JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2005, **84,** 243–267 NUMBER 2 (SEPTEMBER)

THE EFFECT OF TEST SCHEDULES ON THE FORMATION OF LINKED PERCEPTUAL CLASSES

LANNY FIELDS, ADRIENNE FITZER, KIMBERLY SHAMOUN,
PRIYA MATNEJA, MARI WATANABE, AND DANIELLE TITTELBACH

QUEENS COLLEGE OF THE CITY UNIVERSITY OF NEW YORK AND
THE GRADUATE CENTER OF THE CITY UNIVERSITY OF NEW YORK

After training conditional discriminations among selected stimuli from two perceptual classes, the emergence of novel relations involving other members of both classes was assessed using cross-class probes. The cross-class probes were presented using one of four different testing schedules. In the 2/9 test, nine different probes were presented in each of two test blocks. In the 6/3 test, three different probes were presented in each of six test blocks. In the 18/1-RND test, each of the 18 cross-class probes was presented in separate test blocks. In the 2/9 and 6/3 tests, the cross-class probes were presented in a randomized order within test block. In the 18/1-RND test, the cross-class probes were presented in a randomized sequence. In the 18/1-PRGM test, however, the cross-class probes were presented in a programmed order (i.e., the values of the stimuli in each cross-class probe were changed systematically in the succession of probe presentations). About 55% of the linked perceptual classes emerged during the 2/9, 6/3, and 18/1-RND tests. Thus the number of different probes in a test block did not influence the emergence of classes as long as the probes were presented in a random order. Virtually all classes emerged during the 18/1-PRGM test. Thus at least one ordered introduction of different cross probes resulted in the reliable emergence of linked perceptual classes. Mechanisms responsible for linked perceptual class formation are discussed along with the relation of these classes to other complex categories.

Key words: linked perceptual classes, testing schedule, cross-class probe, mutual selection and class definition, computer keyboard responding, college students

A perceptual class is defined in terms both of structure and function. Structurally, a class consists of the contiguous stimuli that can be arrayed along a physically or psychometrically defined dimension. Functionally, the stimuli act as members of a class when they all occasion either the same response or the mutual selection of each other in the absence of a direct history of reinforcement (Belanich & Fields, 2003; Fields & Reeve, 2000; Fields, Reeve, Adams, Brown, & Verhave, 1997; Keller & Schoenfeld, 1950; Reeve & Fields, 2001; Wright, Cook, Rivera, Sands, & Delius, 1988). This must occur despite the fact that some of the stimuli in the putative class are discriminable from each other in the same (Bhatt, Wasserman, Reynolds, & Knauss, 1988; Fields, Matneja, Varelas, et al., 2002; Fields, Reeve, et al., 2002) or different circumstances (Honig & Stewart, 1988; Lea, 1984; Wasserman, Keidinger, & Bhatt, 1988).

In addition to being related to each other, the stimuli in a perceptual class can be, and typically are, related to stimuli in at least one other distinct perceptual class. When that occurs, the stimuli in the two distinct perceptual classes function as members of a single *linked perceptual class* (Fields & Reeve, 2001). One example of a linked perceptual class is the range of pictures of monkeys (one perceptual class) and the word monkey written in many fonts, sizes, or degrees of distortion (another perceptual class). Another example would be the three-dimensional images of leopards and the variety of sounds made by leopards. A third example would be X-rays of malignant breast tumors and a variety of lumps in breast tissue sensed by palpation. These examples suggest the ubiquity of linked perceptual classes that can be found in natural settings.

This research was conducted with support from Contract DASW01-96-K-0009 from the U.S. Army Research Institute, and by PSC-CUNY Research Awards 68547, 69567, and 61617. We thank Xiqiang Zhu for his assistance in the development of the software used to conduct the experiment and analyze the data reported herein. We thank Tricia Moss, Michelle Garruto, and Michael Marroquin for their able assistance with the data analysis and critical commentary during the preparation of the manuscript.

Reprints can be obtained from Lanny Fields, Department of Psychology, Queens College/CUNY, 65-30 Kissena Boulevard, Flushing, New York 11367 (e-mail: Lanny_Fields@qc.edu).

doi: 10.1901/jeab.2005.45-03

To date, only one empirical study has described the formation of linked perceptual classes. These classes were formed from pairs of distinct perceptual classes labeled A1`, A2`, B1`, and B2`, as illustrated in Figure 1 (Fields, Matneja, et al., 2002). Three key stimuli were defined for each class. One endpoint of each perceptual class was called the anchor (X^a). The other endpoint of each class, that most removed from the anchor yet still judged to be related to the anchor, was called the boundary (X^b) stimulus. The midpoints of each class (X^m) were perceptually equidistant from the anchor and the boundary stimuli in that class. The A1` and B1` classes were linked by establishing conditional discriminations between the anchors of the classes and the boundaries of the classes with $A1^a \rightarrow B1^a$ and $A1^b \rightarrow B1^b$ conditional discriminations. Likewise, the A2` and B2` classes were linked by establishing conditional discriminations between the anchors of the classes and the boundaries of the classes with $A2^a \rightarrow B2^a$ and $A2^b \rightarrow B2^b$ conditional discriminations.

The emergence of a linked perceptual class was assessed with the presentation of 18 cross-class probes that consisted of all of the possible combinations of the anchor, midpoint, and boundary stimuli from the two potentially related perceptual classes, all of which are listed in the lower section of Figure 1. During A`→B` probes, the anchor, midpoint, and boundary stimuli from the A classes were presented as samples in combination with pairs of the anchor, midpoint, and boundary stimuli from the B classes presented as comparisons. Likewise, during B`→A` probes, the anchor, midpoint, and boundary stimuli from the B classes were presented as samples in combination with pairs of the anchor, midpoint, and boundary stimuli from the A classes as comparisons. A linked perceptual class was said to have emerged when at least 17 of the 18 cross-class probes occasioned class-consistent comparison selections. Thus the emergence of a linked perceptual class was documented by the mutual selection of the stimuli in the two classes. Because two linked perceptual classes could have been formed for each of the 5 subjects, a maximum of 10 linked perceptual classes could have emerged in the experiment. The initial presentation of the cross-class tests evoked class-consistent comparison selections for 7 of the possible 10 linked perceptual classes. This study measured the formation of linked perceptual classes with a single set of training and testing procedures. It did not, however, identify any variables that might have influenced the formation of those classes. The identification of those variables would increase our understanding of the factors responsible for the establishment of these categories that appear to be ubiquitous in natural settings,

Many studies have shown that parameters of training and testing can influence the formation of a stimulus class. The present study explored the effects of two testing variables on the formation of linked perceptual classes: probe types per test block, and programmed introduction of probes of different types. Prior research has shown that the likelihood of equivalence class formation is an inverse function of the number of different emergent relations probes that are presented in a test block (Adams, Fields, & Verhave, 1993a; Buffington, Fields, & Adams, 1997; Fields, Landon-Jimenez, Buffington, & Adams, 1995; Fields, Reeve, Rosen, et al., 1997;). Similar effects might be expected for linked perceptual class formation.

In the Fields, Matneja, et al. (2002) study, testing involved the presentation of three different types of cross-class probes in each of six successive test blocks. The combination of probe types per block and number of blocks, however, could be manipulated systematically. The present experiment required that 18 different types of cross-class probes be presented to document the formation of a linked perceptual class. The formation of a linked perceptual class could be evaluated with the presentation of X probe types per block in X/18 successive test blocks. The present experiment studied the effects of test blocks that contained one, three, and nine different probe types per block.

In the Fields, Matneja, et al. (2002) study, the different probe types used in a block were presented in a randomized order. Prior research has shown that the likelihood of equivalence class formation is influenced by the order of introducing the different emergent relations probes (Adams et al., 1993a). Thus ordering the presentation of the cross-class probes might enhance the formation of linked perceptual classes. The effect of the ordered introduction of cross-class probes on

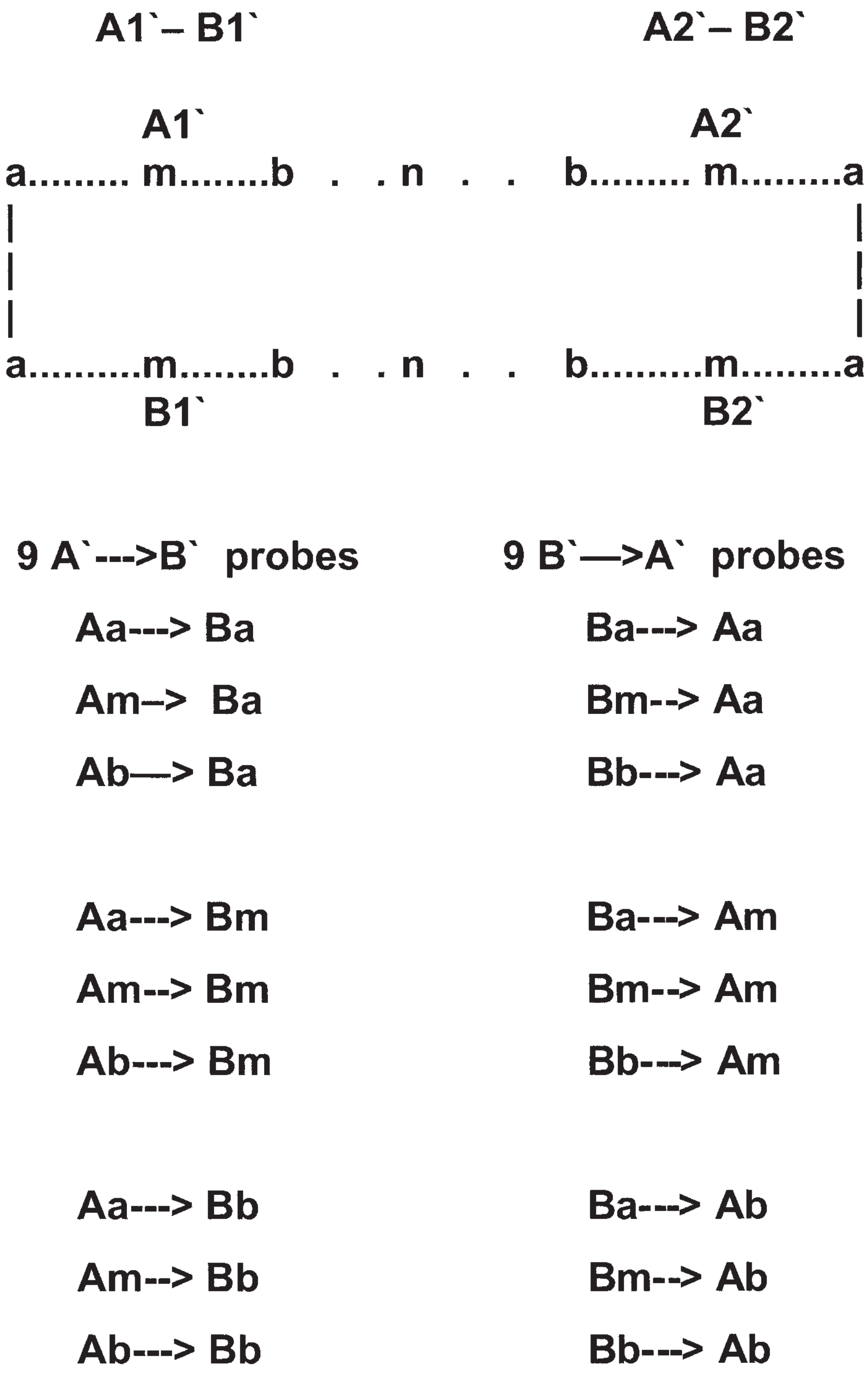

Fig. 1. Schematic diagram of two linked perceptual classes constructed from classes consisting of stimuli at opposite ends of two distinct stimulus domains. The relations shown at the bottom of the figure include the cross-class probes needed to assess the emergence of conditional relations among key stimuli in each perceptual class.

the formation of linked perceptual classes was evaluated by a comparison of two testing conditions. In each, one type of cross-class probe was presented per test block in a sequence of 18 test blocks. In one condition, the different probe types were presented in a randomized sequence. In the other condition, the different probe types were presented in an ordered and programmed sequence. If ordering is a critical variable, then the formation of linked perceptual classes should be enhanced when testing is conducted in an ordered sequence.

METHOD

Participants

Participants were 34 undergraduate students enrolled in an advanced experimental psychology course at Queens College/CUNY. They reported no familiarity with the research area and had not participated in prior experiments in the laboratory. Upon completion of the experiment, students received course credit from their instructors. The experiment lasted 3 to 4 hr per student depending on performance, and was divided into two sessions conducted on different days within a 1-week period. Volunteers were randomly assigned to four groups of $N = 12$, although unequal groups resulted from some subjects failing to return for the second laboratory visit. It is important to note, however, that attrition occurred prior to any experimental manipulation. Thus subject mortality could not be attributed to the experimental manipulations, and it is unlikely that mortality was responsible for differences in outcomes across groups.

Apparatus and Stimuli

The experiment was conducted on an IBM®-compatible computer that displayed all stimuli on a color monitor. Responses consisted of touching specific keys on a standard keyboard. The experiment was controlled by custom software that programmed all stimulus presentations and recorded all keyboard responses.

All stimuli were presented in colored squares without a contrasting border (2.5 by 2.5 cm) on the computer monitor. Stimuli from six domains were used in the experiment. Preliminary training was conducted with stimuli in four domains: female–male, abstract pictures, truck–car, and banded-elevation satellite images of areas of North Korea and Germany. These domains were referred to as W, X, Y, and Z, respectively. The main part of the experiment was conducted with stimuli from two domains, Tree and Cat images in the A domain and banded-elevation satellite images of areas of Haiti and California in the B domain. The *endpoints* of each domain were the images illustrated in Figure 2. Although presented as black and white images in Figure 2, the stimuli in all domains were presented as multicolored RGB 24-bit images.

Stimuli that varied systematically between the endpoints of each domain were created with a commercially available morphing software program (Figuracion, 1998). The intermediate stimuli, which were called *variants,* were produced by superimposing the endpoint stimuli of a domain and changing the relative salience of each. Thus the variants were arrayed along a dimension between the endpoint stimuli of a domain. The software assigned values 000 and 500 to the endpoint stimuli on the satellite-based domain and generated 498 variants between these endpoints. The unit values assigned to the variants varied from 001 to 499 units and indicated relative position along the morphed dimension generated by the software. The variants used in various parts of the experiment were the morphed images with unit values of 030, 070, 100, 130, 170, 210, 250, 280, 310, 340, 370, 390, 430, and 470.

The software assigned unit values 00 and 50 to the endpoint stimuli on the TreeCat domain and generated 49 variants between these endpoints. The numerical values assigned to the variants varied from 01 through 49 and indicated relative position along the morphed dimension generated by the software. The variants used in various parts of the experiment were the morphed images with unit values of 03, 06, 09, 12, 15, 18, 21, 25, 28, 31, 34, 37, 40, 43, and 47.

The stimuli at one end of a domain were assigned lower unit values and will be referred to collectively as Class 1 stimuli. The stimuli at the other end of a domain were assigned higher unit values and will be referred to collectively as Class 2 stimuli. Classes 1 and 2 also will be referred to as the Low and High classes respectively. Preliminary training involved the use of stimuli in the W through Z

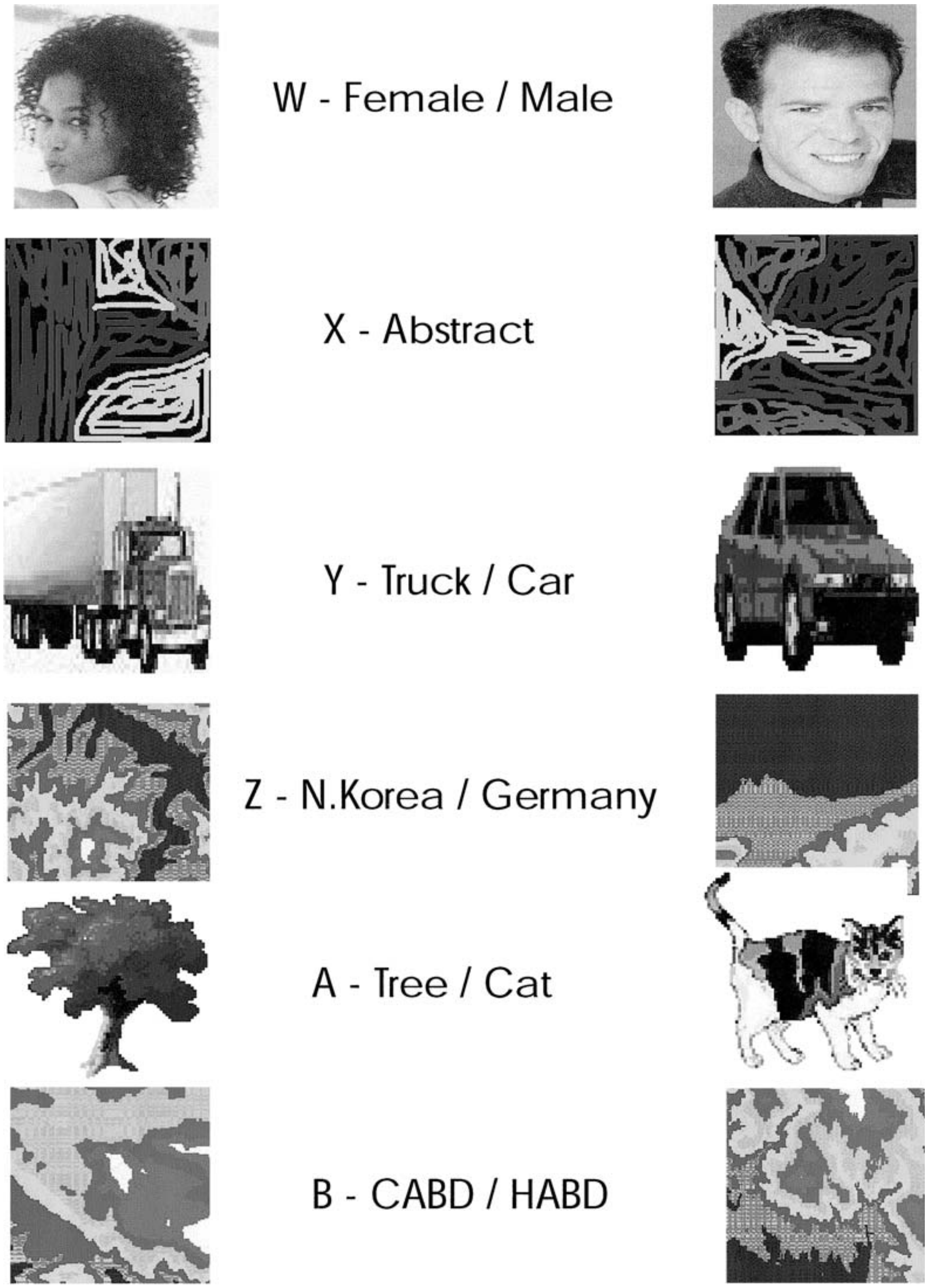

Fig. 2. The stimuli that were the endpoints of all domains used in the experiment. Domains W, X, Y, and Z were used in preliminary training. The endpoints of Domain W were clear images of a male and a female face. The endpoints of Domain X were clear images of two abstract pictures. The endpoints of Domain Y were clear images of a truck and a car. The endpoints of Domain Z were clear images of banded-elevation satellite images of areas of North Korea and Germany. The stimuli in Domains A and B were used in the experiment proper. The endpoints of Domain A were clear images of a tree and a cat. The endpoints of Domain B were clear images of banded-elevation satellite images of areas of Haiti and California.

domains. The anchors, midpoints, and boundaries of two classes in each domain and the ''neither'' stimulus (see below) in each domain were those determined by Fields, Matneja, et al. (2002). First, the endpoints and 14 variants mentioned above were printed on 2.5 by 2.5 cm pieces of paper. The endpoints of a domain were referred to as the anchor stimuli (a) of Classes 1 and 2 and were designated $X1^a$ and $X2^a$, respectively. One anchor stimulus was placed on a table, and a laboratory assistant selected the variant most distant from the anchor in the morphed dimension that was viewed as being related to that anchor stimulus. The same procedure was repeated with the other anchor stimulus. The variants that were selected were referred to as the boundary (b) stimuli of each class and were designated $X1^b$ and $X2^b$ for Classes 1 and 2, respectively. The anchor and boundary stimuli from one class then were placed on the table and the assistant selected the variant that was perceptually equidistant from the anchor and boundary for that class. The same procedure was repeated with the other anchor stimulus. The stimuli that were selected were referred to as the midpoints (m) and were designated as $X1^m$ and $X2^m$ for Classes 1 and 2, respectively. Finally, the boundary stimuli for the two classes were placed on the table and the laboratory assistant selected the variant that was perceptually equidistant from the boundaries of the respective classes. This stimulus was referred to as the *neither* stimulus (n) for the domain and was designated as X^n. It was called the neither stimulus because it was not a member of either Class 1 or Class 2.

The stimuli were sorted in this manner by four laboratory assistants. There was little variation in the values determined by each laboratory assistant for each variant. The specific variants designated as midpoint and boundary stimuli for both classes in the W through Z domains and the neither stimuli in these domains were determined by averaging the values obtained from the sorts conducted by the laboratory assistants. Figure 3 shows the anchor, midpoint, and boundary stimuli for Classes 1 and 2 for Domains W through Z, along with the neither stimulus in each of those domains. Figure 3 also shows some representative stimuli from domains A and B. These stimuli, however, were not necessarily the midpoints, boundaries, and neither stimuli. The stimulus values assigned as the midpoint, boundary, and neither stimuli in these domains were based on the performances of individual subjects as described in Phase 3, below. Note that the terms ''midpoint'' and ''boundary'' are used here in two different ways. In the context of the W, X, Y, and Z domains, the stimulus values denoted by the terms were experimenter-defined, and were the same for all subjects. In the context of the A and B domains, the stimulus values denoted by the terms were determined by the performances of individual subjects.

Procedure

Trial format and responses within a trial. All trials used a matching-to-sample format (Cumming & Berryman, 1965). A trial began when ''Press ENTER'' appeared on the screen. Pressing the enter key cleared the screen and displayed a sample stimulus at the top center of the monitor. Pressing the space bar displayed two comparison stimuli at the bottom left and right corners while the sample remained on the screen. During trials in which the third comparison was programmed, the words ''If NEITHER press 4'' appeared between the two other comparisons.

During a trial, the left or right comparison was selected by pressing the 1 or 2 key, respectively. Pressing the 4 key was the response that selected the neither comparison, when available. A comparison selection cleared the screen and immediately displayed a feedback message centered on the screen. When informative feedback was scheduled, a ''RIGHT'' or ''WRONG'' message appeared, depending on the accuracy of the comparison selection. The message remained on the screen until the R (for RIGHT) or W (for WRONG) key, respectively, was pressed. During some training and all testing trials, uninformative feedback was scheduled following a comparison selection. This consisted of a dashed line surrounding the letter E (- - E - -) that signaled the end of a trial. This cue remained on the screen until the participant pressed the E key, which was used as an observing response for the uninformative feedback. After an appropriate observing response, the screen was cleared and the next trial began (Fields et al., 1995).

Trial block structure and feedback contingencies. Each phase of training and testing

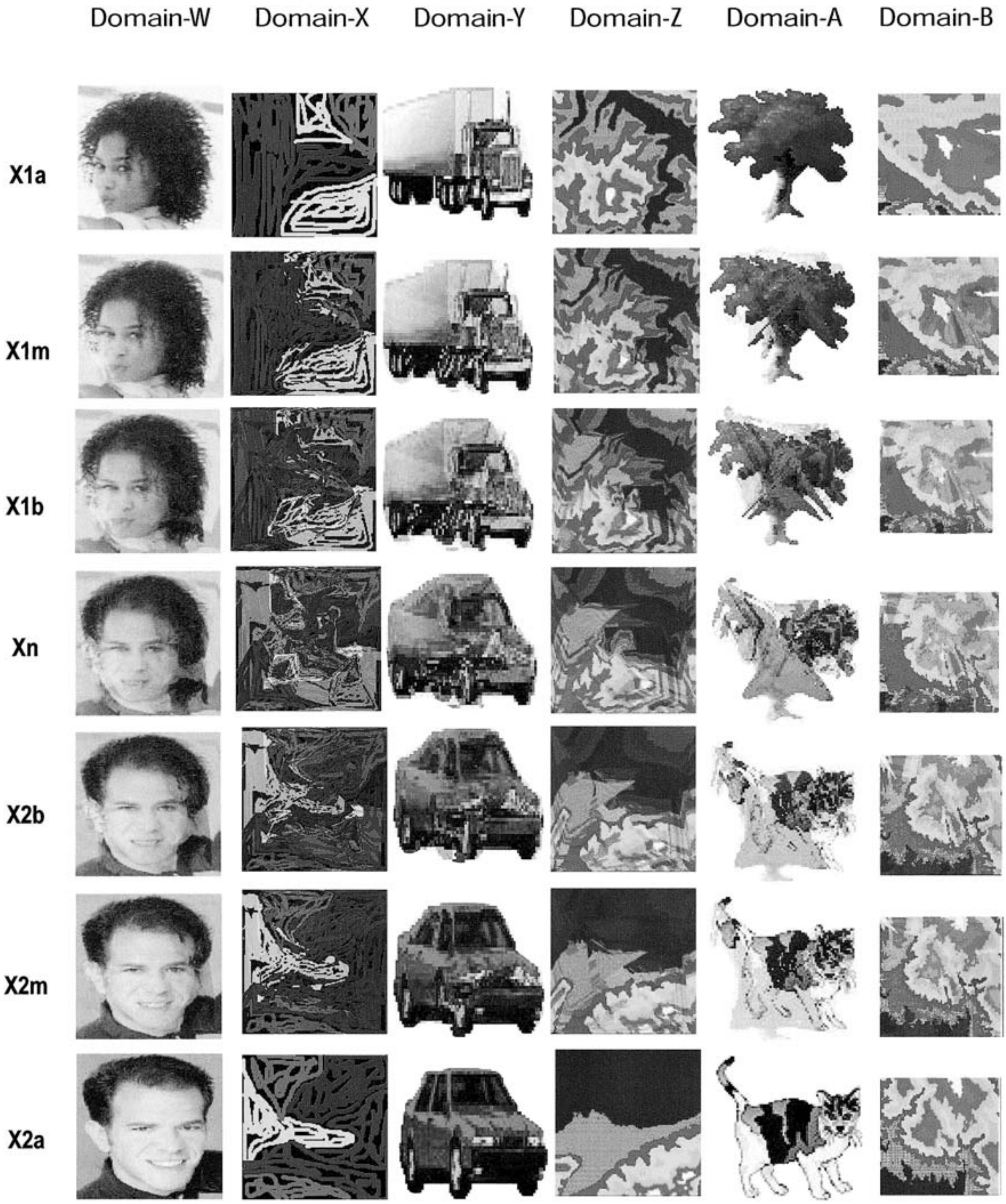

Fig. 3. Anchor, midpoint, and boundary stimuli for the classes at each end of the W through Z domains along with the neither stimulus in the respective domains. The anchor, midpoint, and boundary stimuli in the low classes in the W through Z domains are shown in rows 1 through 3, respectively. The variants assigned as the neither stimuli (X_n) in the domains are shown in the fourth row for the W through Z domains, respectively. The boundary, midpoint, and anchor stimuli in the high classes in domains W through Z are shown in rows 5 through 7, respectively. Variants in the A and B domains are illustrated in the last two columns. The anchor stimuli for the low and high classes in these two domains are illustrated in rows 1 and 7. The stimuli illustrated in rows 2 through 6 are variants that fall between the anchor stimuli but are not necessarily the midpoints, boundaries, or neither stimuli for the A and B domains. The variants that served those functions were determined by a participant's performance in the three-choice generalization tests.

was conducted with blocks of trials. In all experimental phases, the trials in a block were presented in a randomized order without replacement. At the start of training, a block was presented repeatedly with informative feedback after each comparison selection until all trials within the block occasioned 100% correct responding. Thereafter, the percentage of trials that occasioned informative feedback in a block was reduced to 75%, 25%, and finally to 0% as long as comparison selections on all trials were accurate. During feedback reduction, the trials that were followed by informative feedback were randomly determined. If 100% correct responding was not achieved within three blocks at a given feedback level during training, the participant was returned to the previous feedback level for that particular block. Each block ended with the presentation of a message that said, ''Press enter to begin the next block.''

Phase 1: Instructions and keyboard familiarization. Prior to the experiment, subjects were presented with the following instructions on the screen:

> Thank you for volunteering to participate in this experiment. PLEASE DO NOT TOUCH ANY OF THE KEYS ON THE KEYBOARD YET! In this experiment you will be presented with many trials. Each trial contains three or four CUES. These will be familiar and unfamiliar picture images. YOUR TASK IS TO DISCOVER HOW TO RESPOND CORRECTLY TO THE CUES. Initially, there will also be INSTRUCTIONS that tell you how to respond to the cues, and LABELS that will help you to identify the cues on the screen. The labels and the instructions that tell you which KEYS to press will slowly disappear. Your task will be to RESPOND CORRECTLY to the CUES and the INSTRUCTIONS by pressing certain keys on the computer's keyboard. The experiment is conducted in phases. When each phase ends, the screen will sometimes tell you how you did. If you want to take a break at any time, please call the experimenter.

After pressing the space bar, students were trained to emit the appropriate keyboard responses to complete a trial. Sixteen trials, each containing three English words such as KING, QUEEN, and CAMEL, were presented. The semantic relatedness between the sample word (e.g., KING) and one of the comparisons (e.g., QUEEN) was used to prompt the selection of the correct comparison. The words RIGHT or WRONG followed each comparison selection (see Fields, Reeve, Adams, et al., 1997, for further details).

Correct responding to the stimuli in a trial during Phase 1 also was facilitated by instructional prompts (e.g., ''Make your choice by pressing 1 or 2'') that were deleted in a serial manner across trials (see Fields, Reeve, Rosen, et al.,1997, or Fields, Adams, Verhave, & Newman, 1990, for further details). Phase 1 ended once the stimuli were presented without prompts and performance exceeded 87.5% accuracy (14 of 16 correct trials) during a single block. In the remaining phases, whenever a participant pressed a nonexperimentally defined key during a trial, the instruction that prompted the appropriate key press during keyboard familiarization (Phase 1) reappeared on the screen for three subsequent trials.

Phase 2: Generalized categorization repertoire: WXYZ(amb-a) training. Fields, Reeve, et al. (2002) found that multiple-exemplar training with stimuli in a number of different domains established a generalized categorization repertoire. Thereafter, stimuli in new domains were spontaneously categorized into two perceptual classes. In addition, subjects' responses were reinforced for the selection of a neither option, also known as a ''default option'' (Innis, Lane, Miller, & Critchfield, 1998) in the presence of stimuli that were between the boundaries of the two putative classes on a domain. Because the use of the neither option ensured that stimuli in one class did not occasion the selection of stimuli in the other class, the two classes were functionally separable and independent of each other (Reeve & Fields, 2001).

In the present experiment, multiple-exemplar training was conducted with the stimuli in the W through Z domains to induce a generalized categorization repertoire. In each domain, the anchor, midpoint, and boundary stimuli from the two experimenter-defined classes were presented as samples. The comparisons consisted of the pair of anchor stimuli from the same domain. Informative feedback (RIGHT or WRONG) was presented for the selection of the comparison from the same class as the sample. All trials also contained a neither option as a third comparison. Informative textual feedback was presented for the selection of the neither comparison when the neither stimulus was presented as the sample. Training was conducted with the stimuli in the W domain first, and then was repeated with the stimuli from the X, Y, and Z domains. This procedure made it likely that the stimuli in the A and B domains would be categorized without direct training.

Phase 3: Perceptual class-width identification. Perceptual classes that emerged from domains A and B were used in the present experiment. Fields, Matneja, Varelas, and Belanich (2003) showed that the width of the same nominal perceptual class can vary depending on whether generalization tests are conducted in *variant-to-base* or *base-to-variant* formats. In the variant-to-base format, a subject is presented with variants as samples and the endpoints of a domain as comparisons. Under these conditions, a subject selects an endpoint stimulus that is from the same ''class'' as the variant. In the base-to-variant format, a subject is presented with the endpoints of a domain as samples with the variants in that domain as comparisons. Under these conditions, a subject selects the variants that are from the same class as the sample stimulus.

Fields and Reeve (2001) showed that two classes along a continuum could be functionally separated from each other if subjects had access to, and used, a neither comparison during generalization test trials. In the current experiment, the widths of the classes at the ends of the A and B domains, A1′, A2′, B1′, and B2′, were determined with generalization tests conducted in the variant-to-base (VB) and base-to-variant (BV) formats. In addition, a neither comparison was included in all generalization tests conducted in both formats to induce separable classes in each domain.

During the VB tests, each of the variants on a domain (e.g., TreeCat-00 through TreeCat-50) was presented as sample stimuli on different trials. The endpoint stimuli from that domain (e.g., TreeCat-00 and TreeCat-50), and the neither option were presented as comparisons on all trials. During the BV tests, the endpoint stimuli (i.e., TreeCat-00 or TreeCat-50) were presented as a sample on different trials. For each endpoint sample, the other endpoint stimulus and the neither option were presented as two of the three comparisons on all trials. Finally, different variants were presented as the third comparison across trials. Contiguous variants were considered part of a putative class if each of them occasioned the selection of a given endpoint stimulus on at least 88% of the generalization test trials. The boundary stimulus for that class was the contiguous variant most removed from the anchor that occasioned the selection of the anchor on at least 88% of the trials. The midpoint stimulus for a class was the variant that was equidistant between the anchor and the boundary stimuli for a class.

The variant-to-base and base-to-variant tests were conducted in separate blocks of trials, each of which included two presentations of all variants. Each block was presented four times in each test format for a total of eight presentations of each variant in each test format. Subjects were presented first with the eight test blocks that contained stimuli in the A domain, and then with eight blocks that contained stimuli from the B domain. For stimuli in a given domain, the variant-to-base and base-to-variant test blocks were presented in simple alternation.

Phase 4: Linkage of classes with cross-class conditional discriminations. Cross-class conditional discriminations were established between the anchor stimuli of A1ˋ and B1ˋ classes and the boundary stimuli of the same two classes. In addition, cross-class conditional discriminations were established between the anchor stimuli of the A2ˋ and B2ˋ classes and the boundary stimuli of the same two classes. On some trials, the anchor stimuli, $A1^a$ or $A2^a$, were presented as samples with the anchor stimuli from the Bˋ classes, $B1^a$ and $B2^a$, as comparisons. Informative feedback was presented for the selection of the comparison with the same class number designation as the sample. On the other trials, the boundary stimuli, $A1^b$ or $A2^b$, were presented as samples with the boundary stimuli from the Bˋ classes, $B1^b$ and $B2^b$, as comparisons. Informative textual feedback was presented for the selection of the comparison with the same class number designation as the sample. Training was conducted in a block that contained 16 trials, and was completed once all of those trials occasioned correct comparison selection in a block that contained no informative feedback.

At the completion of training, the symmetrical properties of the A→B conditional discriminations were assessed with blocks that

Table 1

Symbolic representation of stimuli used in the cross-class probes. Each line indicates the stimuli used in two cross-class probes. Both probes share the same set of comparison stimuli, but the positive comparison is different for each class. Superscripts a, m, and b designate anchor, midpoint, and boundary functions for each stimulus, respectively. Sa designates sample stimuli, Co+ designates positive comparisons, Co− designates negative comparisons, and NC designates the neither comparison.

	Class 1 probes				Class 2 probes			
Test format	Sa	Co+	Co−	Co−	Sa	Co+	Co−	Co−
Aa→Ba	$A1^a$	$B1^a$	$B2^a$	NC	$A2^a$	$B2^a$	$B1^a$	NC
Am→Ba	$A1^m$	$B1^a$	$B2^a$	NC	$A2^m$	$B2^a$	$B1^a$	NC
Ab→Ba	$A1^b$	$B1^a$	$B2^a$	NC	$A2^b$	$B2^a$	$B1^a$	NC
Aa→Bm	$A1^a$	$B1^m$	$B2^m$	NC	$A2^a$	$B2^m$	$B1^m$	NC
Am→Bm	$A1^m$	$B1^m$	$B2^m$	NC	$A2^m$	$B2^m$	$B1^m$	NC
Ab→Bm	$A1^b$	$B1^m$	$B2^m$	NC	$A2^b$	$B2^m$	$B1^m$	NC
Aa→Bb	$A1^a$	$B1^b$	$B2^b$	NC	$A2^a$	$B2^b$	$B1^b$	NC
Am→Bb	$A1^m$	$B1^b$	$B2^b$	NC	$A2^m$	$B2^b$	$B1^b$	NC
Ab→Bb	$A1^b$	$B1^b$	$B2^b$	NC	$A2^b$	$B2^b$	$B1^b$	NC
Ba→Aa	$B1^a$	$A1^a$	$A2^a$	NC	$B2^a$	$A2^a$	$A1^a$	NC
Bm→Aa	$B1^m$	$A1^a$	$A2^a$	NC	$B2^m$	$A2^a$	$A1^a$	NC
Bb→Aa	$B1^b$	$A1^a$	$A2^a$	NC	$B2^b$	$A2^a$	$A1^a$	NC
Ba→Am	$B1^a$	$A1^m$	$A2^m$	NC	$B2^a$	$A2^m$	$A1^m$	NC
Bm→Am	$B1^m$	$A1^m$	$A2^m$	NC	$B2^m$	$A2^m$	$A1^m$	NC
Bb→Am	$B1^b$	$A1^m$	$A2^m$	NC	$B2^b$	$A2^m$	$A1^m$	NC
Ba→Ab	$B1^a$	$A1^b$	$A2^b$	NC	$B2^a$	$A2^b$	$A1^b$	NC
Bm→Ab	$B1^m$	$A1^b$	$A2^b$	NC	$B2^m$	$A2^b$	$A1^b$	NC
Bb→Ab	$B1^b$	$A1^b$	$A2^b$	NC	$B2^b$	$A2^b$	$A1^b$	NC

contained 12 baseline review trials ($A^a \rightarrow B^a$ and $A^b \rightarrow B^b$) and 12 symmetry probe trials for each trained conditional discrimination ($B^a \rightarrow A^a$ and $B^b \rightarrow A^b$). The criterion for demonstrating the emergence of symmetry involved the selection of the set-consistent comparisons on at least 94% of the trials in a test block.

Phase 5: Test schedules to evaluate linked perceptual class formation. The emergence of relations between members of the A` and B` classes was assessed with the performances occasioned by nine cross-class probes conducted in an A`-B` format and nine other symmetrical cross-class probes conducted in a B`-A` format. As seen in Table 1, all of the A`-B` probes included samples that were the anchor, midpoint, and boundary variants from the Tree and Cat classes (the A classes) and the comparisons that were pairs of the anchor, midpoint, or boundary stimuli from the two satellite-based classes (the B classes). All of the B`-A` probes included samples that were the anchor, midpoint, and boundary variants from the two satellite-based classes and comparisons that were pairs of the anchor, midpoint, or boundary stimuli from the TreeCat classes. All cross-class probe trials also included a neither option as a third comparison, which enabled a participant to indicate that the sample was not related to the other two comparison stimuli.

Although the values assigned to the anchor stimuli were constant in all A`-B` and B`-A` probes, this was not necessarily the case with the midpoint and boundary stimuli. The actual values of the midpoint and boundary stimuli for the same class that were used in the cross-class probes were those obtained from individual subjects in the variant-to-base and base-to-variant tests. When midpoint or boundary stimuli were presented as samples, their values were those obtained from the variant-to-base tests. When midpoint or boundary stimuli were presented as comparisons, their values were those obtained from the base-to-variant tests.

The 18 cross-class probes used to measure the emergence of linked perceptual classes were presented in one of four different test

Table 2

Testing schedules. Each column designates a separate testing schedule. The column heading indicates the number of blocks followed by the number of different probe types per block. Dotted lines separate the blocks in the test. Blocks were presented in the order indicated in a column. Trials within a block were presented in randomized order.

2/9		6/3		18/1-RND	18/1-PRGM
1	Aa–Ba	1	Aa–Ba	Ba–Aa	Aa–Ba
				………	………
	Am–Ba		Am–Ba	Aa–Bm	Am–Ba
				………	………
	Ab–Ba		Ab–Ba	Bm–Ab	Ab–Ba
		………	………	………	………
	Aa–Bm	2	Aa–Bm	Am–Ba	Aa–Bm
				………	………
	Am–Bm		Am–Bm	Bm–Am	Am–Bm
				………	………
	Ab–Bm		Ab–Bm	Am–Bb	Ab–Bm
		………	………	………	………
	Aa–Bb	3	Aa–Bb	Bb–Ab	Aa–Bb
				………	………
	Am–Bb		Am–Bb	Ba–Ab	Am–Bb
				………	………
	Ab–Bb		Ab–Bb	Ab–Bb	Ab–Bb
………	………	………	………	………	………
2	Ba–Aa	4	Ba–Aa	Ba–Am	Ba–Aa
				………	………
	Bm–Aa		Bm–Aa	Ab–Bm	Bm–Aa
				………	………
	Bb–Aa		Bb–Aa	Aa–Ba	Bb–Aa
		………	………	………	………
	Ba–Am	5	Ba–Am	Bm–Aa	Ba–Am
				………	………
	Bm–Am		Bm–Am	Ab–Ba	Bm–Am
				………	………
	Bb–Am		Bb–Am	Am–Bm	Bb–Am
		………	………	………	………
	Ba–Ab	6	Ba–Ab	Bb–Am	Ba–Ab
				………	………
	Bm–Ab		Bm–Ab	Bb–Aa	Bm–Ab
				………	………
	Bb–Ab		Bb–Ab	Aa–Bb	Bb–Ab
………	………	………	………	………	………

schedules. Each schedule is distinguished in terms of the number of different probes that were included in a block of test trials, and the order in which the probes were introduced across test blocks. Table 2 lists the number of blocks presented in each testing schedule and the particular cross-class probes included in each test block. All of the trials in a block were presented with no differential feedback and in a randomized order without replacement.

Each cross-class probe for a given linked perceptual class was presented eight times in a test block. Each presentation of a probe is called a trial. Each testing schedule involved the presentation of a total of 144 trials. In the 2/9 test, each of the two blocks contained 72 trials for a total of 144 trials. In the 6/3 test, each of the six blocks contained 24 trials for a total of 144 trials. In both 18/1 tests, each of the 18 blocks contained eight trials for a total of 144 trials.

The 2/9 test involved the presentation of two test blocks with nine different cross-class probes in each block. The first test block contained the nine probes in the A`-B` format. This was followed by the second test block that contained the nine probes presented in the B`-A` format. Each test block involved the randomized presentation of trials corresponding to nine different types of cross-class probes.

The 6/3 test involved the presentation of six test blocks, each of which contained three different cross-class probes. The first three test blocks contained A`-B` probes, and the last three contained B`-A` probes. All three of the A`-B` test blocks included the presentation of the anchor, midpoint, and boundary stimuli from the A domain as samples. In the first block, the anchor stimuli from the B domain were presented as comparisons. In the second block, the midpoint stimuli from the B domain were presented as comparisons. In the third block, the boundary stimuli from the B domain were presented as comparisons. The subsequently presented three B`-A` blocks had the same organization as the A`-B` block with one exception; the B` stimuli served as samples and the A` stimuli served as comparisons. Trials corresponding to the three different types of cross-class probes in each test block were presented in a randomized order.

The 18/1-RND test involved the presentation of each of the 18 cross-class probes in separate sequentially presented blocks. The order of presenting the probe types was randomized across test blocks in terms of (a) the A` or B` class from which the samples were drawn, (b) the value of the sample stimuli, (c) the A` or B` class from which the comparisons were drawn, and (d) the value of the comparison stimuli.

The 18/1-PRGM test involved the presentation of one probe type per test block where the probes were introduced in a highly programmed and systematic order. Specifically, the first nine test blocks were in A′-B′ format. The anchors, midpoints, and boundaries of the A1` and A2` classes were the sample stimuli in the first, second, and third test blocks, respectively. All three of these blocks contained the anchor stimuli from the B1` and B2` classes as the comparisons. Then the anchors, midpoints, and boundaries of the A1` and A2` classes were the sample stimuli in the fourth, fifth, and sixth test blocks, respectively. All three of these blocks contained the midpoint stimuli from the B1` and B2` classes as comparisons. Finally, the anchors, midpoints, and boundaries of the A1` and A2` classes were the sample stimuli in the seventh, eighth, and ninth test blocks, respectively. All three of these blocks contained the boundary stimuli from the B1` and B2` classes as comparisons. This entire sequence was repeated in test blocks 10 to 18 with the A` and B` stimuli reversed as samples and comparisons. Although the order of introducing the probes across blocks was highly systematized, trials were randomized within a block.

RESULTS

Generalization and perceptual class emergence. To illustrate the structure of the data set, Figure 4 (left column) shows the results of the variant-to-base tests conducted with the Tree–Cat-based variants for a representative participant, Subject 2371. The variants that ranged in value from TreeCat-50 to TreeCat-37 occasioned the selection of TreeCat-50 on at least 88% of trials. Therefore, they functioned as members of the B2` class with TreeCat-37 as the boundary stimulus for that class. In like manner, the variants that ranged in value from TreeCat-00 to TreeCat-18 occasioned the selection of TreeCat-00 on at least 88% of trials. Therefore, they functioned as members of the B1` class with TreeCat-18 as the boundary stimulus for that class.

The selection of TreeCat-50 declined systematically as the variants moved increasingly beyond the boundary stimuli of the B1` class. This decrement was accompanied by a complementary increase in the selection of the neither comparison. It was not accompanied, however, by the selection of TreeCat-00, the other TreeCat comparison. Likewise, the selection of TreeCat-00 declined systematically as the variants moved increasingly beyond the boundary stimuli in the B2` class. This decrement in responding was accompanied by a complementary increase in the selection of the neither comparison, but not by the selection of TreeCat-50, the other TreeCat comparison. Thus the widths of the B1` and B2` classes were functionally separate and independent of each other. As such, one class was not defined as the complement of the other.

When the base-to-variant test involved the presentation of the TreeCat-00 stimulus as the sample (panels in the middle column), the variants that ranged from TreeCat-00 to TreeCat-25 were selected on at least 88% of trials in the presence of TreeCat-00 and, therefore, were functioning as members of the B1′ class with TreeCat-25 as the boundary value for the low-TreeCat class. The likelihood of selecting the neither comparison increased

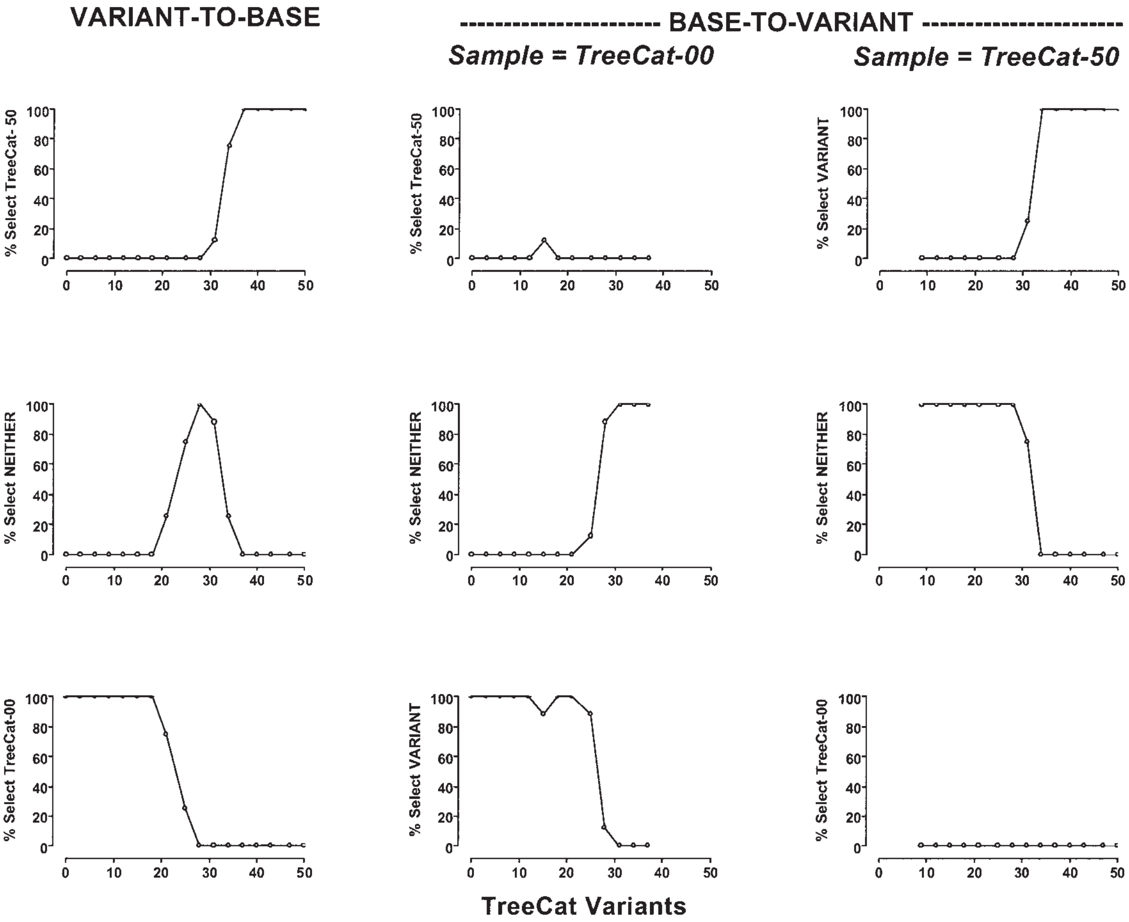

Fig. 4. The results of the variant-to-base and base-to-variant tests during Phase 3 for Participant 2371. Left column: Results of the variant-to-base tests, expressed as the likelihood of selecting TreeCat-00, the neither comparison, and TreeCat-50 as functions of the value of the TreeCat variants presented as samples. The results of the base-to-variant tests are presented in the two remaining columns. Middle column: Likelihoods of selecting the TreeCat variants, the neither comparison, and negative comparison in the presence of TreeCat-00 as functions of the values of the TreeCat variants presented as comparisons. Right column: Likelihoods of selecting the TreeCat variants, the neither comparison, and the negative comparison in the presence of TreeCat-50 as functions of the values of the TreeCat variants presented as comparisons.

in a complementary manner as the value of the variant moved increasingly beyond the value of the boundary stimulus and reached asymptote at TreeCat-28. In addition, the negative comparison (Co−) was rarely selected, regardless of the comparison variant.

When the base-to-variant test involved the presentation of the TreeCat-50 stimulus as the sample (panels in the right column), the variants that ranged from TreeCat-50 to TreeCat-34 were selected on at least 88% of trials in the presence of TreeCat-50. Therefore, they were functioning as members of the B2′ class with TreeCat-34 as the $B2^b$, the boundary value for the high-TreeCat class. The likelihood of selecting the neither comparison increased in a complementary manner with declines in the selection of the other two comparisons, and reached asymptote at TreeCat-31. Finally, the Co− was never selected on any trial regardless of the comparison stimulus value. Thus the variants that were in each class in the same domain did not overlap, were not complements of each other, and were functionally independent and separable from each other.

Table 3

Boundaries of perceptual classes A1`, A2`, B1`, and B2`, measured with variant-to-base (VB) and base-to-variant (BV) tests for subjects receiving different cross-class test schedules. See text for explanation of testing-schedule labels.

		VB		BV		VB		BV	
Schedule	Subject	A1`	A2`	A1`	A2`	B1`	B2`	B1`	B2`
18/1-PRGM	2435	12	43	09	43	070	340	070	430
	2436	12	40	15	40	170	340	170	340
	2398	21	34	21	34	210	280	130	280
	2414	09	43	09	43	070	430	100	430
	2400	18	37	15	40	210	340	130	340
	2432	09	43	09	43	130	250	170	250
	2409	18	34	18	34	170	280	170	280
	2418	15	43	18	43	130	310	130	340
18/1-RND	2471	21	34	18	34	170	340	170	340
	2472	15	43	15	43	170	340	170	340
	2475	12	43	15	43	170	250	100	250
	2544	18	43	15	43	130	340	170	340
	2545	21	40	25	37	130	340	130	370
	2554	18	37	18	37	170	340	170	340
	2548	15	37	15	34	170	280	210	250
	2477	09	43	12	43	070	430	070	390
	2486	21	34	28	31	170	340	210	340
	2547	18	40	15	40	130	340	130	340
	2549	18	34	15	37	130	250	130	340
	2481	18	34	21	34	100	280	130	280
6/3	2366	18	40	18	37	210	340	210	340
	2364	18	34	18	34	170	310	170	340
	2371	18	37	25	34	170	280	100	210
	2369	12	40	18	37	130	280	210	280
	2362	21	37	18	37	170	280	170	310
	2411	15	40	12	37	170	340	070	390
	2425	21	40	18	43	130	430	100	310
2/9	2431	18	34	18	37	130	310	170	310
	2399	06	43	12	43	170	310	170	340
	2410	18	34	18	34	170	340	130	340
	2428	12	43	12	43	130	340	100	390
	2413	18	40	21	40	170	340	170	340
	2427	18	43	15	37	170	390	170	310
	2444	15	43	09	43	100	340	130	310
	MEAN	16	39	16	39	149	326	145	327
	min-max[a]	00	50	00	50	000	500	000	500
	Mean Width[b]	16	11	16	11	149	174	145	173

[a] Minimum and maximum values assigned to the respective endpoint stimuli on each domain.

[b] Average width of stimulus classes at each end of a domain.

Boundary stimuli of perceptual classes. Table 3 shows the boundary stimuli for each class for each subject obtained using the VB and BV tests. For the A domain, which had endpoint values of 0 and 50, the boundary stimuli of the A1` and A2` classes averaged 16 and 39 units respectively, and were separated by an average of 23 units. For the B domain, which had endpoint values of 0 and 500, the boundary stimuli of the B1` and B2` classes averaged 147 and 327 units respectively, and were separated by an average of 180 units. Because the boundaries of the two classes are separated by variants that are not members of either class, the classes on each domain are functionally independent of each other.

Response speed in perceptual classes. One defining characteristic of a perceptual class is that some of the class members must be discriminable from each other. That defining property of the perceptual classes studied in the present experiment was evaluated by measuring the response speeds occasioned by the anchor, midpoint, and boundary stimuli

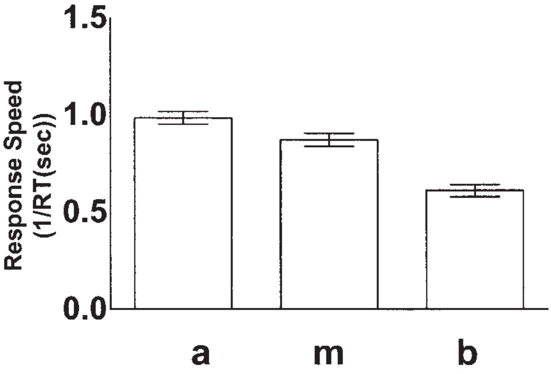

Fig. 5. Mean response speeds occasioned by the anchor (a), midpoint (m), and boundary (b) stimuli in the perceptual classes. See text for explanation of data aggregation. Error bars show ± 1 *SE*.

in the putative A1`, A2`, B1`, and B2` classes (for justification, see Fields, Reeve, et al., 2002). Response speed was measured as the reciprocal of the time (in seconds) that separated the sample-observing response from the selection of a comparison. Averages were computed across participants, domains, classes in a domain, and type of generalization test because systematic differences were not correlated with any of these factors. Response speeds were averaged separately for the anchor, midpoint, and boundary stimuli. An analysis of variance showed significant differences in the response speeds occasioned by the anchor, midpoint, and boundary stimuli in the perceptual classes, $F = (2, 213) = 42.41$, $p < .0001$.

Figure 5 shows that the average response speed was fastest for the anchor stimuli in the emergent TreeCat- and satellite-based classes, was slower for the midpoint stimuli, and was slowest for the boundary stimuli. Newman-Keuls post hoc tests of pair-wise comparisons showed significant differences in the response speeds occasioned by the anchor and midpoint ($q = 4.39$, $p < .01$), midpoint and boundary ($q = 8.42$, $p < .001$), and anchor and boundary stimuli ($q = 12.82$, $p < .001$). Thus the anchor, midpoint, and boundary stimuli in a class were all discriminable from each other. The response speed data, then, suggest that the stimuli in the classes satisfied one of the defining properties of stimuli in a perceptual class, that of discriminability.

Acquisition of cross-class conditional relations. Table 4 shows that most participants learned the $A^a \rightarrow B^a$ and $A^b \rightarrow B^b$ conditional discriminations in a few trial blocks and maintained mastery level performances during the reduction of feedback. In addition, very few test blocks were needed to pass the tests that evaluated the symmetrical properties of the stimuli in the conditional discriminations. Thus conditional relations had been established between the anchor stimuli in each pair of A` and B` perceptual classes and between the boundary stimuli in the same pairs of classes. Because there were no systematic differences in these data, any differential effects of test schedules on linked perceptual class formation could not be attributed to the establishment of the cross-class conditional relations.

Cross-class test performances. The formation of a linked perceptual class was assessed by the performance occasioned by the eighteen cross-class probes. The performance occasioned by these probes is presented in the vertical pairs of ''dot plots'' in Figures 6 through 9. In each pair, the upper- and lower-dot plots depict the data obtained in A`-B` and B`-A` tests, respectively. The *x* axis of each dot plot depicts the anchor, midpoint, and boundary stimuli used as comparisons. The *y* axis of each dot plot depicts anchor, midpoint, and boundary stimuli used as samples. The positions of the stimuli on each axis are scaled in an ordinal manner, with the anchors at the origins of each axis. The actual values of each stimulus are listed at the positions represented by the anchor, midpoint, and boundary for each class on each axis. The stimulus domain indicated on each axis differs with testing format. In the A`-B` tests, the TreeCat and Satellite stimuli were displayed on the ordinate and abscissa, respectively. In the B`-A` tests, the Satellite and TreeCat stimuli were displayed on the ordinate and abscissa, respectively.

Each data point in a dot plot represents the performance of a given cross-class probe. For example, the middle dot on the uppermost row of the dot plot represents the performance occasioned by the Ab-Bm probe. The actual performance occasioned by each probe is indicated by the darkness of the dot at each data point. A black dot represents 88% to 100% class-consistent responding, a white dot represents 0% to 12% class-consistent responding, and dots in one of the three variations of gray represent other ranges of responding that

Table 4

Acquisition: Number of trial blocks needed to acquire the $A^a \rightarrow B^a$ and $A^b \rightarrow B^b$ conditional discriminations and pass the $B^a \rightarrow A^a$ and $B^b \rightarrow A^b$ symmetry tests for each participant. The training columns indicate the number of blocks needed to reach mastery under four levels of feedback (%FB). The value in the column labeled #Blks indicates the number of blocks needed to pass the symmetry test. The values in the columns labeled Blks 1 and 2 indicate the percentage of trials that occasioned correct comparison selection in each symmetry test block.

		Number of training blocks				Symmetry tests		
Test schedule	Subject	100%	75%	25%	0%	#Blks	Blk1	Blk2
18/1-PRGM	2435	3	1	1	1	1	100	
	2436	32	1	1	1	1	100	
	2398	2	1	2	1	1	100	
	2414	3	1	1	2	1	97	
	2400	6	1	1	1	1	100	
	2432	7	1	1	1	2	92	94
	2409	1	1	1	1	1	100	
	2418	2	1	1	1	1	100	
18/1-RND	2471	1	1	2	1	1	100	
	2472	2	1	1	1	1	98	
	2475	10	2	1	1	1	100	
	2544	5	1	1	1	1	98	
	2545	2	2	1	1	1	100	
	2554	4	1	1	1	1	100	
	2548	3	1	1	1	1	100	
	2486	5	1	1	1	1	100	
	2547	4	1	1	1	1	100	
	2549	17	1	1	1	1	98	
	2481	9	1	1	1	1	98	
6/3	2366	3	1	1	1	1	100	
	2364	13	1	1	1	2	92	94
	2371	4	2	1	1	2	92	100
	2369	5	1	2	2	1	100	
	2362	7	1	2	1	1	97	
	2411	5	1	1	1	1	95	
	2425	5	2	1	1	1	100	
2/9	2431	5	1	1	1	1	95	
	2399	8	1	1	1	1	98	
	2410	4	1	1	1	1	100	
	2428	3	1	1	1	1	95	
	2413	6	1	1	1	1	100	
	2427	2	1	1	1	1	97	
	2444	3	1	1	1	1	98	
ALL	MEAN	5.8	1.1	1.1	1.1	1.1	98.2	

are greater than 12% and less than 88% class-consistent comparison selection. The continuous line that connects the data points demarcates all of the stimuli from one perceptual class that theoretically occasion the selection of all of the stimuli from the other perceptual class on at least 88% of the cross-class test trials. A linked perceptual class was deemed to have been formed when a subject responded with at least 88% accuracy on at least 17 of the 18 cross-class probes presented in the pair of A`-B` and B`-A` dot plots.

Figures 6 through 9 show the outcomes of the tests presented to each subject exposed to the 2/9, 6/3, 18/1-RND, and 18/1-PRGM test schedules, respectively. To illustrate, in the 2/9 condition (Figure 6), Subject 2410 demonstrated class-consistent performance on all 18 cross-class probes for both Class 1 and Class 2. Therefore, this subject formed two linked perceptual classes. In the 6/3 condition (Figure 7), Subject 2371 demonstrated class-consistent performance on 17 out of the 18 cross-class probes for both Class 1 and Class 2.

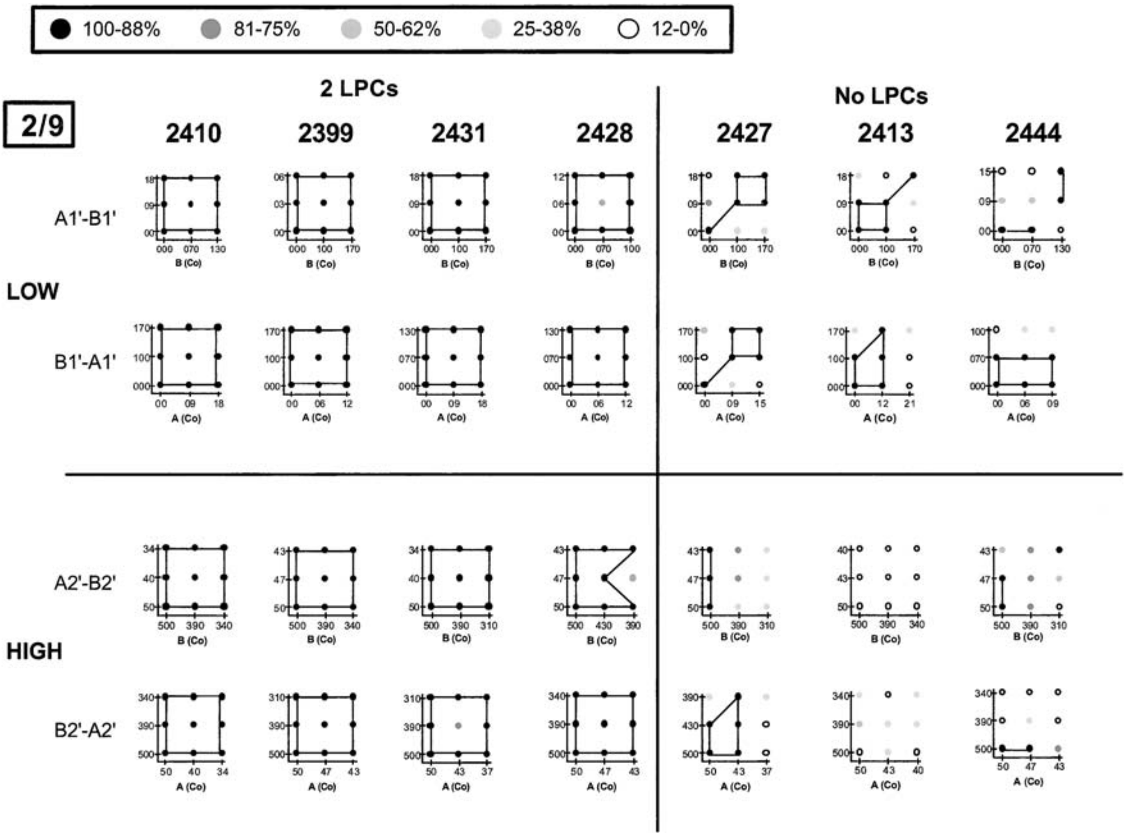

Fig. 6. Performance occasioned by all cross-class probes presented in the A`-B` and B`-A` formats when using the 2/9 test schedule. See text for explanation of graph format. Each column contains data for 1 subject. Top two rows: Outcomes of the A`-B` and B`-A` tests for the low A/B classes. Bottom two rows: Outcomes of the A`-B` and B`-A` tests for the high A/B classes. The vertical line separates subjects who formed two linked perceptual classes (to the left of the line) from those who did not form any linked perceptual classes (to the right of the line).

This subject's performance also illustrated the formation of two linked perceptual classes. In the 18/1-RND condition (Figure 8), Subject 2549 responded in a class-consistent manner to only 10 out of the 18 cross-class probes from the putative Class 1, and 7 out of the 18 cross-class probes from the putative Class 2. Therefore, neither of the linked perceptual classes emerged during the test. In the 18/1-PRGM condition (Figure 9), the test performance of Subject 2418 showed that all of the A2`-B2` and B2`-A2` probes occasioned class-consistent responding, whereas only 14 of the 18 A1`-B1` and B1`-A1` probes occasioned class-consistent responding. These results indicated the emergence of only one of two linked perceptual classes.

In a number of tests, a particular cross-class probe occasioned the selection of the class-consistent comparison on no more than 12% of the test trials. This could indicate a very high likelihood of selecting the comparison from the other potentially linked perceptual class or of selecting the neither comparison. When the class-consistent comparisons were selected on no more than 12% of the test trials, the subjects chose the neither comparison in 82%, 90%, 100%, and 100% of trials during the 2/9, 6/3, 18/1-RND, and 18/1-PRGM tests, respectively. Therefore, comparison selections that were not class-indicative involved the selection of the neither option rather than a comparison from the opposing class.

Effect of test schedule on the emergence of linked perceptual classes. The two panels in Figure 10 summarize the effects of the four different test schedules on the emergence of linked perceptual classes. Two possible linked perceptual

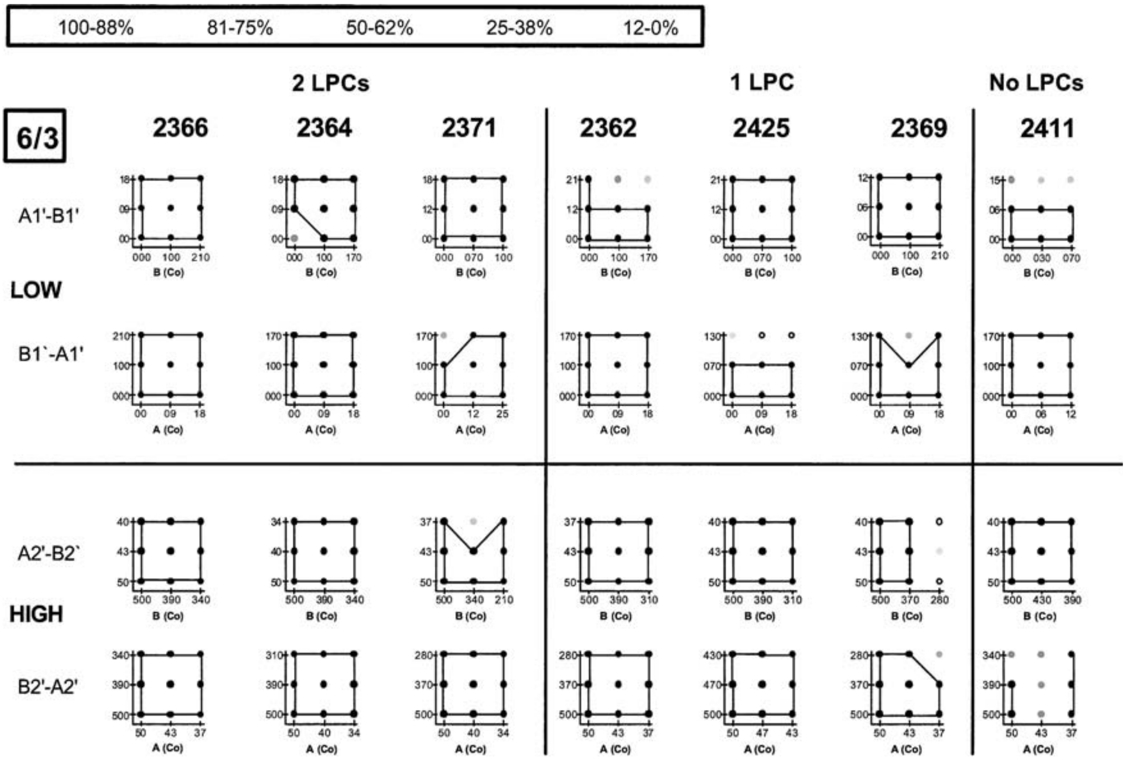

Fig. 7. Performance occasioned by all cross-class probes presented in the A`-B` and B`-A` formats when using the 6/3 test schedule. See text for explanation of graph format. Each column contains data for 1 subject. Top two rows: Outcomes of the A`-B` and B`-A` tests for the low A/B classes. Bottom two rows: Outcomes of the A`-B` and B`-A` tests for the high A/B classes. The left vertical line separates subjects who formed two linked perceptual classes (to the left of the line) from those who formed only one (to the right of the first line). The second vertical line separates subjects who formed one linked perceptual class from the one who did not form either linked perceptual class (to the right of the second line).

classes could emerge for each subject. Thus a total of 16, 14, 16, and 24 linked perceptual classes could emerge during the 2/9, 6/3, 18/1-RND, and 18/1-PRGM tests, respectively.

The upper panel of Figure 10 presents the percentage of possible linked perceptual classes that emerged in each group. Approximately 55% to 65% of the linked perceptual classes emerged when class formation was assessed with the 2/9, 6/3, and 18/1-RND test schedules. The small differences in yields during 2/9, 6/3, and the 18/1-RND tests were not significantly different according to chi-square comparisons. In contrast, testing with the 18/1-PRGM schedule resulted in the emergence of 98% of linked perceptual classes. The percentage of linked perceptual classes that emerged during 18/1-PRGM differed significantly from that in the 2/9, 6/3, and 18/1RND groups, X^2 (1) = 4.051, $p < .01$ in all cases. Therefore, the ordered introduction of different cross-class probes influenced the emergence of linked perceptual classes.

The measure used in the upper panel was a typical group-based measure of class formation. Those data, however, do not necessarily show how the test schedules influence class formation by individual subjects. The data in the lower panel of Figure 10 show such an effect by plotting the percentage of subjects in each group who formed two linked perceptual classes as a function of testing schedule. When the different types of probes were presented in a randomized order within a test block, between 40% and 60% of the subjects formed both linked perceptual classes. When only one probe per test block was scheduled, a modest percentage of subjects formed both classes when the order of probe presentation was unsystematic, whereas most subjects formed

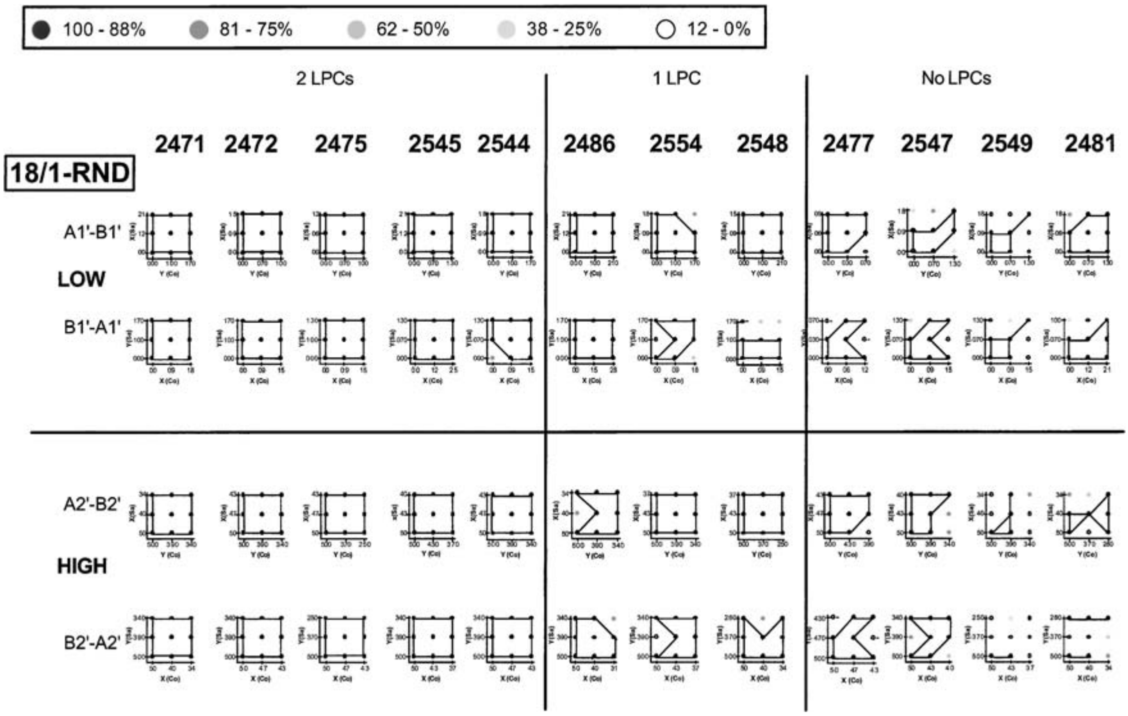

Fig. 8. Performance occasioned by all cross-class probes presented in the A′-B′ and B′-A′ formats when using the 18/1-RND test schedule. See text for explanation of graph format. Each column contains data for 1 subject. Top two rows: Outcomes of the A′-B′ and B′-A′ tests for the low A/B classes. Bottom two rows: Outcomes of the A′-B′ and B′-A′ tests for the high A/B classes. The left vertical line separates subjects who formed two linked perceptual classes (to the left of the line) from those who formed only one (to the right of the first line). The second vertical line separates subjects who formed one linked perceptual class from those who did not form either linked perceptual class (to the right of the second line).

both linked perceptual classes when the probes were presented in the ordered sequence.

To summarize, two dependent measures were used to evaluate the effects of testing schedule on the formation of linked perceptual classes: (a) the percentage of linked perceptual classes that emerged in a group of subjects, and (b) the percentage of subjects in a group who showed the emergence of both linked perceptual classes. With both measures, similar intermediate yields were obtained with the 2/9, 6/3, and 18/1-RND testing schedules, each of which involved the randomized presentation of different cross-class probes across test blocks. In contrast, much higher yields were obtained with the 18/1-PRGM testing schedule where testing involved the programmed presentation of cross-class probes across test blocks.

Failed emergence of linked perceptual classes. Table 5 summarizes the loci of errors that occurred when a linked perceptual class did not emerge. The baseline review probes and their symmetrical counterparts occasioned 57% of the errors that occurred in the 18/1-PRGM group, but only 15%, 21%, and 9% of the errors that occurred in the 2/9, 6/3, and 18/1-RND groups, respectively. When the novel cross-class probes were considered, the 2/9, 6/3, and 18/1-RND tests occasioned twice as many errors as did the 18/1-PRGM test. When the baseline and symmetry trials were compared across testing schedules, trials that contained boundary stimuli as samples and comparisons (bb) produced about 4 times more errors than did the analogous trials that contained anchor stimuli as samples and comparisons (aa).

DISCUSSION

Perceptual classes. Three criteria must be satisfied to conclude that stimuli are functioning as members of a class (Fields & Reeve,

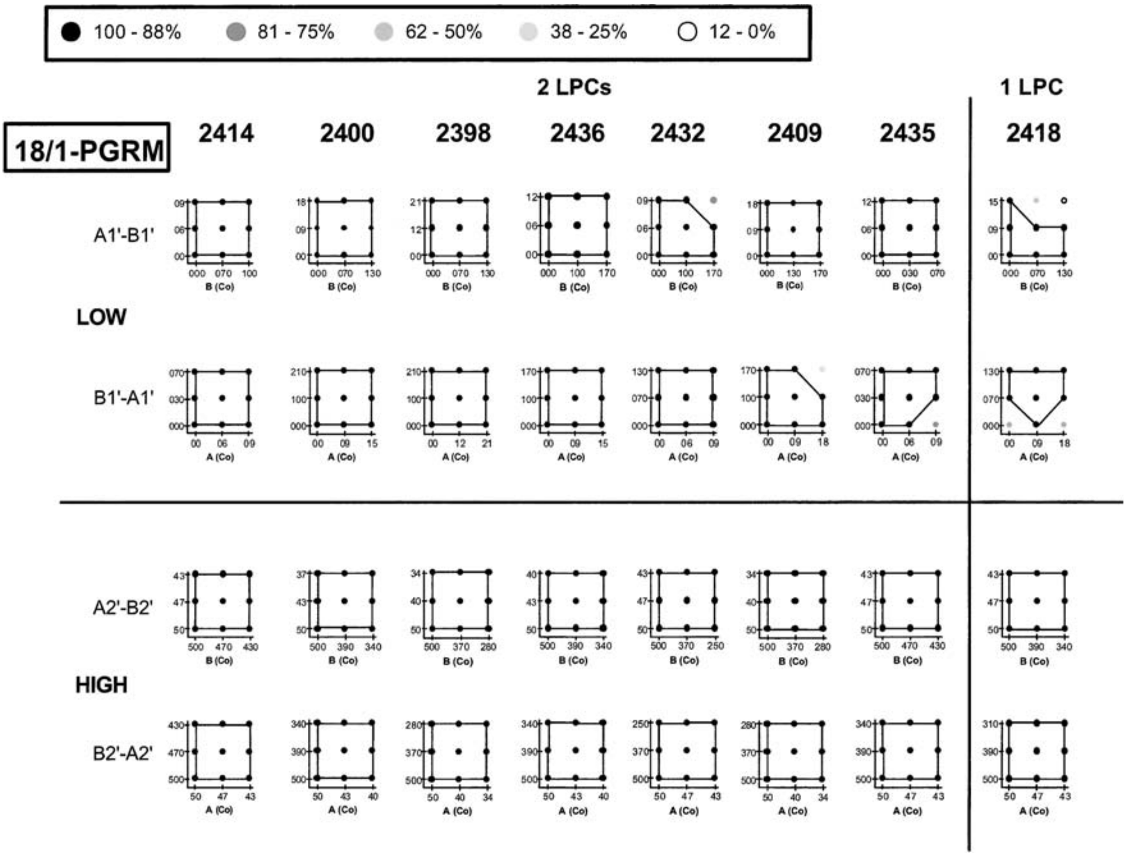

Fig. 9. Performance occasioned by all cross-class probes presented in the A`-B` and B`-A` formats when using 18/1-PRGM test schedule. See text for explanation of graph format. Each column contains data for 1 subject. Top two rows: Outcomes of the A`-B` and B`-A` tests for the low A/B classes. Bottom two rows: Outcomes of the A`-B` and B`-A` tests for the high A/B classes. The vertical line separates subjects who formed two linked perceptual classes (to the left of the line) from the one who did not form any linked perceptual classes (to the right of the line).

2001; Lea, 1984; Reeve & Fields, 2001; Wasserman et al., 1988): All of the stimuli in a set must occasion the selection of the same comparison with similar high probabilities, the stimuli in different sets must occasion the selection of different comparisons, and many of the stimuli in a set must be discriminable from each other.

The performance occasioned by the variants in the primary generalization tests satisfied these criteria, demonstrating the emergence of two perceptual classes each in the A and B domains. While class membership was determined from the conditional selections made during the generalization tests, within-class discriminability was determined at the same time by the measurement of response speeds occasioned by the conditional selections. Therefore, the membership of variants in a perceptual class could not have been due to a failure to discriminate among those variants (Fields, Matneja, et al., 2002; Fields & Reeve, 2001; Lashley & Wade, 1946).

Functional independence of linked perceptual classes. The results of the primary generalization tests demonstrated the functional independence of the perceptual classes at each end of the A domain and at each end of the B domain (Fields et al., 2003; Reeve & Fields, 2001; Sidman, 1987; Wasserman et al., 1988). Thus the linked perceptual classes, which were constituted of the previously mentioned classes, also would have to be functionally independent of each other (Fields, Matneja, et al., 2002; Fields & Reeve, 2001; Innis et al., 1998; Wasserman et al., 1988). This inference

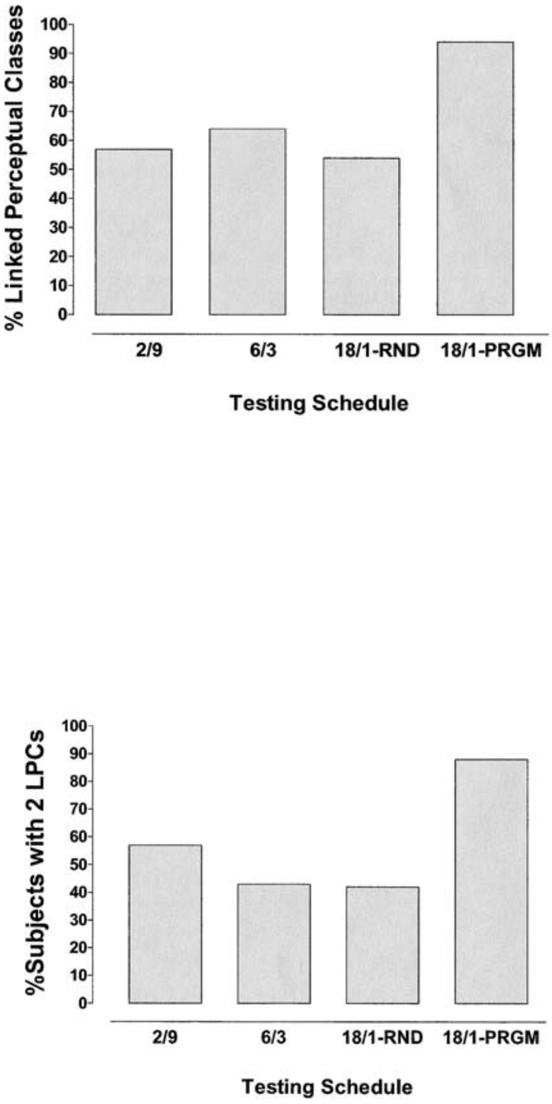

Fig. 10. The effects of the 2/9, 6/3, 18/1-RND, and 18/1-PRGM test schedules in the emergence of linked perceptual classes. Top: Effects of the test schedules on the percentage of possible linked perceptual classes that emerged for all of the subjects in a test condition. Bottom: Percentage of subjects in a test condition who showed the emergence of both potential linked perceptual classes.

was confirmed by the cross-class probe data which showed that subjects typically selected the neither comparisons when they did not select comparisons that came from the same set as a sample stimulus. Thus two lines of evidence support the view that the linked perceptual classes in the present experiment were functionally independent of each other.

Ordering of probes and linked perceptual class formation. The effect of randomization was revealed through a comparison of the effects of the two 18/1 test schedules on the emergence of linked perceptual classes. In the 18/1-RND test, the values of the sample and comparison stimuli did not vary systematically across the succession of test blocks. In the 18/1-PRGM test, the values of the sample and comparison stimuli were varied systematically across the succession of test blocks. The randomly sequenced probe types resulted in the emergence of about half of the linked perceptual classes by some of the subjects. The programmed sequence of probe types resulted in the emergence of almost all possible linked perceptual classes by most subjects. Thus a programmed presentation of cross-class probes played an important role in the immediate emergence of linked perceptual classes.

The results of the present experiment showed that the programmed introduction of the cross-class probes in a specific order maximized the formation of linked perceptual classes. The cross-class probes, however, could have been presented in many other systematic orders. Additional research will be needed to determine whether all or some of them would maximize the immediate emergence of linked perceptual classes. If some testing orders maximize yields and others do not, the differences in testing orders might well clarify the behavioral processes responsible for the emergence of linked perceptual classes.

Generalization and the effects of test schedules. Primary generalization probably played a significant role in the differential effects of the test schedules on the emergence of linked perceptual classes. The 18/1PRGM schedule could have maximized the emergence of classes through generalization in the following manner. In the first three test blocks (i.e., aa, ma, and ba), the probes differed minimally from the previous cross-class probe by the value of an adjacent sample stimulus (e.g., anchor to midpoint to boundary). Small, albeit discriminable, differences in stimulus value maximize the likelihood of generalization between stimuli. Thus, in the 18/1PRGM test, the small changes in stimulus value across successive test blocks were likely to maximize

Table 5

Error analysis: Number of cross-class probes that did not occasion mastery-level performances in each test schedule. Data were summed across all classes and subjects in a test condition. The last column lists the proportion of errors for a test of a given type relative to all errors in the group. Tests are grouped in terms of the stimuli used as sample and comparison (a = anchor, b = boundary, m = midpoint). For example, ''m-a'' indicates all probes in which a midpoint stimulus was the sample and an anchor stimulus was the comparison.

		Test type					
Test schedule	Sa-Co	A1-B1	B1-A1	A2-B2	B2-A2	Sum	Proportion of all errors
18/1-PRGM	a-a	0	1	0	0	1	0.14
	b-b	2	1	0	0	3	0.43
	m-a	0	0	0	0	0	0.00
	b-a	0	1	0	0	1	0.14
	a-m	0	0	0	0	0	0.00
	m-m	0	0	0	0	0	0.00
	b-m	0	0	0	0	0	0.00
	a-b	0	1	0	0	1	0.14
	m-b	1	0	0	0	1	0.14
18/1-RND	a-a	0	1	0	0	1	0.02
	b-b	1	1	1	1	4	0.07
	m-a	0	1	2	4	7	0.13
	b-a	3	5	2	2	12	0.22
	a-m	0	0	1	0	1	0.02
	m-m	0	0	0	2	2	0.04
	b-m	2	3	1	1	7	0.13
	a-b	3	2	3	3	11	0.20
	m-b	1	3	2	3	9	0.17
6/3	a-a	1	0	0	0	1	0.03
	b-b	2	1	1	2	6	0.18
	b-a	2	3	0	2	7	0.21
	m-a	0	1	0	1	2	0.06
	a-m	0	1	0	1	2	0.06
	m-m	1	0	0	2	3	0.09
	b-m	3	2	1	2	8	0.24
	a-b	0	1	1	0	2	0.06
	m-b	0	1	1	1	3	0.09
2/9	a-a	0	0	1	1	2	0.03
	b-b	0	2	3	3	8	0.12
	m-a	2	1	1	2	6	0.09
	b-a	3	3	2	3	11	0.16
	a-m	1	1	3	1	6	0.09
	m-m	2	0	3	3	8	0.12
	b-m	2	1	3	2	8	0.12
	a-b	3	2	3	3	11	0.16
	m-b	1	1	4	3	9	0.13

the generalization of class-based performances from test block to the next. In addition, the comparisons used in the first three test blocks were the most discriminable stimuli in their classes (i.e., the anchor stimuli), thereby maximizing the likelihood of emitting class-based performances. Finally, the comparison sets used in successive sets of three tests blocks differed incrementally from one 3-block cluster to the next. This, too, most likely maximized generalization of class-based performances by the comparisons. The combined effects of these schedule-determined operations on generalization maximized the emergence of linked perceptual classes under the 18/1PRGM testing schedule.

When the three other testing schedules were used, many different cross-class probes were introduced in a randomized sequence in the same block. This procedure increased the disparity between stimuli that were presented on adjacent trials, of necessity reduced the effects of generalization of class-consistent comparison selection across test trials, and consequently reduced the likelihood of linked perceptual class formation.

Probe types per block and linked perceptual class formation. The results of prior research suggested that the likelihood of linked perceptual class formation should be an inverse function of the number of different types of cross-class probes included in a test block. In the current experiment, however, the same modest yields were obtained over a nine-fold variation in the number of different probe types presented in a test block (nine probes per block to one probe per block). Clearly, then, the number of different probe types in a test block was not a variable that was a determinant of linked perceptual class formation.

The same modest yields obtained when the test blocks contained nine, three, or one probe type per test block suggests that the yields were determined by some factor that was constant across these test conditions. One such factor was the randomized presentation of trials of different probe types. This occurred within test blocks in the 2/9 and 6/3 tests, and across test blocks in the 18/1-RND test. Therefore, it is plausible to conclude that the randomized presentation of trials for different probe types was responsible for the constancy of the yields across testing schedules.

Nonemergence of linked perceptual classes. In each of the test schedules, some failures of class formation occurred because of a breakdown of the previously established cross-class conditional discriminations and their symmetrical counterparts. Furthermore, most of these breakdowns occurred when the probes consisted of boundary rather than anchor stimuli. This suggests that all of the stimuli in a perceptual class were not equally related to each other. Perhaps, then, linked perceptual class formation could be enhanced by more training of the boundary-to-boundary than anchor-to-anchor conditional discriminations and the overtraining of both of these relations.

Other failures of linked perceptual class formation were due to errors occasioned by the 14 novel cross-class probes. These errors could reflect the absence of an emergent linkage among the stimuli in two intact perceptual classes, and/or a breakdown of the underlying perceptual classes. These two sources of failure were not isolated in the current experiment because the integrity of the perceptual classes was not evaluated during the cross-class tests.

Linked perceptual classes and generalized equivalence classes. A fully elaborated generalized equivalence class consists of at least three perceptual classes, the members of which all occasion the mutual selection of each other (Fields & Reeve, 2001). For example, such a class might consist of many pictures of domestic cats (Perceptual Class A`), the many sounds made by domestic cats (Perceptual Class B`), and the word cat written in many different fonts (Perceptual Class C`). Theoretically, such a class can be established by the formation of an equivalence class consisting of one member of each of the three perceptual classes by training conditional discriminations between the anchor stimuli of Perceptual Classes A`, B`, and C`: that is, A^a-B^a and B^a-C^a. Although fully elaborated generalized equivalence classes provide a behavior analytic model of the complex categories that emerge in natural settings, to date, experiments have not explored the formation of such classes. Rather, they have explored the formation of a minimally elaborated generalized equivalence class: an equivalence class that is linked to at least one perceptual class (Adams, Fields, & Verhave, 1993b; Barnes & Keenan, 1993; Belanich & Fields, 2003; Branch, 1994; Cowley, Green, & Braunling-McMorrow, 1992; DeGrandpre, Bickel, & Higgins, 1992; Fields, Adams, Brown, & Verhave, 1993; Fields & Reeve, 2000; Fields, Reeve, Adams, & Verhave, 1991; Fields, Reeve, Adams, et al., 1997; Haring, Breen, & Laitinen, 1989; Lane, Clow, Innis, & Critchfield, 1998; Mackay, Stromer, & Serna, 1997; Rehfeldt & Hayes, 2000).

The linked perceptual classes described in the current study share many of the functional properties of fully elaborated generalized equivalence classes in the sense that both involve the emergence of relations between the members of distinct perceptual classes. Indeed, Fields and Reeve (2001) proposed that a linked perceptual class is the minimal form of a fully elaborated generalized equivalence class. Cross-class probes like those described in the current experiment could be used to measure the emergence of fully elaborated generalized equivalence classes, and the testing schedules used to program the introduction of the cross-class probes also might influence the formation of fully elaborated generalized equivalence classes. Thus, after training A^{a}-B^{a} and B^{a}-C^{a}, as described in the preceding paragraph, the emergence of a fully elaborated generalized equivalence class could be documented by the presentation of the following set of cross-class probes: A`-B`, B`-A`, B`-C`, C`-B`, A`-C`, and C`-A`. In addition, the programmed introduction of these probes might well enhance the formation of fully elaborated generalized equivalence classes.

Complex categories and generalized equivalence classes. Generalized equivalence classes are defined in terms of their structure and function. Structurally, such a class contains some stimuli that bear a physical resemblance to each other and others that are perceptually disparate. Functionally, after the establishment of a few relations among the stimuli in the set, all of the remaining stimuli occasion the mutual selection of each other without benefit of direct training. In addition, a response trained to one class member is evoked by the remaining stimuli in such a class. These structural and functional properties also characterize the stimuli that constitute the complex categories that are found in natural setting (Lane, Clow, Innis, & Critchfield, 1998) and approximations to these classes that have been called natural kinds (Gelman, 1988), fuzzy superordinate classes (Rosch & Mervis, 1975), semantic memory networks (Collins & Quillian, 1969), or amodal relations (Bahrick & Pickens, 1994), and have been said to reflect the process of intersensory perception (Lewkowicz, 1994). Given the structural and functional similarities of generalized equivalence classes and classes denoted by the latter terms, the parameters identified in the present experiment might also shed light on the environmental variables that influence the formation of complex categories, regardless of denotation.

REFERENCES

Adams, B. J., Fields, L., & Verhave, T. (1993a). Effects of test order on intersubject variability during equivalence class formation. *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.

Bahrick, L. E., & Pickens, J. N. (1994). Amodal relations: The basis for intermodal perception and learning in infancy. In D. J. Lewkowicz & R. Likliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 205–233). Englewood, NJ: Erlbaum.

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.

Belanich, J., & Fields, L. (2003). Generalized equivalence classes as response transfer networks. *The Psychological Record, 53,* 373–413.

Bhatt, R. S., Wasserman, E. A., Reynolds, W. F., Jr., & Knauss, K. 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,* 219–234.

Branch, M. (1994). Stimulus generalization, stimulus equivalence, and response hierarchies. In S. C. Hayes, L. J. Hayes, M. Sato & K. Ono (Eds.), *Behavioral analysis of language and cognition* (pp. 51–70). Reno, NV: Context Press.

Buffington, D. M., Fields, L., & Adams, B. J. (1997). Enhancing the formation of equivalence classes by pretraining of other equivalence classes. *The Psychological Record, 47,* 1–20.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240–248.

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.

Cumming, W. W., & Berryman, R. (1965). The complex discriminative operant: Studies of matching-to-sample and related problems. In D. I. Mostofsky (Ed.), *Stimulus Generalization* (pp. 284–33). Stanford, CA: Stanford University Press.

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.

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

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., 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., Matneja, P., Varelas, A., & Belanich, J. (2003). Mutual selection and membership in open-ended classes: Variant-to-base and base-to-variant testing. *The Psychological Record, 53,* 287–311.

Fields, L., Matneja, P., Varelas, A., Belanich, J., Fitzer, A., & Shamoun, K. (2002). The formation of linked perceptual classes. *Journal of the Experimental Analysis of Behavior, 78,* 271–290.

Fields, L., & Reeve, K. F. (2000). Synthesizing equivalence classes and natural categories from perceptual and relational classes. In J. C. Leslie & D. Blackman (Eds.), *Issues in experimental and applied analyses of human behavior* (pp. 59–84). Reno, NV: Context Press.

Fields, L., & Reeve, K. F. (2001). A methodological integration of generalized equivalence classes, natural categories, and cross modal perception. *The Psychological Record*, *51*, 67–88.

Fields, L., Reeve, K. F., Adams, B. J., Brown, J. L., & Verhave, T. (1997). Predicting the extension of equivalence classes from primary generalization gradients: The merger of equivalence classes and perceptual classes. *Journal of the Experimental Analysis of Behavior*, *68*, 68–92.

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

Fields, L., Reeve, K. F., Matneja, P., Varelas, A., Belanich, J., Fitzer, A., & Shamoun, K. (2002). The formation of a generalized categorization repertoire: Effect of training with multiple domains, samples, and comparisons. *Journal of the Experimental Analysis of Behavior*, *78*, 291–314.

Fields., L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., & Hobbie, S. A. (1997). Using the simultaneous protocol to study equivalence class formation: The facilitating effects of nodal number and size of previously established equivalence classes. *Journal of the Experimental Analysis of Behavior*, *67*, 367–389.

Figuracion, D. (1998). Morph, Version 2.5. San Diego, CA: Gryphon Software Corp.

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

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.

Honig, W. K., & Stewart, K. E. (1988). Pigeons can discriminate locations presented in pictures. *Journal of the Experimental Analysis of Behavior*, *50*, 541–551.

Innis, A., Lane, S. D., Miller, E. R., & Critchfield, T. S. (1998). Stimulus equivalence: Effects of a default-response option on emergence of untrained stimulus relations. *Journal of the Experimental Analysis of Behavior*, *70*, 87–102.

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

Lane, S. D., Clow, J. K., Innis, A., & Critchfield, T. S. (1998). Generalization of cross-modal stimulus equivalence classes: Operant processes as components in human category formation. *Journal of the Experimental Analysis of Behavior*, *70*, 267–280.

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. Roitblat, T. G. Bever & H. S. Terrace (Eds.), *Animal cognition* (pp. 263–276). Hillsdale, NJ: Erlbaum.

Lewkowicz, D. J. (1994). Development of intersensory perception in human infants. In D. J. Lewkowicz & R. Likliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 165–203). Englewood, NJ: Erlbaum.

Mackay, H. A., Stromer, R., & Serna, R. W. (1997). Emergent behavior and intellectual functioning: Stimulus classes, generalization, and transfer. In S. Soraci & W. J. McIlvane (Eds.), *Perspectives on fundamental processes in intellectual functioning* (pp. 287–310). Greenwich, CT: Ablex Publishing.

Reeve, K. F., & Fields, L. (2001). Perceptual classes established with forced-choice primary generalization tests and transfer of function. *Journal of the Experimental Analysis of Behavior*, *76*, 95–114.

Rehfeldt, R. A., & Hayes, L. J. (2000). The long-term retention of generalized equivalence classes. *The Psychological Record*, *50*, 405–428.

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

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis*, *22*, 1–8.

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.

Wright, A. A., Cook, R. G., Rivera, J. J., Sands, S. F., & Delius, J. D. (1988). Concept learning by pigeons: Matching-to-sample with trial-unique video picture stimuli. *Animal Learning and Behavior*, *16*, 436–444.

Received June 24, 2003
Final acceptance June 28, 2005