of which contained three nonsense words and one line. One class contained the longest line used in the generalization test, and the other contained the shortest line. After equivalence class formation, additional primary generalization tests of line length were conducted, first alone and then intermixed with additional generalization tests of symmetry and equivalence.

Method

Subjects

Six undergraduate students at Queens College/CUNY were recruited from an introductory psychology class. None of the subjects were familiar with the research area. Subjects received partial course credit upon completion of the experiment. Credit did not depend on performance during the experiment. The experiment was completed in about five 1-hr experimental sessions scheduled over a 3-week period.

Apparatus and Stimuli

All phases of the experiment were conducted with an IBM®-compatible computer that displayed the stimuli on a monochrome monitor. Responses consisted of touching specific keys on the computer keyboard. The experiment was controlled by software that programmed all stimulus presentations and recorded all keyboard responses.

Each of the two classes was composed of three nonsense words and a line. In Class 1, the nonsense words were LEQ (A1), HUK (B1), and POV (C1); in Class 2, the nonsense words were MEV (A2), GUQ (B2), and ZOJ (C2). The D stimulus in each class was a line. Each line was composed of a contiguous horizontal string of ASCII character 176, where length was identified by the number of ASCII units in the string. Each character was 3 mm wide and 5 mm high on the computer screen. The Class 1 line (D1) was 1 unit long; the Class 2 line (D2) was 25 units long. The sample stimuli used in the generalization tests were 25 lines that varied in length from 1 to 25 units, in 1-unit increments. Eleven representative lines are illustrated in Figure 1.

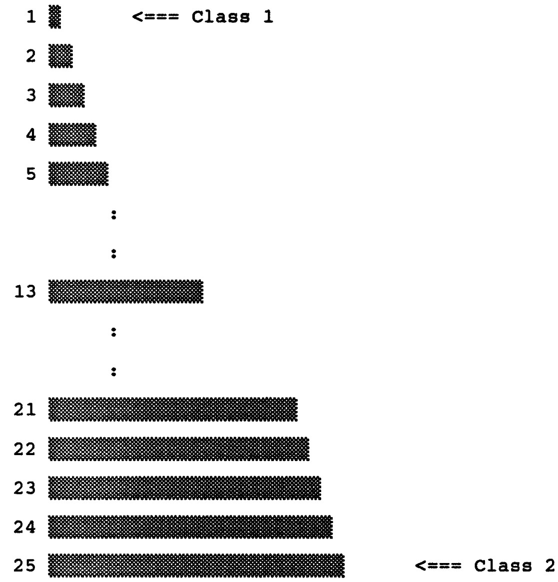


Fig. 1. The lines used as the D stimuli in Classes 1 and 2 and some representative test lines of intermediate length that were used as samples in the generalization tests. The numeral to the left of each line indicates its length measured by number of contiguous ASCII 176 characters used for its construction.

Procedure

Trial format, contingencies, and responses within a trial. All training trials and testing trials were conducted in a matching-to-sample format. Each trial began when "Press ENTER" appeared on the screen. Pressing the enter key removed the message and displayed a sample. Pressing the space bar in the presence of a sample added the comparison stimuli. On trials in which two comparisons were scheduled to appear, they formed an isosceles triangular array with the sample at the vertex and the comparisons at the corners of the base. When three comparisons were scheduled to appear, pressing the space bar also added the message, "If NEITHER, press 4," centered on the screen 1 cm beneath the other comparisons. This message was the neither comparison. Subjects pressed the 1 key to choose the comparison on the left, the 2 key to choose the comparison on the right, or the 4 key to choose the neither comparison. Each of these choices cleared all stimuli from the screen and produced a feedback message.

On all training trials, only one comparison was correct; this was called the positive comparison. The remaining comparisons were incorrect, each of which was called a negative comparison. On trials that received informative feedback, if the positive comparison was chosen, "RIGHT" appeared and remained there until the subject pressed the R key. If the negative comparison was chosen, the message "WRONG" appeared and remained there until the subject pressed the W key. When noninformative feedback was scheduled, the stimulus "—" was presented following the selection of a comparison to signal the end of a trial and remained on screen until the subject pressed the E key. The E key was used because it is located between the R and W keys on the standard keyboard. After the appropriate response (R, W, or E) was made, the screen was cleared, and the next trial began (Fields, Landon-Jimenez, Buffington, & Adams, 1995).

Trial block structure and feedback contingencies. Each phase of training and testing was conducted in blocks. Each block contained trials that were presented in a random order without replacement. At the start of training, a block of trials was presented repeatedly until all trials within the block occasioned 100% correct responding (the mastery criterion unless otherwise indicated). During these blocks, informative feedback was provided after each trial. Thereafter, the percentage of trials in a training block that occasioned informative feedback was reduced to 75%, then to 25%, and finally to 0% over successive blocks as long as performance within a block was maintained at 100% accuracy. If the mastery criterion was not reached within three blocks, the subject was returned to the previous feedback level. During all test blocks, however, each choice response was followed by noninformative feedback.

Phase 1: Keyboard familiarization. Subjects were trained to make the appropriate keyboard responses in the presence of each cue used within a trial. In each three-word trial, the semantic relation between the sample word and one of the comparisons was used to prompt the selection of the correct comparison. The stimuli used in each trial in Phase 1 are listed in Table 1. Each trial was presented once per block. Informative feedback was scheduled on all trials during this phase.

Correct responding to the stimuli in a trial was also facilitated by the presence of instructional prompts, which were deleted in a serial manner across trials (Buffington, Fields, & Adams, 1997; Fields, 1980; Fields, Adams, Verhave, & Newman, 1990) as illustrated in Figure 2. Phase 1 ended once the stimuli were presented without prompts and performance exceeded 85% accuracy (14 of 16 correct trials) during a single block. For the remainder of the experiment, if one of the trained keyboard responses was not made in the presence of the appropriate stimulus, the instruction that prompted the appropriate response reappeared on the screen for that trial as well as the next two trials. One example would be presenting the prompt "Press R to continue" if the subject incorrectly pressed the W key in the presence of the feedback word "RIGHT."

Phase 2: Identity conditional discrimination training. Subjects received identity conditional discrimination training with the 1-unit and 25-unit lines. On each trial either the 1-unit or the 25-unit line was presented as a sample. The comparisons consisted of the 1-unit and the 25-unit lines only. The stimuli used in each triad are listed symbolically in Table 1. When 100% feedback was scheduled, each block contained 16 trials; thus, each triad was presented eight times per block. When 75% to 0% feedback was scheduled, each block contained eight trials; thus, each triad was presented four times per block. Each of the two comparisons appeared equally often on the left and the right in a block. A correct response was the selection of the comparison that was the same as the sample. Feedback

Table 1

The stimuli used as samples (Sa), positive comparisons (Co+), and negative comparisons (Co−) in Phases 1, 2, and 4. NC indicates the neither comparison, which consisted of the written message, "If NEITHER, press 4." The stimuli in Phase 2 are represented nominally.

Sa	Co+	Co−	Co−
Phase 1: Keyboard familiarization			
ALCOHOL	DRUNK	MOUSE	
ANT	BEE	COW	
CANARY	SPARROW	STARS	
CAT	MOUSE	DRAGONS	
COMETS	STARS	FATHER	
DOG	WOLF	DARK	
DUNGEONS	CHAINS	PENCIL	
EGGS	BACON	SPARROW	
KINGS	QUEENS	CAMELS	
LIGHT	DARK	SOCK	
MOTHER	FATHER	BACON	
MUD	PIG	HAT	
PAPER	WRITE	OCEAN	
RED	COLOR	PEAR	
SOAP	WATER	THAT	
THIS	THAT	KINGS	
Phase 2: Identity conditional discrimination training			
1-unit	1-unit	25-unit	
25-unit	25-unit	1-unit	
Phase 4: Neither comparison training			
DOG	WOLF	DARK	NC
RED	COLOR	PEAR	NC
MUD	PIG	HAT	NC
MOTHER	FATHER	BACON	NC
ANT	BEE	COW	NC
DUNGEONS	CHAINS	PENCIL	NC
SHOE	SOCK	DARK	NC
COMETS	STARS	FATHER	NC
SOAP	NC	COMPUTER	TRASH
PAPER	NC	COFFEE	OCEAN
CANARY	NC	FIRE	STARS
KINGS	NC	TRUCK	ROCK

was scheduled as outlined in *Trial Block Structure and Feedback Contingencies.*

Phase 3: Two-choice primary generalization test. Once the identity conditional discriminations had been established, a primary generalization test for line length was conducted. In the block of trials used for this test, each of the 25 lines, which varied from 1 to 25 units in length, was presented as a sample on two trials. The 1-unit and 25-unit lines were used as the comparisons on all trials. For each sample, each comparison was presented once on the left and once on the right. Each block contained 50 trials; five blocks were presented for a total of 250 trials. Noninformative feedback was provided for all comparison selections.

Phase 4: Training the use of the neither comparison. When only two comparisons are provided, as in the Phase 3 generalization test, any sample must be assigned to one comparison or the other. If an intermediate-length line does not appear to be perceptually similar to either comparison, subjects cannot assign that stimulus to either of the available comparisons (Fields, Adams, Brown, & Verhave, 1993). To allow for such an option, Phase 4 introduced the neither comparison as a third response on all test trials. Training the appropriate use of the neither comparison began with the presentation of the following instructions on the computer monitor:

> You will learn how to use the number 4 key to select a third choice. Please select the com-

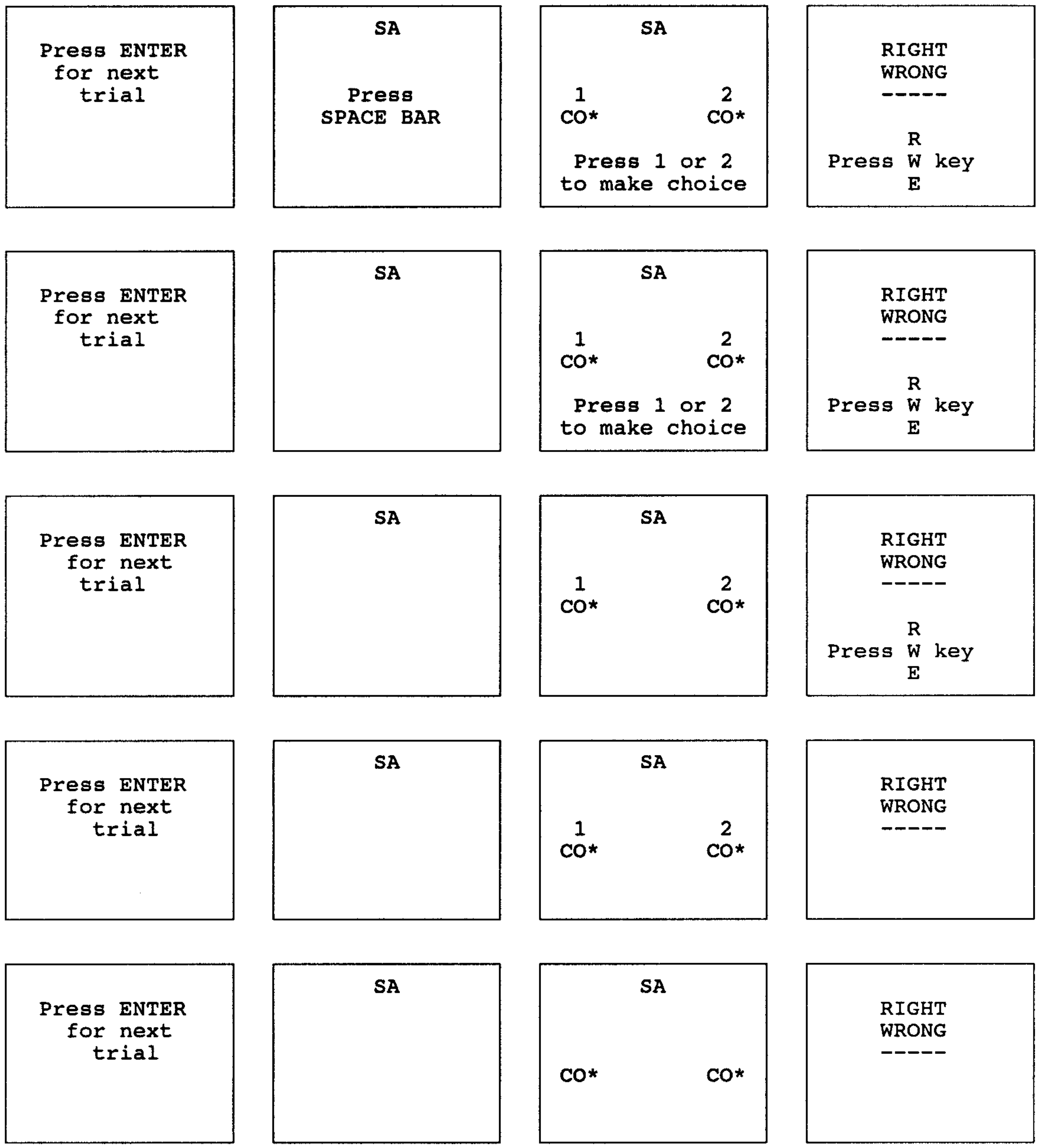


Fig. 2. Sequential changes in the stimuli and the prompts that are presented during a trial are illustrated across a row. Trial onset begins with the leftmost frame. Deletion of prompts in successive blocks of trials is illustrated in successive rows. Messages are as indicated in each frame. SA represents the location of the sample stimuli. CO* represents the location of the comparisons. After a comparison was selected, only one of the three feedback messages was presented on the screen, although all three possibilities are included in Figure 2. The response that terminated the feedback message corresponds to pressing R for right, W for wrong, and E for —(i.e., no informative feedback).

parison stimulus that is related to the sample. If they are not related to the sample, press the number 4 key. Thank you for your cooperation.

The trials used in Phase 4 are listed in Table 1. Each trial was presented once per block. On all trials, the sample and two of the comparisons were words. In addition, the neither comparison was available on all trials. On some trials one of the word comparisons (e.g., COMETS) was semantically related to the sample (e.g., STARS), and its selection occasioned informative feedback. On other trials, neither of the two comparison words was related to the sample, in which case the selection of the neither comparison occasioned informative feedback (McIlvane & Stoddard, 1981; Urcuioli & Nevin, 1975). Feedback was reduced as indicated in the *Trial Block Structure and Feedback Contingencies* section. After the completion of Phase 4, the neither comparison was then included on all trials presented in Phases 5, 16, and 17 only.

Phase 5: Three-choice primary generalization test. Primary generalization gradients were obtained by presenting the same type of test block used in Phase 3, with the addition of the neither comparison on all trials. The use of the neither comparison during the test allowed subjects to select one of the line comparisons in the presence of a given sample without forcing a complementary change in the likelihood of selecting the other line comparison. The availability of the neither comparison, therefore, made the gradients measured with 1-unit and 25-unit lines potentially independent of each other.

Phases 6 through 13: Three-member equivalence class formation. Two three-member equivalence classes were established by the presentation of trial blocks using a simple-to-complex training and testing protocol (Adams, Fields, & Verhave, 1993a; Fields, Newman, Adams, & Verhave, 1992; Fields et al., 1991; Lynch & Cuvo, 1995; Schusterman & Kastak, 1993). The stimuli used in each block are listed symbolically in Table 2. The protocol involved training AB, testing BA symmetry, training BC, testing CB symmetry, reviewing BA and CB, testing AC transitivity, and then testing CA equivalence. Finally, all of the probes for emergent relations and the trained relations (as mentioned above) were reviewed in a mixed test. All test blocks contained 50% training trials. No informative feedback was provided on any trials presented in the test blocks for emergent relations.

Phases 14 and 15: Expansion to four-member equivalence classes. After the formation of the three-member classes, CD conditional relations were trained. The expansion of class size was then assessed with DC, BD, AD, DB, and DA probes, all of which were presented in a random order in the same test block.

Table 2

Symbolic representation of stimulus triads in Phases 6–15. In each trial block, each positive comparison (Co+) appeared equally often on the left and right. The asterisk indicates that when less than 100% feedback was scheduled, each triad appeared only four times in a block. The plus indicates that when less than 100% feedback was scheduled, each triad appeared only twice. Triads are listed for Class 1 only. In each phase, parallel triads were also presented for Class 2.

Sa	Co+	Co−	Presented
Phase 6: Train AB			
A1	B1	B2	8*
Phase 7: BA symmetry test			
A1	B1	B2	8
B1	A1	A2	8
Phase 8: Train BC			
A1	B1	B2	4+
B1	C1	C2	4+
Phase 9: CB symmetry test			
A1	B1	B2	4
B1	C1	C2	4
C1	B1	B2	8
Phase 10: BA and CB symmetry tests			
A1	B1	B2	4
B1	C1	C2	4
B1	A1	A2	4
C1	B1	B2	4
Phase 11: AC transitivity test			
A1	B1	B2	4
B1	C1	C2	4
A1	C1	C2	8
Phase 12: CA equivalence test			
A1	B1	B2	4
B1	C1	C2	4
C1	A1	A2	8
Phase 13: Symmetry, transitivity, and equivalence tests			
A1	B1	B2	4
B1	C1	C2	4
B1	A1	A2	2
C1	B1	B2	2
A1	C1	C2	2
C1	A1	A2	2
Phase 14: Train CD			
A1	B1	B2	2
B1	C1	C2	2
C1	D1	D2	6
Phase 15: Symmetry, transitivity, and equivalence tests			
B1	A1	A2	2
C1	B1	B2	2
A1	C1	C2	2
C1	A1	A2	2
D1	C1	C2	4
B1	D1	D2	4
A1	D1	D2	4
D1	B1	B2	4
D1	A1	A2	4

Maintenance of control by the previously established emergent relations was assessed with BA, CB, AC, and CA probes presented in the same test block. No training trials were included, and no informative feedback was presented for selections in the test block.

Phase 16: Primary generalization test after equivalence class formation. The Phase 16 primary generalization test used the Phase 5 stimuli and procedure. This test assessed generalization among all test lines and the 1- and 25-unit lines that had become members of Classes 1 and 2, respectively.

Phase 17: Primary generalization test and generalization tests of symmetry and equivalence. Phase 17 assessed the extent to which the 25 different lines occasioned selection of each member of the two equivalence classes. Four types of tests were conducted. All 25 lines were used as samples in each of the four tests. D1 and D2 were the comparisons in the primary generalization test (DD). C1 and C2 were the comparisons in the generalization test of symmetry (DC). B1 and B2 and A1 and A2 were the comparisons in the generalization tests of equivalence (DB and DA). Each of the four tests occurred in separate blocks; each block was presented five times in the order DB, DC, DA, DD, DC, DB, DD, DA, DD, DA, DC, DB, DA, DD, DB, DC, DB, DA, DC, and DD. Each test block contained 50 trials. Each of the 25 lines was presented as a sample on two trials. Each comparison appeared equally often in the left and right positions.

Results

The results of Experiment 1 focus on three major outcomes: (a) the effects of the neither comparison on the breadth of the generalization gradients obtained prior to the formation of the equivalence classes, (b) the extent to which the performances observed on the primary generalization test conducted prior to the formation of equivalence classes predicted the performances on the same tests conducted after the formation of equivalence classes, and (c) the extent to which the performances occasioned by the primary generalization test obtained prior to equivalence class formation predicted the performances on the generalization tests of symmetry and equivalence.

Figure 3 illustrates the effects of the neither comparison on primary generalization test performances conducted prior to equivalence class formation. Regardless of the availability of the neither comparison, the longest of the test lines typically evoked selection of the 25-unit comparison line, as seen in the right column. With reductions in the lengths of the test lines, selection of the 25-unit comparison line declined systematically and eventually reached zero. Similarly, the shortest of the test lines evoked selection of the 1-unit comparison line on nearly all trials, as seen in the left column. With increases in the lengths of the shorter test lines, selection of the 1-unit comparison line declined systematically and eventually reached zero.

The middle column of Figure 3 shows the likelihood of selecting the neither comparison during the Phase 5 generalization test. When the neither comparison was available, it was used in varying degrees by each subject. For Subjects 268, 269, and 267, the availability of the neither comparison reduced the range of longer test lines that occasioned the selection of the 25-unit comparison line, and also reduced the range of shorter test lines that occasioned the selection of the 1-unit comparison line. As the sample line length decreased from 25 units, a reduction in the selection of the 25-unit comparison was complemented by a corresponding increase in the selection of the neither comparison. As the sample line length increased from 1 unit, a decrease in the selection of the 1-unit comparison was complemented by a corresponding increase in the selection of the neither comparison.

For Subjects 270 and 266, the introduction of the neither comparison reduced the range of longer test lines that occasioned the selection of the 25-unit comparison line. It did not, however, influence the range of shorter lines that occasioned the selection of the 1-unit comparison line. Although the neither comparison was substituted for the selection of the 25-unit comparison in the presence of some of the longer lines, it was not substituted for the selection of the 1-unit comparison for the shorter lines. For Subject 265, the use of the neither comparison, however, did not influence the range of stimuli that usually occasioned the selection of the 25-unit or the 1-unit comparison lines.

For Subjects 269, 268, 266, and 267, differ-

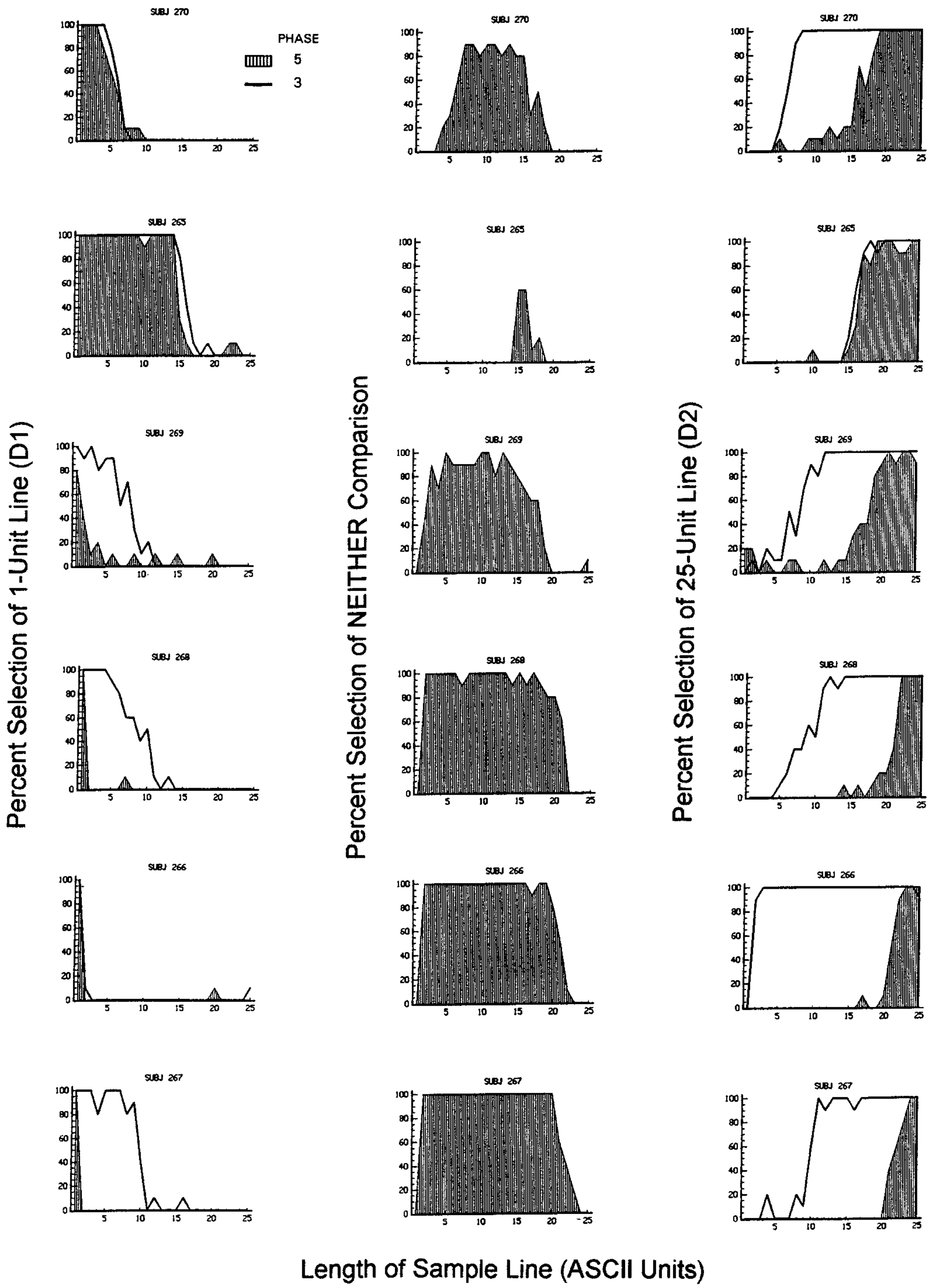


Fig. 3. The effect of introducing the neither comparision on primary generalization test performances conducted prior to equivalence class formation. The graphs in a row are for 1 subject. Each graph in the left and right columns contains a primary generalization gradient obtained when two comparisons were available in Phase 3 (line) and a gradient obtained when three comparisons were available in Phase 5 (filled area). The generalization gradients depicted in the right column show the likelihood of selecting the 25-unit comparison line. The generalization gradients in the left column show the likelihood of selecting the 1-unit comparison line. The graphs in the middle column show the likelihood of selecting the neither comparison during the generalization test that contained three comparisons. The legend indicates the phase number from which each gradient was obtained.

Table 3

Number of blocks needed to learn each baseline relation and pass each emergent relations test for all subjects in Experiment 1. AB, BC, and CD are the baseline relations that were trained. BA and CB were symmetry tests. 3MIX and 4MIX were test blocks that contained many different emergent relations probes used to assess three- and four-member equivalence class formation. The numbers below the types of relations indicate the percentage of trials in a block that occasioned informative feedback.

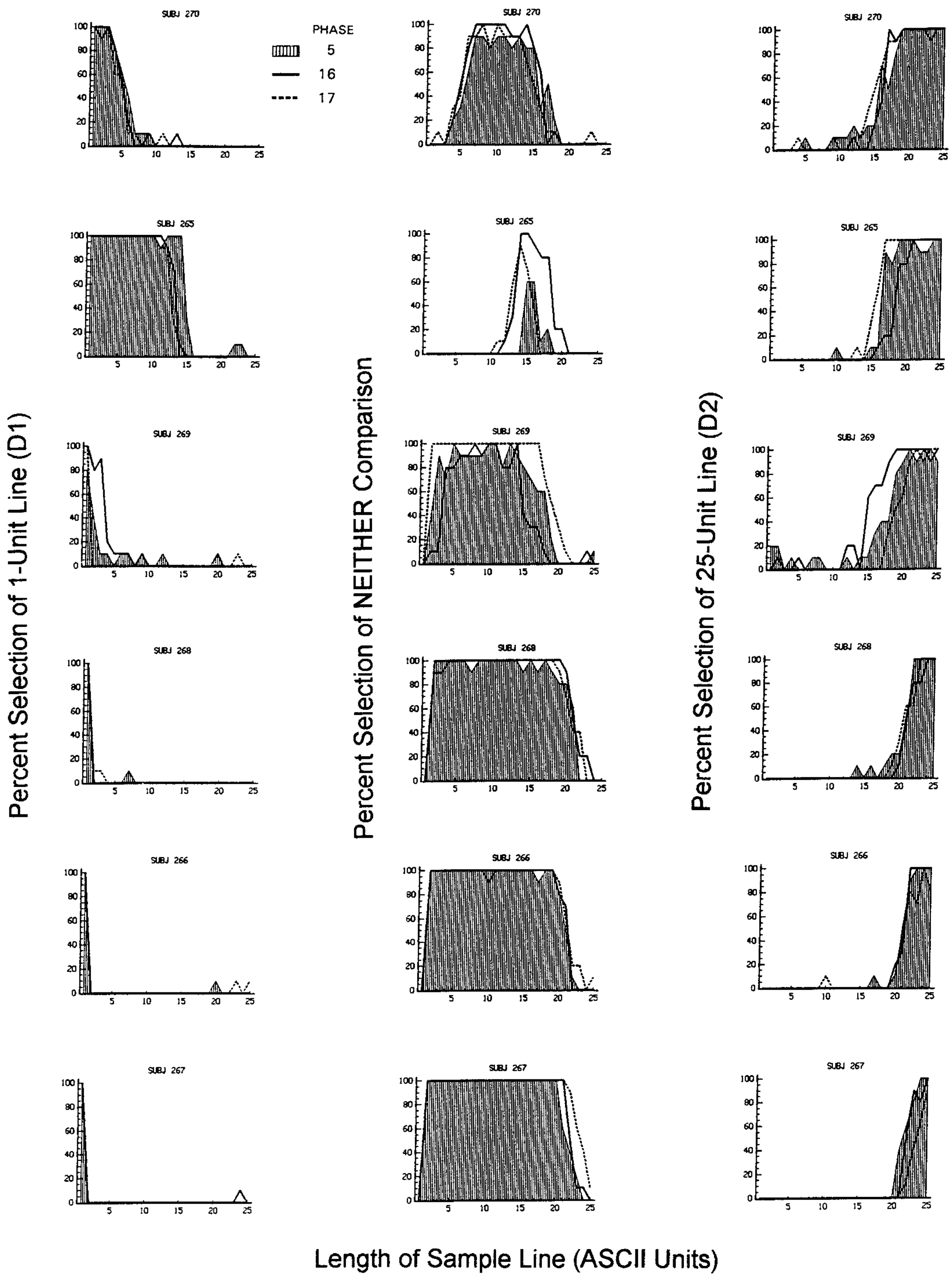
	Trial block type and percentage of feedback																		
Subject	AB 100	AB 75	AB 25	AB 0	BA 0	BC 100	BC 75	BC 25	BC 0	CB 0	BA & CB 0	AC 0	CA 0	3MIX 0	CD 100	CD 75	CD 25	CD 0	4MIX 0
270	2	1	1	1	1	2	1	1	1	1	1	1	2	1	2	4	1	1	2
265	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1
269	5	1	1	1	3	3	1	1	1	1	2	1	2	1	3	2	1	1	2
268	2	1	1	1	1	1	1	1	1	1	1	3	1	1	3	4	1	1	2
266	2	1	1	1	1	2	1	1	1	2	1	1	3	1	1	1	1	3	1
267	3	1	1	1	1	2	1	1	1	1	1	2	1	1	2	1	1	1	1
Average	2.5	1.0	1.0	1.0	1.3	1.8	1.0	1.0	1.0	1.2	1.2	1.7	1.7	1.0	2.0	2.2	1.0	1.3	1.5

ent ranges of adjacent test lines occasioned the exclusive selection of the neither comparison. For Subjects 270 and 265, however, no test lines occasioned the exclusive selection of the neither comparison. Instead, a range of adjacent test lines of intermediate length evoked the selection of the neither comparison or the 25-unit comparison. A separate range of adjacent test lines occasioned selection of the neither comparison or the 1-unit comparison. For all subjects, then, the test lines that occasioned the selection of the 1-unit comparison rarely occasioned the selection of the 25-unit comparison. Thus, the introduction of the neither comparison led to a functional separation of the generalization gradients occasioned by the selection of the 1-unit and the 25-unit lines.

Table 3 contains data showing the formation of equivalence classes for each subject. All 6 subjects formed the two four-member equivalence classes. All subjects acquired the baseline relations in one to three training blocks when 100% feedback was scheduled. Conditional discriminative control was maintained while feedback was decreased from 75% to 0% of the trials in a training block. Most emergent relations tests were passed in the first or second presentation of a test block. There was little intersubject variation in blocks to learn baseline relations or in blocks to pass given emergent relations tests.

Figure 4 compares performances observed during the three primary generalization tests conducted when the neither comparison was included as a selection option. The right column of Figure 4 shows the likelihood of selecting the 25-unit (Class 2) comparison line in the presence of the various test lines used as samples in each generalization test. For all Class 2 gradients, the longer test lines typically occasioned selection of the 25-unit comparison line that was a member of Class 2. As the length of the test line decreased, the likelihood of selecting the 25-unit line (the Class 2 comparison) declined systematically to zero. For all subjects except 267, the likelihood of selecting the Class 2 comparison line for a given sample test line did not differ systematically across the three gradients; in addition, the range of test lines that always occasioned the selection of the Class 2 line did not vary systematically across the three gradients. The range of stimuli that functioned in that manner, however, varied across subjects. Only for Subject 267 did repetition of the primary generalization test result in a clear sharpening of the generalization gradient.

The left column of Figure 4 indicates the likelihood of selecting the 1-unit (Class 1) comparison line in the presence of the various test lines. For Subjects 270 and 265, the test lines closest in length to the Class 1 line almost always occasioned selection of the Class 1 line comparison. As the length of the sample test line increased, the likelihood of selecting the Class 1 line comparison de-


PHASE
5
16
17
SUBJ 270
SUBJ 265
SUBJ 269
SUBJ 268
SUBJ 266
SUBJ 267
Percent Selection of 1-Unit Line (D1)
Percent Selection of NEITHER Comparison
Percent Selection of 25-Unit Line (D2)
Length of Sample Line (ASCII Units)

clined in gradual fashion to zero. In contrast, for Subjects 268, 266, and 267, the 1-unit comparison line was selected only when the 1-unit line was presented as a sample. Subject 269 showed very sharp overlapping gradients. The possible basis for this restricted generalization will be considered in the Discussion section. Thus, the shape of the generalization gradients, including the range of test lines that always occasioned selection of the Class 1 comparison, varied across subjects. They did not, however, differ systematically across the three generalization tests.

The middle column of Figure 4 illustrates the likelihood of selection of the neither comparison in the presence of each test line. As the length of the test line decreased from the 25-unit line, the systematic decline in the selection of the 25-unit comparison was complemented by an increase in the selection of the neither comparison. Conversely, as the length of the test line increased for the 1-unit line, the systematic decline in the selection of the 1-unit comparison was complemented by an increase in the selection of the neither comparison. Thus, for all subjects, a range of test lines of intermediate length almost always occasioned the selection of the neither comparison. This range, however, differed across subjects.

Figure 5 compares the results of the primary generalization test of line length obtained in Phase 5 with the tests of primary generalization, symmetry, and equivalence obtained in Phase 17. For all of the Class 2 gradients, the test lines closest in length to the Class 2 line almost always occasioned selection of all Class 2 word or line comparisons. As the length of the test line decreased, selection of the Class 2 word and line comparisons declined systematically and rapidly to zero.

For a given subject, a given test line evoked very similar likelihoods of selecting the comparisons that were the Class 2 words and the Class 2 line; in addition, the same range of test lines almost always occasioned the selection of all Class 2 stimuli. That range, however, varied across subjects. Thus, the likelihood of selecting a line or a word from Class 2 in the post-class-formation generalization tests was highly predicted by the performances observed in the primary generalization tests conducted prior to equivalence class formation.

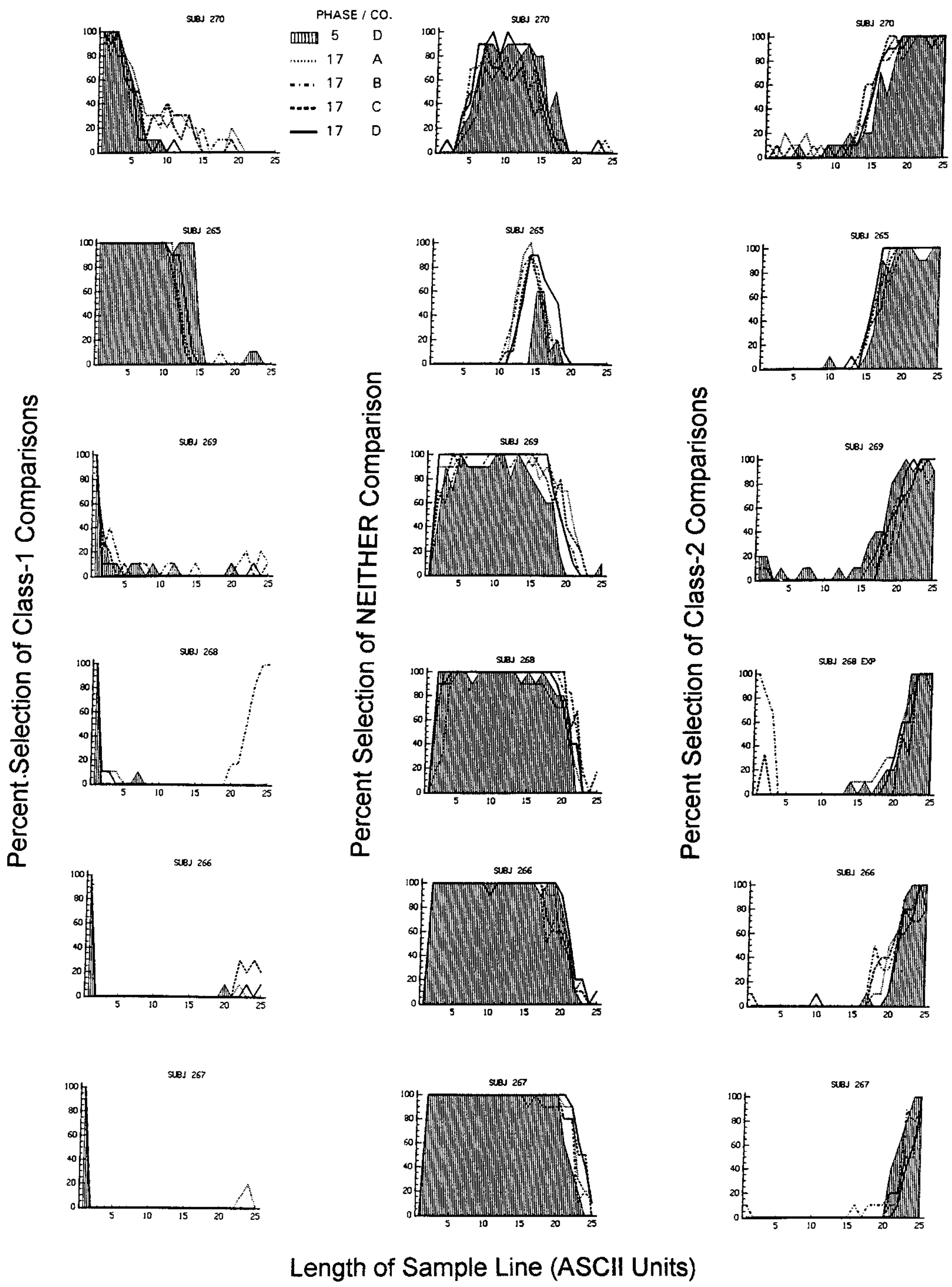
The left column of Figure 5 provides similar information on the likelihood of selecting the Class 1 comparisons (words and line) in the presence of the test lines. For all subjects, the Class 1 comparisons (words and line) were almost always selected when the 1-unit line was the sample. For Subjects 270 and 265, somewhat longer lines also occasioned selection of the Class 1 comparisons. For all subjects, however, as the length of the test lines increased, selection of the Class 1 comparisons declined systematically to zero. The range of test lines that almost always occasioned selection of the Class 1 comparisons did not vary across the test types. Thus, the Class 1 line or the Class 1 words were selected with the same likelihood in the presence of a given test line.

To summarize, the likelihood of selecting a line or a word from Class 1 in the post-class-formation generalization tests was highly predicted by the performances observed in the primary generalization tests conducted prior to equivalence class formation. The range of test lines that always occasioned selection of the Class 1 comparisons, however, varied across subjects.

The middle column of Figure 5 illustrates the likelihood of selecting the neither com-

←

Fig. 4. A comparison of the performances observed during primary generalization tests conducted prior to and after equivalence class formation. Each row of graphs shows data for 1 subject. The three functions in each graph represent the results of the primary generalization test conducted prior to equivalence class formation in Phase 5 (filled area), the primary generalization test conducted in isolation immediately after equivalence class formation in Phase 16 (solid line), and the primary generalization test conducted in the context of the generalization tests of symmetry and equivalence in Phase 17 (dashed line). The graphs in the right column show the likelihood of selecting the 25-unit comparison line in the presence of the test lines used as samples in each generalization test. The graphs in the left column contain generalization test data that indicate the likelihood of selecting the 1-unit comparison line in the presence of the test lines. The graphs in the middle column illustrate the likelihood of selecting the neither comparison in the presence of each test line. The legend indicates the phase number from which each gradient was obtained.


PHASE / CO.
5 D
17 A
17 B
17 C
17 D
SUBJ 270
SUBJ 265
SUBJ 269
SUBJ 268
SUBJ 268 EXP
SUBJ 266
SUBJ 267
Percent Selection of Class-1 Comparisons
Percent Selection of NEITHER Comparison
Percent Selection of Class-2 Comparisons
Length of Sample Line (ASCII Units)

parison in the presence of each test line. For all subjects, as the length of the test line decreased from the 25-unit line, the systematic decline in the selection of the 25-unit comparison was complemented by a corresponding increase in the selection of the neither comparison. As the test lines became shorter, the neither comparison was selected almost exclusively for some continuous range of test lines.

As test line length decreased further, two different patterns of usage emerged. For Subjects 270 and 265, there was a systematic decline in the selection of the neither comparison that was complemented by a corresponding increase in the selection of the 1-unit comparison. Eventually, the remaining range of short lines always occasioned the selection of the 1-unit comparison. For Subjects 269, 268, 266, and 267, however, once the neither comparison had been selected almost exclusively, it continued to be selected for all shorter test lines, with the exception of the 1-unit sample line. The 1-unit sample line never occasioned selection of the neither comparison.

The only anomaly in these results was a reversal in the generalization tests of the DB equivalence relation for Subject 268. The B2 comparison was selected when the D1 line was presented as a sample. As the length of the test line increased, the likelihood of choosing B2 decreased systematically to zero. The intermediate test lines occasioned the selection of the neither comparison. An increase in line length resulted in an increase in the likelihood of selecting B1. Finally, the longest test lines occasioned the exclusive selection of the B1 stimulus. Although they were not experimenter defined, these systematic performances suggest that the gradient in which B1 was chosen can be compared to the gradients in which A2, C2, and D2 were chosen. If this is done, the gradient obtained when measuring the choice of B1 overlapped the gradients obtained when measuring the choice of A2, C2, and D2. Likewise, the gradients obtained when measuring the choice of B2 overlapped the gradients obtained when measuring the choices of A1, C1, and D1. When compared in this way, the generalization gradients for emergent relations also overlapped, although not to the same degree as with the other subjects. These stable, isolated, and idiosyncratic reversals of an emergent relation have also been reported by Pilgrim and Galizio (1990, 1995), Pilgrim, Chambers, and Galizio (1995), and Saunders, Saunders, Kirby, and Spradlin (1988, 1990). In addition, the fact that the generalization gradients were consistent with the reversal provides additional evidence of the stability of the reversal of isolated emergent relations.

To summarize, the primary generalization gradients of line length obtained prior to equivalence class formation were highly predictive of the post-class-formation primary generalization gradients and generalization gradients of symmetry and equivalence. The range of variants that occasioned the selection of the long line during the generalization tests conducted prior to equivalence class formation was very similar to the line variants that functioned as members of the equivalence class that contained the long line as a class member. Likewise, the range of variants that occasioned the selection of the short line during the generalization tests conducted prior to equivalence class formation were very similar to the line variants that functioned as members of the equivalence class that contained the short line as a class member. Thus, many line variants became in-

←

Fig. 5. A comparison of the results of the primary generalization test of line length obtained in Phase 5 prior to equivalence class formation with the post-class-formation tests of primary generalization, symmetry, and equivalence obtained in Phase 17. Each row of graphs shows data for 1 subject. The five functions in each graph represent the results of the primary generalization test conducted prior to equivalence class formation (filled area), as well as results of the post-class-formation DD primary generalization test (solid line), the generalization test of DC symmetry (dashed line), the generalization test of DB equivalence (alternation of dots and dashes), and the generalization test of DA equivalence (dots). The graphs in the right column show the likelihood of selecting the word and line comparisons that were members of Class 2 in the presence of the test line used as samples. The graphs in the left column indicate the likelihood of selecting the Class 1 comparisons (words and line) in the presence of the test lines. The middle column illustrates the likelihood of selecting the neither comparison in the presence of each test line. The legend indicates the phase number from which each gradient was obtained as well as the letter designation of the comparisons (CO) presented in each type of test.

terchangeable with the words in an equivalence class. The fact that all subjects selected the neither comparison for some intermediate range of test lines led to the functional separation of the generalization gradients based on the measurement of the Class 2 and the Class 1 stimuli.

Discussion

The results of Experiment 1 raise a number of issues. These include (a) the effects of the neither comparison on generalization, (b) the asymmetry of Class 1 and Class 2 gradients, (c) the substitutability of stimuli in equivalence classes, (d) the prediction of the generalization gradients of emergent relations, and (e) the prediction of test lines that functioned as members of a generalized equivalence class.

The effects of the neither comparison on generalization. For many subjects, the introduction of the neither comparison resulted in a narrowing of primary generalization gradients. Nonoverlapping ranges of test lines occasioned the selection of either the Class 1 comparison and the neither comparison or the Class 2 comparison and the neither comparison. Therefore, the selection of the neither comparison led to the functional separation of the primary generalization gradients that measured the behavioral similarity of the test lines to the 1-unit and 25-unit lines. As a consequence, the extension of each equivalence class was measured independently of the other. These results replicate those previously reported in Experiment 2 by Fields, Adams, Brown, and Verhave (1993).

Asymmetry of Class 1 and Class 2 gradients. For 4 of the subjects in Experiment 1, the line and words in Class 1 were selected only in the presence of the 1-unit test line during the generalization tests. In contrast, a range of different test lines occasioned the selection of the line and the words in Class 2. These data suggest a perceptual discontinuity between the 1-unit line and the remaining test lines. One likely account can be adduced by considering that stimulus orientation, a feature of the test lines that covaries with length, can be used to dichotomize the test lines. As illustrated in Figure 1, the lines that are 3 to 25 units in length are all rectangular in shape and are wider than they are tall; thus, they appear to be horizontally oriented. The 2-unit line appears to be square; it is a neutral point with respect to horizontal-vertical orientation and has no dominant orientation. The 1-unit line is taller than it is wide; unlike all of the other test lines, it appears to be vertically oriented.

The performances occasioned by the 1-unit line relative to all others suggests that these 4 subjects were responding to the orientation of the test lines rather than to their absolute length. Previous studies have shown that different subjects can and will attend to different aspects of complex stimuli presented under the same experimental conditions (Cheng & Spetch, 1995; Wright, Cook, & Kendrick, 1989). Because the reinforcement contingencies used in Experiment 1 did not rule out attention to line orientation, some subjects may have attended to orientation where possible, and then to the length for the lines that had the same orientation. The 1-unit line was responded to as a singular vertically oriented stimulus, whereas the remaining horizontally oriented stimuli were responded to according to length, with the 2-unit line serving as its endpoint. This interpretation would also provide a plausible account of similar asymmetries in the generalization gradients of emergent relations noted by Fields, Adams, Brown, and Verhave (1993) and Fields et al. (1996).

Substitutability of stimuli in equivalence classes. Fields, Adams, and Verhave (1993), Hayes (1991), and Sidman (1994) noted that a test other than the emergent relations tests used to define the emergence of an equivalence class should be conducted to confirm the interchangeability of the stimuli in the class. The performances during the generalization tests of emergent relations provided such a demonstration. After the formation of equivalence classes, all of the stimuli in each equivalence class were selected with the same likelihood in the presence of a given test line. These performances provided a demonstration of the interchangeability of the stimuli in the equivalence classes and replicated the findings reported by Fields, Adams, Brown, and Verhave (1993). The fact that the class members functioned interchangeably, regardless of the likelihood of selecting all class members in the presence of a given test line, attests to the generality and robustness of the substitutability of the stimuli in an equivalence class (de Rose, McIlvane, Dube, Galpin, & Stoddard, 1988; Fields, Adams, Brown, & Verhave, 1993; Fields, Adams, Verhave, & Newman, 1993; Hayes, Kohlenberg, & Hayes, 1991; Mackay & Sidman, 1984; Wulfert & Hayes, 1988).

Predicting the generalization of emergent relations. The results of Experiment 1 demonstrated the essential overlap of performance across six generalization tests, all of which involved the presentation of test lines as samples. Some of the test lines in the pre-class-formation primary generalization test almost always occasioned the selection of the 1-unit or 25-unit comparison line. These same test lines always led to the selection of the corresponding class words in the post-class-formation generalization tests. The performances on the pre-class-formation tests, then, predicted rather precisely the test lines that came to function as members of the equivalence classes. Thus, the results of Experiment 1 show that the extension of equivalence class membership to the variants of one member of the class can be predicted by the prior measurement of primary generalization among the variants of the class member.

Predicting membership of a generalized equivalence class. A generalized equivalence class consists of the members of an equivalence class along with the dimensional variants of each class member, each of which occasions the selection of the remaining class members in emergent relations tests (Adams et al., 1993b; Fields et al., 1996). The generalization tests of emergent relations identified lines of intermediate length that functioned as members of Class 1 or Class 2. These test lines and the original members of the equivalence classes constituted two generalized equivalence classes. The performances observed in the pre-class-formation primary generalization test of line lengths predicted rather precisely the range of test lines that came to function as members of each generalized equivalence class.

EXPERIMENT 2

The results of Experiment 1 showed that some test lines functioned as members of each equivalence class. There are two possible accounts for the inclusion of the test lines in an equivalence class. On the one hand, the test lines that were functioning as members of each equivalence class, although different physically, might not have been discriminable from each other. In that case, the extension of membership in the equivalence class would reflect a subject's failure to distinguish the test lines from the line that was a member of the equivalence class. This argument is similar to that proposed by Lashley and Wade (1946) to account for constancy of responding across stimuli presented in primary generalization tests. If the extension of an equivalence class includes test lines that were not distinguishable from the lines that are members of that class, such an extension would be trivial (Lea, 1984; Wasserman, Kiedinger, & Bhatt, 1988).

On the other hand, the test lines that were functioning as class members could have been discriminable from each other. If so, each of these lines would have evoked the same selection responses, even though the response had been trained to occur in the presence of only the line that was a member of an equivalence class. Therefore, this set of discriminable lines would be functioning as a perceptual category (Adams et al., 1993b; Goldiamond, 1962; Keller & Schoenfeld, 1950; Lea, 1984; Wasserman et al., 1988). Because the test lines were arrayed along a simple physically defined dimension, length, the set of discriminably different lines that occasioned the same response would constitute a dimensionally defined class, one type of perceptual category. If so, the extension of each equivalence class in Experiment 1 reflected its merger with a dimensionally defined class.

Identification of the appropriate account of the results of Experiment 1 requires an independent measure of the discriminability of the lines that functioned as members of each equivalence class. Therefore, discriminability functions were obtained in Experiment 2 and were compared with the generalization gradients obtained in Experiment 1. The comparison determined whether the extension of the equivalence classes reflected the failure to discriminate among line lengths or the merger of each equivalence class with a dimensionally defined class of lines.

METHOD

Subjects and Apparatus

The 3 participants in Experiment 2, Subjects 265, 267, and 268, had participated in

Experiment 1. The apparatus used in Experiment 1 was also used in Experiment 2. The stimuli used in Experiment 2 were the 25 line lengths used in Experiment 1.

Procedure

Subjects were presented with eight repetitions of the trial block used in the Phase 5 primary generalization test in Experiment 1. In Experiment 2, however, the selection of the 1-unit line was reinforced when the sample was the 1-unit line; the selection of the 25-unit line was reinforced when the sample was the 25-unit line. The selection of the neither comparison was reinforced when the 2-through 24-unit lines were presented as samples. Feedback was given on all trials for the first two blocks and on 50% of the trials for a third block. No informative feedback was presented for the remaining five blocks. These contingencies were intended to maximize the discrimination between the intermediate-length lines and the 1-unit and 25-unit lines that were members of Classes 1 and 2, respectively (D. Blough & Blough, 1977; Hamilton & Coleman, 1933; Wright, 1972; Wright & Cumming, 1971).

RESULTS

The likelihood of choosing the 1-unit and the 25-unit lines was measured in the presence of each test line. High levels of stimulus control developed during the first few training trials and were maintained during all subsequent test blocks.

Figure 6 contains data for the 3 subjects in Experiment 2. The discriminability functions obtained for the selection of the 25-unit comparison were similar across subjects. When the 25-unit line was the sample, all subjects selected the 25-unit comparison line on almost all trials. The 25-unit comparison was selected with decreasing frequency in the presence of increasingly shorter sample lines. These stimuli were increasingly discriminable from the 25-unit line and were also discriminable from each other. The Class 2 comparison was never selected in the presence of test lines shorter than the 21-unit line by Subjects 267 and 268 and the 22-unit line by Subject 265. Therefore, the Class 2 line was completely discriminable from all test lines shorter than 21 units.

The discriminability functions obtained with the selection of the 1-unit comparison were the same for all subjects. The 1-unit comparison was always selected in the presence of the 1-unit sample and was never selected in the presence of any other test line. Thus, the 1-unit line was completely discriminable from all other test lines.

When the discriminability functions are compared with their corresponding generalization gradients, similar results were obtained in Class 1 for Subjects 268 and 267 and in Class 2 for Subject 267. The discriminability function overlapped with the generalization gradients. None of the test lines functioned as members of the equivalence classes. Rather, the likelihood of selecting Class 1 or Class 2 stimuli in the generalization tests reflected the discriminability of the test lines from the lines that were equivalence class members.

In the generalization tests obtained for Subject 265, the 21- to 17-unit lines usually occasioned the selection of the Class 2 comparisons and the 13- to 2-unit lines always occasioned the selection of the Class 1 comparisons. These test lines, however, never occasioned the selection of a class-based comparison in the discriminability test. Thus, the primary generalization gradients obtained in Classes 1 and 2 were much broader than the corresponding discriminability functions obtained for Classes 1 and 2, respectively.

For Subject 265, during the generalization tests, the Class 2 line was always selected in the presence of the 24- and 21-unit test lines. During the discriminability tests, however, the same test lines occasioned the selection of the Class 2 line with successively lower likelihoods. A similar pattern was observed in the Class 2 data for Subject 268. These data show that the test lines were discriminable from the 25-unit line that was a member of Class 2 and were also discriminable from each other.

DISCUSSION

Each discriminability function obtained in Experiment 2 identified test lines that were discriminable from the lines that were directly trained as equivalence class members in Experiment 1. For Classes 1 and 2 with Subject 265 and for only Class 2 with Subject 268, the test lines that functioned as class members in the generalization tests were clearly discriminable from the line that was directly trained

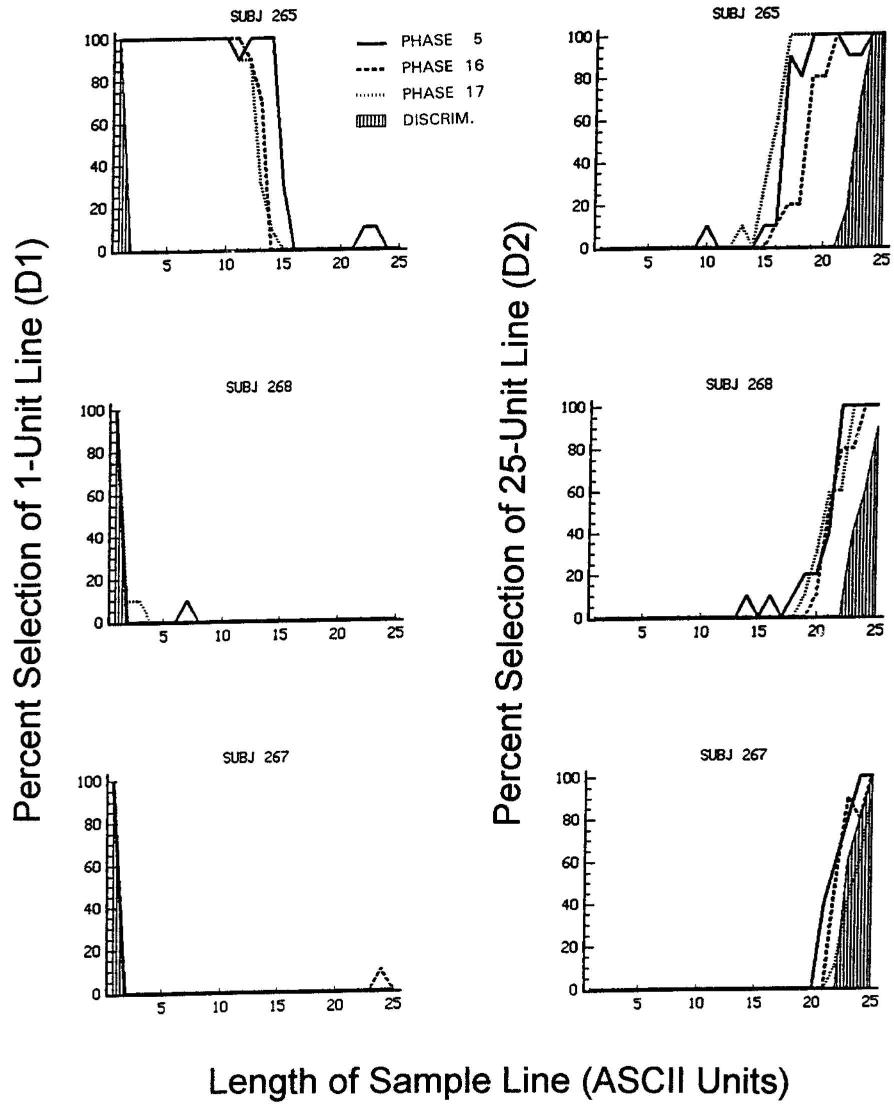


Fig. 6. Discriminability of test lines from 1-unit and 25-unit lines compared to the primary generalization gradients obtained in Experiment 1. The graphs in a row are for the same subject. The graphs in the right column contain data based on the selection of the Class 2 line in the presence of each test line. The graphs in the left column contain data based on the selection of the Class 1 line in the presence of each test line. Each graph contains three functions. The function defined by the shaded area shows the discriminability functions obtained in Experiment 2. These functions were obtained by averaging across the last five test blocks. Data were averaged because there were no systematic trends across individual test blocks. The other functions in each graph are the results of the three primary generalization tests conducted before and after class formation in Experiment 1 (Phases 5, 16, and 17). The legend indicates the phase number from which each gradient was obtained in Experiment 1 as well as the discriminability function (DISCRIM) obtained in Experiment 2.

as a class member. The extension of each equivalence class, then, could not have resulted from a failure to discriminate lines of different lengths. Rather, these lines were functioning as members of a stimulus class in Experiment 1 (Keller & Schoenfeld, 1950; Lea, 1984; Wasserman et al., 1988) that was dimensionally defined. For these cases, the extension of the equivalence classes observed in Experiment 1 resulted from a merger of the equivalence class with a dimensionally defined stimulus class (McIlvane, Dube, Green,

& Serna, 1993; Stromer, Mackay, & Remington, 1996).

The 1-unit line was a member of both the dimensional class of ‘‘short lines’’ and Equivalence Class 1; likewise, the 25-unit line was a member of the dimensional class of ‘‘long lines’’ and Equivalence Class 2. The 1-unit and 25-unit stimuli functioned to link each dimensional class with an equivalence class. The 1-unit and 25-unit stimuli were functioning as nodal stimuli to link classes of different types. More commonly, nodal stimuli have been implicated in the linkage of stimuli drawn from different conditional relations (Fields & Verhave, 1987; Fields et al., 1984) and in the linkage of two equivalence classes (Saunders et al., 1988; Sidman, Kirk, & Willson-Morris, 1985; Williams, Saunders, Saunders, & Spradlin, 1995). The linking of different types of stimulus classes, then, is another function that can be served by nodal stimuli.

Step-like discriminability functions were obtained for Subjects 267 and 268 when measuring the selection of the 1-unit comparison line. The 1-unit line was fully discriminated from all other test lines; thus, there was no extension of Equivalence Class 1 to intermediate line lengths. These performances probably reflected a discrimination of the 1-unit line from all others based on the perceived orientation of the lines rather than line length. The 1-unit line functioned as a rectangle with a vertical orientation that was categorically different from the other lines, which functioned as rectangles with horizontal orientations. There was no dimensional class of stimuli with vertical orientations that could merge with Equivalence Class 1.

Finally, in Experiment 2, the Class 2 data for Subject 267 showed that the 24- to 21-unit test lines were discriminable from each other and from the 25-unit line that was a member of Class 2. The discriminability function in Experiment 2 overlapped the last primary generalization gradient obtained in Experiment 1. In Experiment 1, then, the selection of the Class 2 comparisons in the generalization tests reflected the discriminability of the test lines from the Class 2 line.

Discriminability functions obtained for a given class in Experiment 2 were very similar across subjects. This overlap suggests that similar discriminability functions would also be obtained for most subjects. These hypothetical functions could be obtained for Classes 1 and 2 by averaging the data obtained from the 3 subjects in Experiment 2. Such hypothetical discriminability functions could be used to make plausible interpretations of the data obtained from Subjects 270, 269, and 266 in Experiment 1, all subjects for whom no discriminability data were obtained. When this approach is used, a visual inspection of Figure 3 suggests that the extension of an equivalence class by merger with a dimensionally defined class may have occurred in Class 2 for Subjects 270, 269, and 266 and in Class 1 for Subject 270. There may have been no extension of Equivalence Class 1 to intermediate-length test lines for Subjects 269 and 266 because the 1-unit line was fully discriminated from all other test lines. These interpreted data provide additional support for the view that the extension of equivalence classes by generalization reflects the merger of an equivalence class with a dimensionally defined class.

GENERAL DISCUSSION

The interpretation of primary generalization gradients. In Experiment 1, two equivalence classes were established with each of 6 subjects for a total of 12 classes. In the generalization tests of emergent relations, some contiguous range of test lines usually occasioned the selection of the comparison stimuli that were members of one equivalence class. That range of test lines could represent stimuli that were discriminable from each other and were functioning as members of a dimensionally defined class; alternatively, those test stimuli might not be discriminable from each other and thus would not be functioning as members of a dimensionally defined class. In seven of the 12 cases (five in Class 2 and two in Class 1), the test lines that occasioned such responding were discriminable from each other and thus were functioning as members of a dimensionally defined class. In the remaining five cases (one in Class 2 and four in Class 1), the test lines that occasioned such responding were not discriminable from each other and thus were not functioning as members of a dimensionally defined class.

This interpretation of the generalization test data obtained in Experiment 1 required a secondary set of measures that independently assessed the discriminability of the test lines that always occasioned the selection of the members of one equivalence class. The discrimination training procedure conducted in Experiment 2 provided evidence of discriminability. (Alternatively, we could have used a pseudodiscrimination training procedure, such as that described by Wasserman et al., 1988, or Lea, 1984, for the same purpose.) In any case, without such a secondary measure, it would not have been possible to determine whether test stimuli that occasioned the same response were discriminable from each other.

The results of Experiments 1 and 2 provide illustrative examples that primary generalization gradients do not necessarily reflect performances occasioned by stimuli that are functioning as members of a dimensionally defined class. On the one hand, high and invariant ceiling performances occasioned by some contiguous test lines could reflect a failure of the subjects to discriminate among those stimuli; those stimuli then would not be functioning as members of a dimensionally defined class. Only the edge of the gradient would reflect the range of stimuli that were discriminable from each other. On the other hand, high and invariant ceiling performances occasioned by some contiguous test lines could reflect membership of those test lines in a dimensionally defined class; in this case, these stimuli would be discriminable from each other. The edge of the gradient would reflect the limit of the dimensionally defined class and would also show that the stimuli were discriminable from each other.

When generalization gradients reflected control of behavior by dimensionally defined classes, the gradients were wider than they would have been had the test stimuli not been related through class membership. Thomas (1993) has also shown that the shape of a generalization gradient is not invariant. Rather, it is influenced by a wide number of experimental parameters as well as the relational properties of the stimuli used for training, testing, or both. Our finding that the shape of a generalization gradient is influenced by the membership of some of the test stimuli in a dimensionally defined class complements the view set forth by Thomas.

Stability of primary generalization gradients. Subject 267 was the only participant in Experiment 1 whose gradient showed a systematic sharpening with repeated testing. This shift in test performance was consistent with the known effects of repeated testing (P. Blough, 1971, 1972; Friedman & Guttman, 1965; Mishkin & Weiskrantz, 1959; Thomas & Barker, 1964), labeling (Dickins et al., 1993; Spradlin & Dixon, 1976), intradimensional discrimination training (Hanson, 1959), and extradimensional discrimination training (Honig, 1969; Thomas et al., 1970) on primary generalization. In contrast, the primary generalization gradients obtained in Experiment 1 for Subjects 270, 265, 269, 268, and 266, did not differ systematically with repeated testing. Such stability is surprising when one considers the effects of repeated testing, label training, or discrimination training on primary generalization test performances.

The stability of the gradients for Subjects 270, 265, 269, 268, and 266 can be understood by considering that different ranges of the test lines for these subjects were functioning as members of dimensionally defined classes. Stimuli that are members of a class occasion performances that tend to remain stable both over time (Saunders et al., 1990; Spradlin, Saunders, & Saunders, 1992) and in a given context (Bush et al., 1989; Lynch & Green, 1991; Meehan & Fields, 1995). The fact that the generalization gradients for these subjects did not vary with repeated testing and were relatively insensitive to the effects of interposed discrimination training is consistent with the view that different stimuli along the length dimension were functioning as members of distinct stimulus classes for Subjects 270, 265, 269, 268, and 266. This analysis also implies that different stimuli along the length dimension for Subject 267 were not functioning as members of distinct stimulus classes.

Extending the domain of generalized equivalence classes. Each generalized equivalence class established in Experiment 1 consisted of perceptually disparate stimuli that were members of an equivalence class and a range of discriminably different variants that occasioned the selection of the members of the equivalence class during generalization tests of emergent relations (Adams et al., 1993b). This set of variants also functioned as members of a dimensionally defined class. Thus, the extension of equivalence class membership reflected its merger with a dimensionally defined class.

A dimensionally defined class, however, is only one type of a perceptual category. Perceptual categories differ in terms of the continua used to array class members. In addition, all of the stimuli in such a class occasion the same response that is trained to some of the stimuli in the set (Goldiamond, 1962; Keller & Schoenfeld, 1950; Lea, 1984; Wasserman et al., 1988). Finally, all of the stimuli in a perceptual category resemble each other in varying degrees.

Some types of perceptual categories are feature classes (McIlvane et al., 1993; Stromer, Mackay, & Remington, 1996), basic level (Rosch & Mervis, 1975) or fuzzy (Wittgenstein, 1953) categories, and polymorphous categories (Jitsumori, 1993, 1994; Lea & Harrison, 1978). In a feature class, a common set of defining features is found in all exemplars. Feature classes are similar to the traditional Aristotelian notion of concept and were studied by Hull (1920) in his classic experiments on concept formation. Examples of feature classes include all words with the letter W and all people with red hair. In a dimensionally defined class, stimuli can be arrayed along a dimension defined by physics. Examples of dimensionally defined classes include hot-cold, long-short, high-low, and heavy-light. Dimensionally defined classes can also be arrayed along some complex mathematically derived dimension such as compactness (Hrycenko & Harwood, 1980). In a basic level or fuzzy category, all exemplars bear a family resemblance to each other but all class members do not share a common defining feature. Instead, the exemplars of a basic level category contain many features, the number of which is unspecified, with each exemplar containing some of the features. Exemplars can be arrayed along a psychometrically defined dimension based on the number of features in an exemplar weighted by the prevalence of each feature among all class members. The perceived resemblance of two exemplars is a direct function of position along such a dimension. Examples of basic level categories include pictures of chairs and pictures of cars. A polymorphous category is a set of stimuli that is defined by exactly n features. Each member of a polymorphous class contains any combination of at least m of the n features. An example of a polymorphous category would be a set of stimuli that contained five shapes, at least two of which were different. Other examples include the use of m of n symptoms to identify individuals with autism and chronic fatigue syndrome.

The results of Experiments 1 and 2 support the view that the extension of an equivalence class by generalization represents the merger of an equivalence class with a dimensionally defined class. By analogy, such extension may also represent the merger of an equivalence class with any type of perceptual class. Regardless of the type of perceptual class, the prior specification of the stimuli that are members of a perceptual class should precisely predict the stimuli that will come to function interchangeably with the members of an equivalence class if one stimulus becomes a member of both classes. Thus, the merger of equivalence classes with a perceptual class may be an additive process in general.

Conversion of close-ended to open-ended categories. Some perceptual classes are open ended (Herrnstein, 1990); they consist of an unlimited number of exemplars that vary along a single physical dimension or along a multiplicity of dimensions. Such categories include, but are not limited to, dimensionally defined classes, basic level or fuzzy classes, and feature classes. In contrast, as usually presented, an equivalence class contains a specific number of members and thus is close ended.

When an equivalence class merges with an open-ended perceptual class, as seen in Experiment 1, the specific number of stimuli in the equivalence class is extended to an indefinite value. At that point, then, a close-ended equivalence class becomes an open-ended category that has an unlimited number of exemplars. Thus, the merger of an equivalence class with a perceptual class blurs the dichotomous characterization of classes as being open ended or close ended.

Naturally occurring categories. Adams et al. (1993b) and Fields et al. (1996) noted the similarity of generalized equivalence classes, naturally occurring categories, superordinate semantic categories (Medin & Smith, 1984; Rosch & Mervis, 1975), and natural kind categories (Gelman, 1988a, 1988b; Gelman & Markham, 1986, 1987). At the level of formal stimulus specification, all of these classes consist of some stimuli that are perceptually disparate and others that are perceptually similar. To illustrate, the superordinate semantic category of FURNITURE would consist of the written words COUCH, TABLE, and LAMP, the same three words as heard, many pictures of different couches, tables, and lamps, as well as many pictures of a given couch, table, or lamp taken from different distances and vantage points. At the level of functional properties, all of the stimuli in any of these classes would have to be interchangeable under some test conditions to conclude that they were all members of the complex class of furniture. Tests analogous to those used in Experiments 1 and 2 would provide a means of assessing such functional interchangeability. Thus, the stimuli that are members of a generalized equivalence class, such as that demonstrated in Experiment 1, bear striking formal and functional similarities to the stimuli that are members of the other classes named above.

If there are no fundamental differences among the so-called types of naturally occurring categories, the variables that account for the formation of generalized equivalence classes would also apply to these other classes (Adams et al., 1993b). Also, membership in these complex classes should be predictable from a prior knowledge of the membership in the perceptual categories that are components of the complex classes. Finally, a uniform account of the emergence of generalized equivalence classes, superordinate semantic categories, and natural kind categories should be provided by a specification of the procedures needed to establish equivalence classes and perceptual classes and to link different kinds of classes.

REFERENCES

Adams, B. J., Fields, L., & Verhave, T. (1993a). The effects of test order on the establishment and expansion of equivalence classes. *The Psychological Record, 43,* 133–152.

Adams, B. J., Fields, L., & Verhave, T. (1993b). Formation of generalized equivalence classes. *The Psychological Record, 43,* 553–566.

Balsam, P. D. (1988). Selection, representation, and equivalence of controlling stimuli. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (pp. 111–166). New York: Wiley.

Barnes, D., & Keenan, M. (1993). A transfer of functions through derived arbitrary and nonarbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior, 59,* 61–82.

Blough, D. (1983). Alternative accounts of dimensional stimulus control. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (pp. 59–74). New York: Harper & Row.

Blough, D., & Blough, P. (1977). Animal psychophysics. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 514–539). Englewood Cliffs, NJ: Prentice Hall.

Blough, P. (1971). The visual acuity of the pigeon for distant targets. *Journal of the Experimental Analysis of Behavior, 15,* 57–67.

Blough, P. (1972). Wavelength generalization and discrimination in the pigeon. *Perception and Psychophysics, 12,* 342–348.

Buffington, D. M., Fields, L., & Adams, B. J. (1997). Enhancing equivalence class formation by pretraining of other equivalence classes. *The Psychological Record, 47,* 69–96.

Bush, K. B. (1993). Stimulus equivalence and cross-modal transfer. *The Psychological Record, 43,* 567–584.

Bush, K. M., Sidman, M., & de Rose, T. (1989). Conditional control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 38,* 29–46.

Cheng, K., & Spetch, M. L. (1995). Stimulus control in the use of landmarks by pigeons in a touch-screen task. *Journal of the Experimental Analysis of Behavior, 63,* 187–201.

Cowley, B. J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to reach name-face matching to adults with brain injuries. *Journal of Applied Behavior Analysis, 25,* 461–475.

DeGrandpre, R. J., Bickel, W. K., & Higgins, S. T. (1992). Emergent equivalence relations between interoceptive (drug) and exteroceptive (visual) stimuli. *Journal of the Experimental Analysis of Behavior, 58,* 9–18.

de Rose, J. C., McIlvane, W. J., Dube, W. V., Galpin, V. C., & Stoddard, L. T. (1988). Emergent simple discrimination established by indirect relation to differential consequences. *Journal of the Experimental Analysis of Behavior, 50,* 1–20.

Dickins, D. W., Bentall, R. P., & Smith, A. B. (1993). The role of individual stimulus names in the emergence of equivalence relations: The effects of interpolated paired-associates training of discordant associations between names. *The Psychological Record, 43,* 713–724.

Fields, L. (1980). Enhanced learning of new discriminations after stimulus fading. *Bulletin of the Psychonomic Society, 15,* 327–330.

Fields, L., Adams, B. J., Brown, J. B., & Verhave, T. (1993). The generalization of emergent relations in equivalence classes: Stimulus substitutability. *The Psychological Record, 43,* 235–254.

Fields, L., Adams, B. J., Buffington, D. M., Yang, W., & Verhave, T. (1996). Response transfer between stimuli in generalized equivalence classes: A model for the establishment of natural kind and fuzzy superordinate categories. *The Psychological Record, 46,* 665–684.

Fields, L., Adams, B. J., Newman, S., & Verhave, T. (1992). Interactions of emergent relations during the formation of equivalence classes. *Quarterly Journal of Experimental Psychology, 45B,* 125–138.

Fields, L., Adams, B. J., & Verhave, T. (1993). The effects of equivalence class structure on test performances. *The Psychological Record, 43,* 697–712.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53,* 345–358.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1993). Are stimuli in equivalence classes equally related to each other? *The Psychological Record, 43,* 85–105.

Fields, L., Landon-Jimenez, D. V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal distance effects after equivalence class formation. *Journal of the Experimental Analysis of Behavior, 64,* 129–146.

Fields, L., Newman, S., Adams, B. J., & Verhave, T. (1992). The expansion of equivalence classes through simple discrimination training and fading. *The Psychological Record, 42,* 3–15.

Fields, L., Reeve, K. F., Adams, B. J., & Verhave, T. (1991). Stimulus generalization and equivalence classes: A model for natural categories. *Journal of the Experimental Analysis of Behavior, 55,* 305–312.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 49,* 317–332.

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.

Friedman, H., & Guttman, N. (1965). Further analysis of the various effects of discrimination training on stimulus generalization gradients. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 255–267). Stanford: Stanford University Press.

Galizio, M., & Baron, A. (1976). Label training and auditory generalization. *Learning and Motivation, 7,* 591–602.

Gelman, S. A. (1988a). Children's expectations concerning natural kind categories. *Human Development, 31,* 28–34.

Gelman, S. A. (1988b). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20,* 65–95.

Gelman, S. A., & Markham, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183–209.

Gelman, S. A., & Markham, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58,* 1532–1541.

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

Guttman, N., & Kalish, H. J. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology, 51,* 79–88.

Hamilton, W. F., & Coleman, T. B. (1933). Trichromatic vision in the pigeon as illustrated by the spectral hue discrimination curve. *Journal of Comparative Psychology, 15,* 183–191.

Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology, 58,* 321–334.

Haring, T. G., Breen, C. G., & Laitinen, R. E. (1989). Stimulus class formation and concept learning: Establishment of within- and between-set generalization and transitive relationships via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior, 52,* 13–26.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Reno, NV: Context Press.

Hayes, S. C., Kohlenberg, B. S., & Hayes, L. (1991). The transfer of specific and general consequential functions through simple conditional equivalence relations. *Journal of the Experimental Analysis of Behavior, 56,* 119–137.

Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition, 37,* 133–166.

Honig, W. K. (1969). Attentional factors governing the slope of the generalization gradient. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 35–62). London: Academic Press.

Honig, W. K., Boneau, C. A., Burstein, K. R., & Pennypacker, H. S. (1963). Positive and negative generalization gradients obtained after equivalent training conditions. *Journal of Comparative and Physiological Psychology, 56,* 111–116.

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36,* 405–445.

Hrycenko, O., & Harwood, D. W. (1980). Judgments of shape similarity in the Barbary dove (*Streptopelia risoria*). *Animal Behaviour, 28,* 586–592.

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs, 28* (1, Whole No. 132).

Jitsumori, M. (1993). Category discrimination of artificial polymorphous stimuli based on feature learning. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 244–254.

Jitsumori, M. (1994). Artificial polymorphous categories in humans and non-humans. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavioral analysis of language and cognition* (pp. 91–106). Reno, NV: Context Press.

Keller, F. S., & Schoenfeld, W. N. (1950). *The principles of psychology.* New York: Appleton-Century-Crofts.

Lashley, K. S., & Wade, M. (1946). The Pavlovian theory of generalization. *Psychological Review, 53,* 72–87.

Lea, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 263–276). Hillsdale, NJ: Erlbaum.

Lea, S. E. G., & Harrison, S. N. (1978). Discrimination of polymorphous stimulus sets in pigeons. *Quarterly Journal of Experimental Psychology, 30,* 521–537.

Lynch, D. C., & Cuvo, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis, 28,* 115–126.

Lynch, D. C., & Green, G. (1991). Development and crossmodal transfer of contextual control of emergent stimulus relations. *Journal of the Experimental Analysis of Behavior, 56,* 139–154.

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

McIlvane, W. J., Dube, W. V., Green, G., & Serna, R. W. (1993). Programming conceptual and communication skill development: A methodological stimulus class analysis. In A. P. Kaiser & D. B. Gray (Eds.), *Enhancing children's communication* (Vol. 2, pp. 242–285). Baltimore, MD: Brookes.

McIlvane, W. J., & Stoddard, L. T. (1981). Acquisition of matching-to-sample performance in severe retardation: Learning by exclusion. *Journal of Mental Deficiency Research, 25,* 33–48.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Reviews of Psychology, 35,* 113–138.

Meehan, E. F., & Fields, L. (1995). Contextual control of new equivalence classes. *The Psychological Record, 45,* 165–182.

Mishkin, M., & Weiskrantz, L. (1959). Effects of cortical lesions in monkeys on critical flicker frequency. *Journal of Comparative and Physiological Psychology, 52,* 660–666.

Pilgrim, C., Chambers, L., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: II. Children. *Journal of the Experimental Analysis of Behavior, 63,* 238–254.

Pilgrim, C., & Galizio, M. (1990). Relations between baseline contingencies and equivalence probe performances. *Journal of the Experimental Analysis of Behavior, 54,* 213–224.

Pilgrim, C., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior, 63,* 225–238.

Rilling, M. (1977). Stimulus control and inhibitory processes. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 432–480). Englewood Cliffs, NJ: Prentice Hall.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 50,* 145–162.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1990). Long term stability of equivalence relations in the absence of training or practice. *American Journal of Mental Retardation, 95,* 291–303.

Schusterman, R., & Kastak, D. (1993). A California sea lion (*Aalophus californianus*) is capable of forming equivalence relations. *The Psychological Record, 43,* 823–840.

Sidman, M. (1990). Equivalence relations: Where do they come from? In H. Lejeune & D. Blackman (Eds.), *Behavior analysis in theory and practice: Contributions and controversies* (pp. 93–114). Hillsdale, NJ: Erlbaum.

Sidman, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

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

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, K. J., & Saunders, R. R. (1992). The stability of equivalence classes. In S. C. Hayes & L. B. Hayes (Eds.), *Understanding verbal relations* (pp. 29–42). Reno NV: Context Press.

Stromer, R., Mackay, H. A., Howell, S. R., & McVay, A. A. (1996). Teaching computer-based spelling to individuals with developmental and hearing disabilities: Transfer of stimulus control to writing tasks. *Journal of Applied Behavior Analysis, 29,* 25–42.

Stromer, R., Mackay, H. A., & Remington, B. (1996). Naming, the formation of stimulus classes, and applied behavior analysis. *Journal of Applied Behavior Analysis, 29,* 409–431.

Thomas, D. R. (1993). A model for adaptation-level effects on stimulus generalization. *Psychological Review, 100,* 658–673.

Thomas, D. R., & Barker, E. G. (1964). The effects of extinction and ''central tendency'' on stimulus generalization in pigeons. *Psychonomic Science, 1,* 119–121.

Thomas, D. R., Freeman, F., Svinicki, J. G., Burr, D. E. S., & Lyons, J. (1970). Effects of extradimensional training on stimulus generalization. *Journal of Experimental Psychology, 83,* 1–21.

Urcuioli, P. J., & Nevin, J. A. (1975). Transfer of hue matching in pigeons. *Journal of the Experimental Analysis of Behavior, 24,* 149–155.

Wasserman, E. A., Kiedinger, R. E., & Bhatt, R. S. (1988). Conceptual behavior in pigeons: Categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 235–246.

Williams, D. C., Saunders, K. J., Saunders, R. R., & Spradlin, J. E. (1995). Unreinforced conditional selection within three-choice conditional discriminations. *The Psychological Record, 45,* 613–628.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: Macmillan.

Wright, A. A. (1972). The influence of ultraviolet radiation on the pigeon's color discrimination. *Journal of the Experimental Analysis of Behavior, 17,* 325–337.

Wright, A. A., Cook, R. G., & Kendrick, D. F. (1989). Relational and absolute stimulus learning by monkeys in a memory task. *Journal of the Experimental Analysis of Behavior, 52,* 237–248.

Wright, A. A., & Cumming, W. W. (1971). Color-naming functions for the pigeon. *Journal of the Experimental Analysis of Behavior, 15,* 7–17.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 40,* 125–144.

Received August 17, 1992
Final acceptance March 18, 1997