CLASSIFICATION OF VOWELS AND CONSONANTS BY INDIVIDUALS WITH MODERATE MENTAL RETARDATION: DEVELOPMENT OF ARBITRARY RELATIONS VIA MATCH-TO-SAMPLE TRAINING WITH COMPOUND STIMULI

SCOTT D. LANE AND THOMAS S. CRITCHFIELD

AUBURN UNIVERSITY

This study explored whether an identity-matching-based stimulus equivalence procedure could be used to teach vowel and consonant stimulus classes to 2 adolescent females with moderate mental retardation. Delayed match-to-sample trials presented a compound sample stimulus consisting of printed letters and a spoken word ("vowel" or "consonant"). The correct comparison stimulus matched only one of the letters in the compound sample. Subsequently, test trials assessed whether arbitrary relations had formed among the individual stimuli from each compound sample and whether stimuli from different compound samples had merged into larger stimulus classes. Both participants acquired five-member classes of vowel and consonant stimuli, which subsequently generalized to vocal classification and to identification in the context of four-letter words. Follow-up tests showed that the generalized performances remained intact after 6 weeks. These procedures suggest an economical approach to stimulus class development.

DESCRIPTORS: identity matching, complex stimuli, stimulus equivalence, developmental disability, computer touch screen

---

Complex stimulus control is integral to classification and other forms of conceptual learning (e.g., Barsalou, 1992; Herrnstein, 1990). Instances in which stimulus control is restricted or *overselective* may impede the acquisition of adaptive skills (Lovaas & Schreibman, 1971), and behavior controlled

by a stimulus complex may not occur when individual elements of the complex are present (Halle & Holt, 1991). Such cases, perhaps common in persons with developmental disabilities, illustrate the limitations of behavior not controlled by multiple stimulus attributes. Accordingly, instructional techniques that seek to promote classification (e.g., conceptual) skills will be most effective by establishing "broad networks of adaptive discriminative control" (Maguire, Stromer, Mackay, & Demis, 1994, p. 753). The practical constraints of classroom environments place a premium on teaching techniques that are able to establish these broad networks of stimulus control with a minimum of training investment. Stimulus equivalence procedures (Sidman & Tailby, 1982) may provide one basis for doing so.

Equivalence procedures have generated much enthusiasm as a potential basis for teaching techniques because they can produce new skills, incorporating large networks of stimulus relations, with notable

economy of training (Sidman, 1994). In persons with mental retardation or brain injury, who may learn slowly by other means, equivalence-derived procedures have been used to promote rapid acquisition of several socially important skills, including the rudiments of spelling, oral reading, classification, and quantification (Kennedy & Serna, 1995; Mackay, 1985; Sidman, 1971; Sidman & Kirk, 1974; Stoddard, Brown, Hurlbert, Manoli, & McIlvane, 1989; Stromer & Mackay, 1992). Unfortunately, relatively little is known about the capacity of equivalence-based procedures to establish important forms of multielement stimulus control, because most of the existing research on equivalence has employed single-element stimuli.

Stimulus equivalence procedures require teaching conditional discriminations that involve overlapping stimulus relations. For example, a student might first be taught to match the letters A and D to the labels "vowel" and "consonant," respectively. In match-to-sample format, given the spoken word "vowel" as sample and the printed letters A and D as comparisons, selection of A would be reinforced. Given the spoken word "consonant" as sample and the same comparisons, selection of D would be reinforced. To create overlapping relations, the student could then be taught to match the labels to different letters. For example, with "vowel" as sample and the printed letters O and V as comparisons, selection of O would be reinforced. With "consonant" as sample and the same comparisons, selection of V would be reinforced. Although only four relations are taught, others often emerge as a result of the training. Of particular importance is the fact that the student may now be able to associate letters that have no direct training history of association. For example, when shown the printed letter A, the student may now reliably select O from among several alternatives (e.g., O and D). When shown

O, the student may reliably select A from among several alternatives. The letters D and V may be matched in a similar format. Further untrained performances, not directly implicated in the definition of equivalence, also might emerge in some cases. The student described above, for example, might now be able to label (name) the letters A and O as vowels and the letters D and V as consonants (e.g., see Sidman, Willson-Morris, & Kirk, 1986).

Equivalence-based procedures thus promote substantial new learning in proportion to training investment, at least where relations among single-element stimuli are concerned. Only recently have studies been designed to investigate whether the same advantages might apply to learning that involves multielement stimuli. A common feature of these equivalence-based studies is that participants are taught to match multielement sample stimuli to single-element comparisons. For the present purposes, the prominent finding has been that untrained relations can emerge among the individual elements of the samples. This can occur when the training requires matching of nonidentical stimuli (e.g., Markham & Dougher, 1993; Stromer, McIlvane, & Serna, 1993; Stromer & Stromer, 1990a, 1990b), or when the training mimics, or is based on, identity matching (Maguire et al., 1994; Schenk, 1993, 1995; Smeets & Striefel, 1994; Stromer & Mackay, 1992, 1993).

To illustrate the identity-based variation of a multielement delayed match-to-sample procedure: A student could be presented, on some trials, with the printed vowels AO as a complex sample stimulus and, after the AO complex was removed, the student would receive reinforcement for selecting A from among the alternatives A and D. On other trials, given the AO sample, the student could be taught to select O from among the alternatives O and V. In both cases, physical similarity (e.g., matching

identical stimuli) serves as the primary basis for correct responding. Subsequently, however, when A is presented alone as a sample stimulus, the student may be able to select O from among the alternatives O and V, despite the fact that, during training, A and O were never presented together as sample and comparison stimuli. To foreshadow the methods of the present study, it should be noted that conditional discrimination procedures require that the same selection be correct and incorrect in different contexts. Thus, the training would also include trials with DV as the sample stimulus. On these trials, the student would select D (not A) or V (not O), producing a D-V relation analogous to the A-O relation. One noteworthy feature of this procedure is that a student who is asked to make responses that can be based on physical similarity alone may also come to demonstrate arbitrary relations between the nonidentical elements of the sample stimulus.

The preceding example may be viewed as speculative. Currently, there exist only a few demonstrations that equivalence-based procedures can be used with applied populations to establish arbitrary relations that involve complex, educationally relevant stimuli. These studies have focused on creating rudimentary spelling skills in persons with developmental disabilities (Stromer & Mackay, 1992, 1993). For example, Stromer and Mackay (1992) used identity-based procedures with complex stimuli to instate arbitrary relations between pictures of objects and the written names of those objects in children and adolescents with developmental disabilities. A typical complex stimulus involved the simultaneous presentation of a picture of a dog and the printed word "dog." When shown this complex stimulus as sample and pictures of a dog and a cat as comparison stimuli, participants were taught to select the picture of the dog. Similarly, when shown the same stimulus and several letters,

the participants were trained to touch the letters that spell "dog." Without additional training, participants were able to select the picture of the dog when presented only with the printed word "dog" as a sample, and to spell "dog" when presented only with the picture of the dog as the sample. These performances emerged for a number of pictures and printed words. The present study, by contrast, was an attempt to use equivalence-based procedures, involving complex stimuli, to establish stimulus classes of vowels and consonants.

This study also explored a means of improving economy of training by capitalizing on the observation, in equivalence studies using single-element stimuli, that separately created stimulus classes that share a common member can merge into a single, larger class. Class-merger outcomes can produce more untrained relations (i.e., new skills) than would be expected within separate, smaller classes (Saunders, Saunders, Kirby, & Spradlin, 1988; Saunders, Wachter, & Spradlin, 1988; Sidman, Kirk, & Willson-Morris, 1985). To illustrate, in a single-element context, a student could be taught, on separate trials, to match the letters A and O to the spoken word "vowel," producing the stimulus class A-O-"vowel." Similar training could create a second stimulus class, E-U-"vowel." Without further training, the student may be able to match the letters E, U, A, and O to one another. The letters, despite having no history of direct association, thus may become related through their shared associate, "vowel." Because this is a conditional discrimination procedure, a second class of stimuli, serving as incorrect selections during vowel training and testing, would serve as correct selections at other times. For example, following training in the format just described, the letters D, V, K, and T could become related through a shared associate, "consonant."

Following from the findings described

above, the present study was an attempt to create a single training and testing package that (a) used educationally related stimuli in an identity-based training procedure involving compound (complex) stimuli to instate arbitrary relations and (b) promoted the merger of stimulus classes through shared members. If successful, such an approach would produce greater economy of training (and acquisition of new skills) than either of the component procedures typically produces individually and could form the basis for effective, economical teaching techniques.

We attempted to generate arbitrary stimulus relations using letters of the alphabet and spoken words ("vowel" and "consonant") as stimuli. Identity-matching-based training with complex sample stimuli was used to teach the baseline stimulus relations (see Stromer & Mackay, 1992, 1993). The training was designed to create two three-member "vowel" stimulus classes and two three-member "consonant" stimulus classes. If these stimulus classes formed, participants would acquire 16 new arbitrary stimulus relations. Furthermore, because each pair of three-member stimulus classes shared a common auditory (spoken) element, they would be expected to merge into a single five-member class. If so, then 16 additional untrained stimulus relations would be acquired without any direct history of association. Thus, the training of eight complex stimulus relations in an identity-based procedure could produce up to 32 arbitrary relations. Finally, as a measure of generalization of effects, the participants completed two novel follow-up tasks in which they could demonstrate up to 40 additional untrained categorization performances related to those just mentioned.

METHOD

Participants, Setting, and Apparatus

Two adolescent females with Down syndrome participated. Patty was 14 years old,

had a mental age equivalent score of 4 years 1 month on the Peabody Picture Vocabulary Test (PPVT), and was enrolled in a middle school special education program. Beth was 12 years old, had a mental age equivalent of 4 years 2 months on the PPVT, and was enrolled in a elementary school special education program. Both participants could name all of the English-language letters used as stimuli in this study. Neither had taken part in any previous stimulus control experiments. Both participants signed assent forms after participating in an initial six-trial tutorial segment. They received $4.00 for each day they took part in the experiment. Visits to the laboratory took place after school and lasted 15 to 30 min. Each day, participants completed between three and nine sessions, determined by the length of sessions and the participant's decision to discontinue participation for that day. Sessions lasted approximately 1 to 12 min, depending on the number of trials in that session. Each participant made about 20 visits over a period of 1 month.

Each participant worked in a quiet room in front of a VGA video monitor equipped with a touch-sensitive screen that recorded responses and speakers that permitted the delivery of auditory stimuli. An experimenter sat approximately 2 m directly behind the participant's work station at a table equipped with an IBM-compatible microcomputer and a monochrome monitor. A custom program written in MicroSoft Visual Basic® controlled experimental procedures and collected the data. On each trial throughout the study, the computer screen displayed a rectangular box near the top of the screen in which sample stimuli were presented and three square boxes near the bottom of the screen in which one or usually two comparison stimuli were presented (see Figure 1). All boxes and stimuli were yellow against a blue background. Only touches inside the
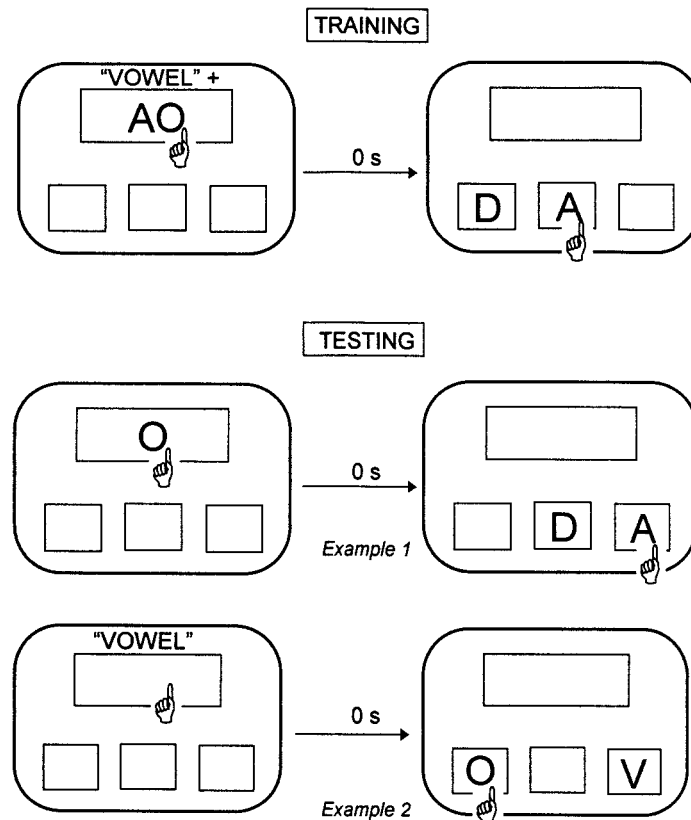
Figure 1.   Diagram of representative training and test trial arrangements used in the study (top, middle, and bottom panels, respectively). Letters in the boxes depict the visual stimuli shown on the computer screen. The "vowel" label above the sample box refers to the spoken word that (along with "consonant" on other trials) was presented simultaneously with the letter or letters.

sample and comparison boxes served as functional responses.

Feedback. A variety of consequences were used. During training, following each match-to-sample trial in which feedback was scheduled, the screen cleared and a yellow "happy" face (correct responses) or a purple "sad" face (incorrect responses) appeared. Simultaneously, a recorded spoken message appropriate to the participant's choice (either "very good" or "sorry, try again") was played through a speaker connected to the computer. A touch anywhere on the feedback screen cleared the display and, after a 2-s intertrial interval, initiated the next trial.

Throughout training and testing, each block of 12 trials was followed by a short break during which the participant was given a cheese puff (which was consumed immediately) and a plastic poker chip, which she deposited into a coffee can (the "bank") through a slot in the top. The session resumed immediately afterward. Chips were exchanged for toys, balloons, posters, and books at the end of each visit to the laboratory.

White chips were given during training sessions that were not yet completed (e.g., following Trial 12 of a 24-trial session) and during all test sessions. When white chips were given, the experimenter said either "Good work," or "You're working very hard," but did not refer to specifics of performance. Red chips were given following

the end of training sessions in which participants did not meet the criterion for mastery. When red chips were given, the experimenter said, "Good try, but you missed a few. Let's see if we can get them all this time, OK?" Blue chips were given following training sessions in which the criterion for mastery was met. When blue chips were given, the experimenter said, "That's great! You were just about perfect that time, so you get blue money." No other differential consequences were associated with the color of the chips.

PHASE 1: TUTORIAL AND PRETESTING

Procedure

Tutorial. The tutorial was designed to familiarize the participants with the identity and arbitrary match-to-sample procedures. Three identity trials and three arbitrary trials were presented. During the three identity trials, a square was presented as the sample. After the participant touched the sample square, the sample box turned blank, and a single comparison square was presented in one of the three comparison boxes. The other boxes were blank. Touching the comparison square was reinforced. During the arbitrary trials, a circle was presented as the sample. When it was touched, it disappeared, and a single triangle appeared in one of the comparison boxes. Touching the triangle was reinforced. Both participants touched the correct comparison box on all six trials, thus assuring that they distinguished between the visual comparison stimuli and the empty boxes.

Pretest: Arbitrary matching. This pretest assessed match-to-sample performance on the arbitrary stimulus relations expected following training with complex stimuli. Included were 32 possible relations among five vowel stimuli (the printed letters A, O, E, and U, plus the spoken word "vowel," which was

presented only as a sample stimulus) and among five consonant stimuli (the printed letters D, V, K, and T, and the spoken word "consonant," which was presented only as a sample stimulus). Figure 1 illustrates the trial types used in the pretest. Each trial began with the presentation of either a spoken word ("vowel" or "consonant") or a printed letter (A, O, E, U, D, V, K, or T). If the sample was a spoken word, the word was presented once after the participant touched the blank sample box. Then, two letters (a vowel and a consonant) were presented as comparisons. Touching the letter that matched the class of the sample was counted as correct. If a printed letter was presented as the sample, touching that letter resulted in its removal and, immediately following, the presentation of a vowel and a consonant as comparisons. Touching the comparison letter that matched the class of the sample letter was recorded as correct. No differential feedback concerning the accuracy of the response was provided. However, after 12 trials the participant was given a cheese puff and a white token to maintain on-task behavior.

Trial types, shown in Table 1, were intermixed randomly throughout the session. Each relation was tested three times per session, with the correct comparison occurring at least once in each comparison location. Data were recorded by computer for two 96-trial sessions.

Pretest: Oral naming. This test assessed the participants' preexperimental ability to orally classify each of the letters as vowel or consonant. The experimenter sat directly in front of the participant with notecards (3 in. by 5 in.), on which was printed one of the eight letters used in the study. The cards were shown one at time. During the first presentation of each card, the experimenter asked, "Is this a consonant or a vowel?" Vowel and consonant cards were presented in a random sequence. During the second

Table 1
Summary of the Pretest Procedures of Phase 1

Trial type	Sample stimulus	Comparison stimuli		Sample stimulus	Comparison stimuli	
Pretest trials for Set 1 training	"vowel"	A̲	D	"consonant"	A	D̲
	"vowel"	O̲	V	"consonant"	O	V̲
	A	O̲	V	D	O	V̲
	O	A̲	D	V	A	D̲
Pretest trials for Set 2 training	"vowel"	E̲	K	"consonant"	E	K̲
	"vowel"	U̲	T	"consonant"	U	T̲
	E	U̲	T	K	U	T̲
	U	E̲	K	T	E	K̲
Pretest trials across Set 1 and Set 2	A	E̲	K	D	E	K̲
	A	U̲	T	D	U	T̲
	O	E̲	K	V	E	K̲
	O	U̲	T	V	U	T̲
	E	A̲	D	K	A	D̲
	E	O̲	V	K	O	V̲
	U	A̲	D	T	A	D̲
	U	O̲	V	T	O	V̲

Note. Underlined letters indicate the correct comparison stimulus.

presentation of each card, the experimenter asked, "Is this a vowel or a consonant?" On subsequent presentations, the experimenter asked, "What is this?" The experimenter presented each of the eight cards three more times, for a total of six trials per letter. No differential feedback was given following these trials. Data were recorded by the experimenter.

During both this pretest and its associated posttest in Phase 3, oral naming data were recorded by a single experimenter, and interrater reliability measures were not obtained. This scoring occurred during short sessions, under minimal time pressure, and involved a small number of response categories that incorporated easily discriminable responses—all factors that normally contribute to accurate ratings by human observers (e.g., Harris & Lahey, 1982; Reid, 1982). However, the absence of interrater reliability measures dictates that the oral naming data of the present study be considered preliminary.

Results

Pretest: Arbitrary matching (Sets 1 and 2). Across individual stimulus relations, perfor-

mances on these pretests ranged from 16% to 83% accurate. No systematic patterns were evident across relations, and, considering all 192 pretest trials combined, performance was near chance for both Patty (45% accurate) and Beth (50% accurate).

Pretest: Oral naming. Each letter was tested six times. As with the arbitrary matching pretests, performance varied nonsystematically across individual stimuli, and considering all pretest trials combined, performance was near chance for both Patty and Beth (50% and 48%, respectively).

PHASE 2: TRAINING AND TESTING

Procedure

Training: Identity matching with complex stimuli, Set 1 vowels and consonants. This training sought to establish the relations that were prerequisite for the formation of two three-member stimulus classes, Set 1 vowels (A, E, "vowel") and Set 1 consonants (D, V, "consonant"). The top panels of Figure 1 depict the basic procedure. After the initial training trials for each letter, on each trial,

the participant's screen showed one sample box and three comparison boxes aligned horizontally below. A complex sample stimulus appeared first. Touching both elements of the sample, in either order, immediately cleared the sample box and produced two comparison stimuli, one of which was identical to an element of the sample. Comparison stimuli appeared in different boxes across trials, with the locations counterbalanced within and across all sessions. A touch anywhere inside one of the comparison boxes was counted as a valid response. Touching the comparison that was identical to one of the sample elements counted as a correct response and was followed by reinforcement. The protocol required touching a blank box to be counted as an incorrect response, but no empty boxes were selected during the experiment.

An important feature of the matching procedure employed here was that sample and comparison stimuli were never simultaneously present. Under such conditions, it is impossible for a participant to predict which sample element will appear as the correct comparison. Thus, the reinforcement contingencies encourage attention to all sample elements that can appear as comparisons (Baron & Menich, 1985; Stromer, McIlvane, Dube, & Mackay, 1993). Accordingly, the present procedure almost certainly promotes different types of stimulus-viewing behavior than traditional identity matching procedures that involve single-element stimuli. Nevertheless, we describe our training procedure as identity based because, throughout training, the correct comparison stimulus was always physically identical to one of the elements in the complex sample. In addition, the limited research available to date suggests that individuals with moderate developmental disabilities acquire this kind of matching readily, just as they typically do in traditional identity matching procedures

(Maguire et al., 1994; Stromer & Mackay, 1992).

The training followed 13 steps that gradually instated the conditional relations that were prerequisite to the formation of stimulus classes. For example, relations between vowels and consonants in Set 1 were trained separately before being combined in review sessions. In addition, early in training, complex sample stimuli contained one printed letter and the spoken word "vowel" or "consonant." As training progressed, the stimuli were expanded to include two letters and a spoken word (as shown in Table 2).

In the final steps of training, participants presented with a three-element complex sample stimulus could routinely select the correct comparison stimulus. For example, given a sample consisting of the printed letters D and V accompanied by the spoken word "consonant," participants could select D as the correct comparison from the choices D and A. An important feature of training in these sessions was that "consonant" was the auditory accompaniment of the visual compound DV, and "vowel" was the auditory accompaniment of the visual compound AO. Note, however, that the training contingencies did not require attention to the auditory stimuli. Sessions in the final three steps of training were 48 trials long, and all of the trained relations involving vowel and consonant stimuli from Set 1 were intermixed. Feedback followed all trials until a participant scored at least 46 of 48 correct in one session. In the final two steps of training, the reinforcement probability was decreased to .70 until the same mastery criterion was met and was then reduced to .25.

Testing: Untrained arbitrary relations, vowels and consonants from Set 1. Upon completion of the training sequence, testing was initiated to assess whether arbitrary relations had been acquired among the elements of the complex sample stimuli. The Set 1 re-

Table 2
Steps and Trial Types Used in Set 1 Training During Phase 2

Step	Sample	Comparisons		Trials per session	Mastery criterion
1	"vowel" + A	<u>A</u>	D	12	6 consecutive correct
2	"vowel" + A + O	<u>A</u>	D	12	12 of 12
3	"vowel" + O	<u>O</u>	V	12	6 consecutive correct
4	"vowel" + A + O	<u>O</u>	V	12	12 of 12
5	"vowel" + A + O	<u>O</u>	V	12	23 of 24
	"vowel" + A + O	<u>A</u>	D	12	
6	"consonant" + D	A	<u>D</u>	12	6 consecutive correct
7	"consonant" + D + V	A	<u>D</u>	12	12 of 12
8	"consonant" + V	O	<u>V</u>	12	6 consecutive correct
9	"consonant" + D + V	O	<u>V</u>	12	12 of 12
10	"consonant" + D + V	A	<u>D</u>	12	23 of 24
	"consonant" + D + V	O	<u>V</u>	12	
11	"vowel" + A + O	<u>A</u>	D	12	46 of 48
	"vowel" + A + O	<u>O</u>	V	12	
	"consonant" + D + V	A	<u>D</u>	12	
	"consonant" + D + V	O	<u>V</u>	12	
12	Same as 11, except that only 70% of responses were reinforced				
13	Same as 11, except that only 25% of responses were reinforced				

Note. Underlined letters indicate the correct comparison stimulus.

lations pretested in Phase 1 were reassessed. The trial format was identical to that of training except that only a single response to the sample-stimulus box was required to produce the comparison stimuli, because the sample stimulus now consisted of one or zero visual elements. A correct response was defined as the selection of a comparison stimulus from the same three-element complex sample that was used during training. For example, with A as the sample, O would be the correct selection from among the comparisons O and V. Similarly, with D as the sample, V would be the correct selection from the comparisons O and V. No feedback followed any trials, but participants were given cheese puffs and white tokens after every 12 trials to keep them on task. Test-trial types were intermingled randomly throughout the session.

All stimulus relations that were tested are shown in Table 1. With one exception, test sessions presented all possible relations between stimuli within Set 1 (both vowels and consonants). It was not practical for the au-

ditory stimuli (the words "vowel" and "consonant") to serve as comparisons (i.e., A-"vowel," O-"vowel," D-"consonant," etc.), because of difficulties inherent in presenting two auditory stimuli simultaneously (see Dube, Green, & Serna, 1993, on assessing auditory stimulus classes). Test sessions included 48 trials, 24 of which employed a single visual sample stimulus and 24 of which employed a single auditory sample stimulus (see Figure 1, middle and bottom panels).

Replication of training and testing with vowels and consonants in Set 2. The 13-step identity matching training procedures used previously were replicated with new vowel stimuli (E-U-"vowel") and new consonant stimuli (K-T-"consonant"). Table 3 shows the trial configurations for Set 2 training. Following training, participants were retested with the Set 2 trials shown in Table 1.

Results

Performance during training. Accuracy was high for both Patty and Beth across all steps

Table 3
Steps and Trial Types Used in Set 2 Training during Phase 2

Step	Sample	Comparisons		Trials per session	Mastery criterion
1	"vowel" + E	E̲	K	12	6 consecutive correct
2	"vowel" + E + U	E̲	K	12	12 of 12
3	"vowel" + U	U̲	T	12	6 consecutive correct
4	"vowel" + E + U	U̲	T	12	12 of 12
5	"vowel" + E + U	U̲	T	12	23 of 24
	"vowel" + E + U	E̲	K	12	
6	"consonant" + K	E	K̲	12	6 consecutive correct
7	"consonant" + K + T	E	K̲	12	12 of 12
8	"consonant" + T	U	T̲	12	6 consecutive correct
9	"consonant" + K + T	U	T̲	12	12 of 12
10	"consonant" + K + T	E	K̲	12	23 of 24
	"consonant" + K + T	U	T̲	12	
11	"vowel" + E + U	E̲	K	12	46 of 48
	"vowel" + E + U	U̲	T	12	
	"consonant" + K + T	E̲	K̲	12	
	"consonant" + K + T	U	T̲	12	
12	Same as 11, except that only 70% of responses were reinforced				
13	Same as 11, except that only 25% of responses were reinforced				

Note. Underlined letters indicate the correct comparison stimulus.

of the training using complex sample stimuli. During training with vowels and consonants from Set 1, Patty averaged 98.4% correct (*SD* = 4.76, range, 75% to 100%). Beth did similarly well (mean percentage correct = 94.69%, *SD* = 8.11, range, 62% to 100%), but during the 11th step of training she had a series of five sessions marked by errors involving incorrect selection of the letter V (71%, 62%, 83%, 93%, 93%, and 100% accurate, respectively). On these trials, she often uttered "V, vowel," suggesting a possible conflict between training contingencies (V as consonant) and her preexperimental repertoire (V as the first letter in the word vowel). During training with vowels and consonants from Set 2, Patty averaged 99.7% correct (*SD* = 1.1, range, 96% to 100%), and Beth averaged 98.2% correct (*SD* = 3.3, range, 89% to 100%). Across the 13 training steps for both Set 1 and Set 2 stimuli, Patty always met the mastery criterion in the first session, and Beth required

an average of 1.46 sessions to meet the mastery criterion.

Performance on tests for untrained arbitrary relations. For economy of presentation, the tested relations are organized by relation type. For example, data from the test trials that presented A as the sample and O as the correct comparison were combined with data from the test trials that presented O as the sample and A as the correct comparison. Given that the stimuli in the tested relations had no history of directional training (e.g., A and O were associated only through their common association with "vowel"), there is no reason to expect differential outcomes when one stimulus served as sample and the other as comparison. Recall that auditory stimuli were presented individually as samples during arbitrary relations tests, but never as comparisons. Thus, in relation types labeled "vowel" or "consonant," the letters served only as comparisons.
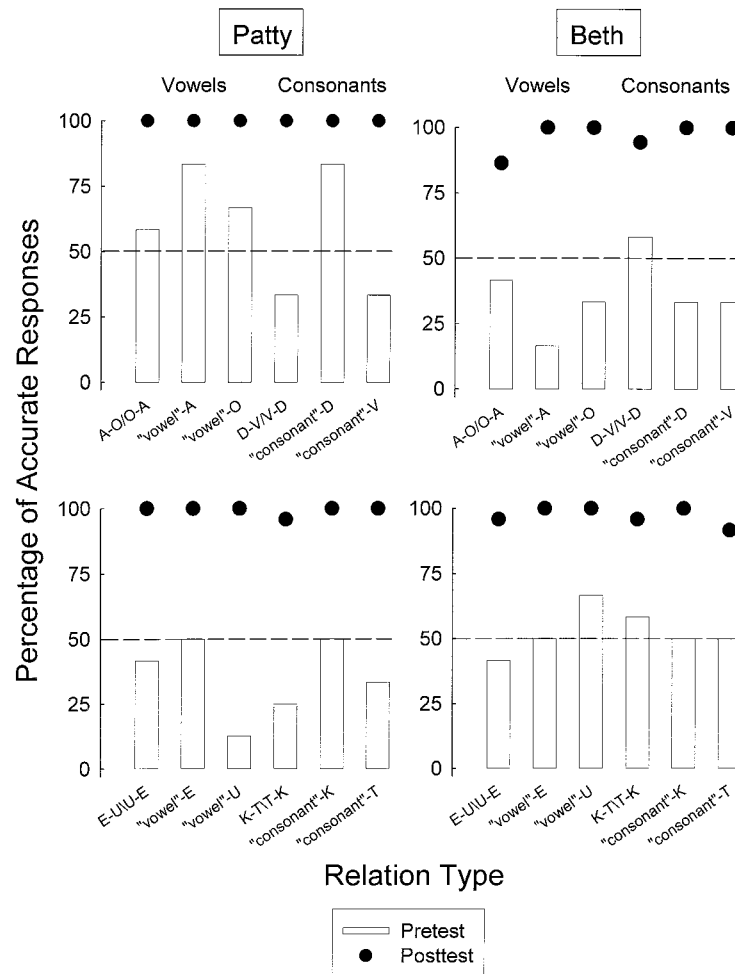
Figure 2 shows the test performances for

Figure 2. Percentage of accurate responses on the pretest (white bars) and within-set arbitrary stimulus relations tests (black circles) following training. Dashed lines at 50% indicate chance levels of performance. The letters tested are depicted under each bar. Except for relations between spoken words and letters, stimuli are grouped by both sample and comparison (e.g., A-O/O-A refers to tests when A was the sample and O was the comparison and vice versa). Spoken words (e.g., "vowel") were presented only as samples (see text for details). The panels show performances on the first set (A, O, D, V; top panels) and second set (E, U, K, T; bottom panels) of vowels and consonants for both participants.

untrained arbitrary stimulus relations among complex stimulus elements for vowels and consonants within Set 1 (A-O-"vowel," D-V-"consonant") and Set 2 (E-U-"vowel," K-T-"consonant"). Test performance was substantially better than pretest performance for all relation types. Thus, both participants demonstrated all possible arbitrary relations between elements of each complex sample. Patty made no errors across the 96 test trials,

and Beth scored above 90% on all relation types except the A-O/O-A relation (86%).

PHASE 3:
TRAINING REVIEW
AND POSTTESTING

Procedure

Training review: All vowels and consonants from Sets 1 and 2. After the arbitrary rela-

tions tests with vowels and consonants from Set 2 were completed, all of the trained relations from Sets 1 and 2 were reviewed in 48-trial sessions consisting of six trials from each trained relation. The trial format was identical to that of earlier training sessions. Accuracy feedback followed every trial until a single session was completed with at least 47 of the 48 trials correct. Feedback probability was lowered to .70, and then to .25, under the same mastery criterion. Neither participant required more than two sessions to achieve mastery or performed worse than 92% accurate in any training review session.

Posttest: Arbitrary relations indicating merger of stimulus classes. These tests assessed whether the stimulus sets with common auditory elements had merged into a single class or category. Specifically, these tests examined whether the three-member class A-O-"vowel" had merged with the three-member class E-U-"vowel" to form the five-member class A-O-E-U-"vowel" and whether the three-member class D-V-"consonant" had merged with the three-member class T-K-"consonant" to form the five-member class D-V-T-K-"consonant." No feedback followed any trials, but participants were still given cheese puffs and white chips after every 12 trials to maintain on-task behavior. All tested relations are shown in Table 1.

Sessions lasted 72 trials and included (a) 48 trials in which a single visual sample stimulus was presented, and the correct comparison was from a different stimulus set than the sample (e.g., relations A-E, A-U, D-T) and (b) 24 trials in which an auditory sample was presented without a visual accompaniment. These auditory-visual stimulus relations also had been tested in previous parts of the study ("vowel"-A, etc.). For the sake of expedience, no other emergent relations from earlier phases were retested. We focused on relations involving auditory stimuli under the assumption that any loss of previously demonstrated relations would be most apparent in cases involving those stimuli, which were not previously implicated in the experimental contingencies.

Posttest: Oral naming. These tests assessed whether the participants could name each individual letter as a vowel or consonant. The procedures used in this phase were identical to those during pretesting of oral naming of the printed letters. As discussed earlier, reliability measures were not obtained on these tests.

Posttest: Identification of letters in word context. These tests assessed whether the participants could identify the letters used in the study as vowels or consonants when they appeared in words. The stimuli were 32 four-letter words, each printed in 48-point font on a sheet of paper about 7 by 12 cm. Each word contained one vowel and one consonant from among the letters used during the study. For example, the word RUST contained the letters U and T. Other words included BITE, SEND, and DOGS. All 16 possible combinations of vowels and consonants (e.g., E, T; E, D; O, T; O, D; etc.) were presented across the 32 words. Each letter from the study was presented an equal number of times. A separate sheet of paper was used for each trial, and each sheet was presented one at a time. The serial position of the letters used in the study varied nonsystematically across the 32 words.

The 32 words were each presented twice. During the first presentation of the 32 words, participants were told to identify any vowels they saw in each word by crossing out the vowel with a red marker. During the second presentation, using the same words presented in the same order, participants were told to identify any consonants they saw in each word by crossing out the consonants with a red marker. This test assessed generalization of the arbitrary performances and served as a form of control for extraexperimental learning. Generalization of the acquired skill to a new context would be

demonstrated if participants marked the appropriate vowel and consonant letters used in training. Control would be demonstrated if participants did not mark other consonants or the other vowel (the letter I).

At a later date, two individual raters who were unaware of the details of the experiment were given the individual sheets produced by each participant. They were told which vowels and consonants were the target letters and then were instructed to score the participants' performances. Agreement among the experimenter's original scoring and the scoring of the two raters was 100%.

Follow-up. To assess the maintenance of the acquired stimulus relations between vowels and consonants that were revealed during testing, the final two posttests (oral naming and identification within words) were repeated approximately 6 weeks after the end of the study. These tests were selected for use at follow-up because, among all of the assessments of the study, they were (a) structurally least similar to the experimental training, and therefore possibly most subject to deterioration over time; and (b) most similar to classroom tasks, and therefore a good estimate of the relevance of the generality of the training to educational activities. Agreement on oral naming follow-up tests was not obtained. Agreement on the test for identification within words was obtained using the same raters and procedure noted above. Agreement among the raters and the experimenter was again 100%.

Results

Posttest: Arbitrary relations indicating merger of stimulus classes. Figure 3 shows outcomes of the tests with arbitrary stimulus relations that would emerge if Vowel Sets 1 and 2 merged into a single vowel class and Consonant Sets 1 and 2 merged into a single consonant class (e.g., A as sample and E as correct comparison, or T as sample and D as correct comparison). The performance of both participants was consistent with merger of the four sets into the two five-member stimulus classes (or categories), A-O-E-U- "vowel" and D-V-K-T-"consonant." Patty and Beth both averaged above 90% accuracy on all relation types. For economy of presentation, the data in Figure 3 were collapsed across relation types in the same manner as in Figure 2.

Posttest: Oral naming. The upper panel of Figure 4 shows performances on tests that assessed whether the match-to-sample training produced oral classification (or naming) of each of the letters as vowel or consonant. Both participants performed nearly perfectly in classifying the letters as vowel or consonant. Most aspects of this untrained naming remained at a high level of accuracy during the follow-up tests. Beth categorized all letters perfectly during the follow-up tests. Patty's accuracy decreased for E (down to 25% accurate) and U (75%), but remained at 100% for the other letters.

Posttest: Identification of letters in word context. The bottom panel of Figure 4 shows performances on tests that assessed generalization of the newly emergent skills to a novel context—identification of the letters as vowels or consonants within the context of four-letter words. Pretests were not given for these skills. During posttesting, Patty identified the letters perfectly as vowels or consonants, and Beth made errors only with regard to V (two errors in eight trials). At follow-up, Patty made only four errors in 64 trials, and Beth made no errors. Notably, neither participant marked any letters that were not used during the initial training. This observation provides assurance that behavior during these generalization tests was controlled by the training procedures rather than by other extraexperimental variables.

DISCUSSION

Stimulus equivalence paradigms are promising in part because of their generativ-
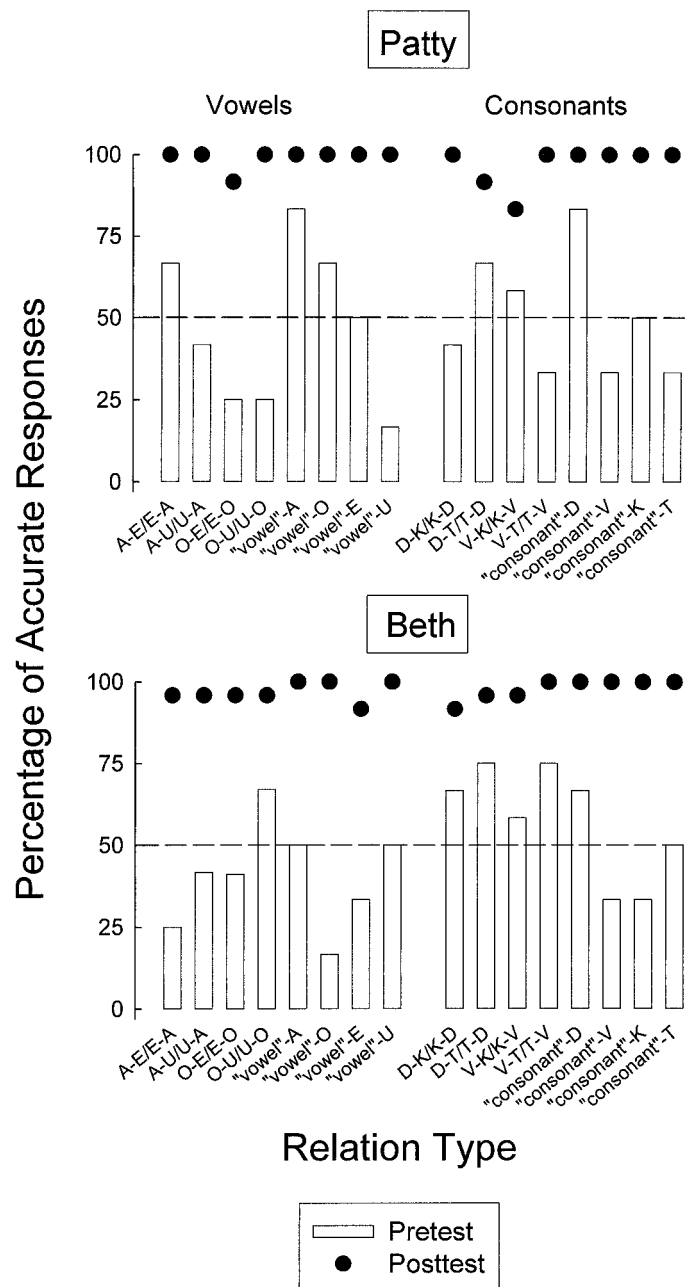
Figure 3.   Percentage of accurate responses on the pretest (white bars) and the posttest (black circles) for mergers of sets sharing the common element "vowel" or "consonant" during training. Dashed lines at 50% indicate chance levels of performance. The letters tested are depicted under each bar. Except for relations between spoken words and letters, stimuli are grouped by both sample and comparison (e.g., A-E/E-A refers to tests when A was the sample and E was the comparison and vice versa).

ity. They facilitate the predictable emergence of many untrained stimulus relations following the explicit training of only a few. By contrast, persons with developmental disabilities are sometimes characterized by a lack of generativity in their learning; that is,
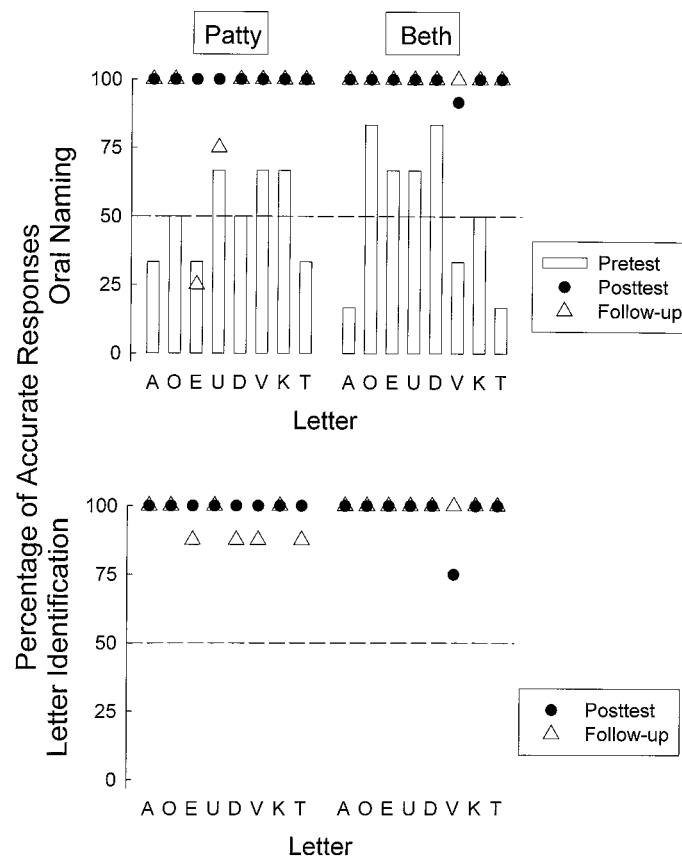
Figure 4. Percentage of accurate responses on the oral naming tests (top panel) and tests for vowel identification within words (bottom panel). In the top panel, white bars show pretest performances, black circles show the posttest results, and white triangles show performances at the 6-week follow-up. Dashed lines show chance levels. In the bottom panel, black circles show results of emergent generalization tests in which participants identified the stimuli in the context of four-letter words. White triangles show performances at the 6-week follow-up (no pretest measures were taken on these tests).

what emerges may not exceed what is explicitly trained. In this context, the present results are noteworthy in three respects.

The first contribution of the present study is that it corroborates the findings of previous studies in demonstrating that identity-based match-to-sample procedures using complex sample stimuli can rapidly instate new skills in individuals with developmental disabilities (Stromer & Mackay, 1992, 1993). Throughout training, participants rarely performed below 90% accuracy in any session, and acquisition of the trained relations often took place in the minimum time

possible within the constraints of the study's mastery criteria.

The second contribution is the demonstration of a potentially effective technique for increasing the generativity of equivalence-based paradigms. The complex sample stimuli used in training consisted of three elements, two printed letters and a spoken label (or category name) for those letters. Previous studies have shown that stimulus classes that share a common element can merge to form a larger class (Saunders, Saunders, Kirby, & Spradlin, 1988; Sidman et al., 1985), and that is what occurred in the

present study. Here, two three-member vowel classes and two three-member consonant classes each merged to form a single five-member class.

Performances on oral naming and letter identification tests provided additional evidence of the generative potency of the present procedures. Previous studies using complex stimuli have produced several types of functionally relevant generative outcomes in addition to those emerging as equivalence classes. For example, Stromer and Mackay (1992) employed an identity-based procedure using complex sample stimuli to produce emergent spelling-related performances using pictures and dictated words. In addition, their data suggest that written spelling, oral spelling, and naming may have supplementally emerged as a result of their procedures. Burke and Cerniglia (1990) reported, anecdotally, that teaching children with autism to follow instructions involving multiple stimulus attributes may have resulted in a spontaneous increase in the complexity of their expressive language. The present study focused primarily on untrained arbitrary relations expected from the identity-based training. Without additional training, Beth and Patty acquired the ability to name the letters used in the study as vowels or consonants and to identify them as vowels or consonants when they were imbedded in four-letter words. All told, by using complex sample stimuli, the training procedure produced 32 arbitrary match-to-sample performances that were not in the participants' repertoires prior to training, plus 40 additional arbitrary classifications involving non-match-to-sample, or generalized, skills.

The third contribution of the present study is that it joins only a handful of others in assessing the durability over time of learning outcomes produced by equivalence-based procedures in persons with developmental disabilities. We conducted follow-up tests of oral naming and identification of letters in words, the two performances with greatest classroom validity and least obvious connection to the relations we explicitly trained. After 6 weeks of no experimental sessions, these performances were largely maintained. Saunders, Saunders, and Spradlin (1990) reported successful 1-year follow-up results with arbitrary relations that involved abstract symbols, but ours is the first study of which we are aware to assess the maintenance of non-match-to-sample (e.g., generalized) performances that resulted from stimulus equivalence procedures.

The training package used in this study, based on the combination of several previous methodological innovations, produced arbitrary conditional stimulus relations with relative efficiency throughout training and testing. The training contingencies required the matching of identical stimulus elements, but participants may have actually learned more than simple identity relations during training. For example, the observing response requirement of touching both visual stimulus elements (letters) at the start of each trial could be considered to be a form of direct training of arbitrary relations because it required attending to the stimulus elements in similar stimulus–response relationships. In addition, the simultaneous presentation of the auditory and visual elements on each training trial suggests the possibility not only of joint control by the auditory and visual stimuli but also of directly training the auditory-visual stimulus relations. That is, the simple temporal pairing of sample stimulus elements (e.g., printed letters and words) could have directly connected them as arbitrary relations; if so, then the auditory-visual relations should be considered as directly trained rather than emergent (see Stromer, McIlvane, & Serna, 1993). These possibilities cannot be evaluated explicitly based on the present design and thus provide important questions for follow-up studies. But for applied practitioners, whether the acquired

arbitrary relations are best described as emergent or trained is less important than the finding that these performances were obtained without the separate, systematic, reinforcement-based training of each relation.

Although the results are tentative, the use of spoken words as members of the stimulus sets may have contributed to the efficacy of the present procedures and therefore is worth considering as a component of other programs designed to teach categorization. Previous studies have shown that the use of dictated names as stimuli can enhance stimulus class formation (Dugdale & Lowe, 1990; Green, Mackay, McIlvane, Saunders, & Soraci, 1990; Sidman et al., 1986). In the training portion of the present study, category words ("vowel" and "consonant") were used as the common associate that linked visual stimuli within each of the categories. Stimulus classes sharing a common auditory stimulus readily merged, despite the fact that (a) stimuli in different classes were never presented together during training and (b) attention to the spoken words was never required by the experimental contingencies. The use of spoken words in the training procedure also necessarily facilitated the emergence of untrained abilities to name each of the trained stimuli as vowels or consonants (for a thorough discourse on this topic, see Horne & Lowe, 1996; Stromer, 1996).

The testing format of the present study was important in ruling out one common source of spurious effects in the acquisition of complex stimulus control. When stimulus control is established based on complex stimuli, only limited features (elements) of the putative complex discriminative stimulus may actually control behavior. This restricted or overselective stimulus control is fairly common in persons with developmental disabilities (Allen & Fuqua, 1985; Lovaas & Schreibman, 1971; see also Reynolds, 1961). When it occurs, a casual evaluation of training effects could suggest more complex stim-

ulus relations than actually exist. In procedures like the present one, therefore, it is necessary to test each possible relation between the complex stimulus elements separately (Stromer, McIlvane, Dube, & Mackay, 1993). For example, in the present study, after training with the A-O-"vowel" complex sample, each possible combination of the sample elements was tested with each element as both sample and comparison. The same approach was used when the posttests for stimulus class mergers were conducted. Had the training produced restricted control based on the complex samples, Beth and Patty would have performed inconsistently on tests for untrained relations. For example, they might have exhibited relations between the printed letters but not between the printed letters and the spoken words. Instead, their near-perfect performance showed that each of the possible binary stimulus relations had indeed formed.

Previous stimulus equivalence studies have shown that procedures that incorporate only two response options during match-to-sample training can produce another type of spurious accuracy based on control by the negative comparison, or exclusion-based responding (Carrigan & Sidman, 1992; Johnson & Sidman, 1993). Thus, it is possible that the participants could have learned not to select the nonidentical stimuli, rather than to select identical stimuli. For example, given A-O-"vowel" as a complex sample stimulus and A and D as comparison stimuli, participants may have learned to "not choose" D, rather than learning the identity relation A-A. Such a strategy could have produced accurate responding throughout training, but it would have produced only chance-level performances on the arbitrary relations tests. This is because these tests included only stimuli that were not identical to the sample. In view of the participants' near-perfect performance on test trials that provided no basis for spurious accuracy

based on exclusion, it seems unlikely that their performance during training reflected control by negative comparison stimuli.

It is important to note that although the present study focused on the training of vowel and consonant classification skills, these procedures should be useful in teaching a variety of important academic categories. We decided to teach vowel and consonant identification after pretesting the participants with several categories of stimuli (including foods and animals). Vowels and consonants were selected because they (a) produced chance performance levels in pretesting; (b) met with parental approval as a useful skill to teach; and (c) were to be part of Beth's school curriculum later in the academic year. Although vowel and consonant recognition are not typical components of language arts programs for persons with mental retardation (whole-word approaches are more common; Buckley, 1995), the skills instated in the present study do have educational relevance. There is evidence, for example, that categorization of word sounds and word components (especially vowels) plays an important role in reading and spelling acquisition in normally developing individuals (Ehri, Wilce, & Taylor, 1987; Goswami, 1993). Consistent with this finding, it has been suggested that this approach of identifying word parts and their relationship to word sounds may be helpful in teaching children with developmental disabilities to read (Folk & Campbell, 1978; Johansson, 1993).

Whatever the specific categories of interest, training programs based on procedures like the present ones might prove to be more productive and less time intensive than some more traditional methods. Teachers could apply these methods to help students with cognitive delays learn to classify stimuli such as foods and animals. A didactic session might, for example, include providing a student with a banana and a carrot, then asking him or her to hold up the relevant stimulus follow-ing the simultaneous presentation of either a banana, an apple, and the spoken word "fruit," or a carrot, a potato, and the spoken word "vegetable." In subsequent sessions, the teacher could ask students to orally label the individual stimuli as fruit or vegetable, or given an array of stimuli, to match those that belong together. Subsequent teaching could emphasize analogous relations with new stimuli (e.g., orange-pear-"fruit," or lettuce-turnip-"vegetable"). The present results would predict spontaneous merger of all fruit and vegetable stimuli into a single large class. In theory, systematic application of such a teaching strategy could promote the exponential growth of stimulus classes. Each additional trained set of stimuli would, in theory, merge into the existing class. The procedures used in this study also might be applied to providing adults with functional, work-related skills such as classifying monetary amounts of equal value (e.g., "one dollar," a one-dollar bill, four quarters, ten dimes, etc.); identifying machine parts of various sizes, like screws and bolts; or grouping tools, such as different forms of wrenches and screwdrivers.

Laboratory procedures have only recently been devised to explore the acquisition of arbitrary relations in identity-based training and the merger of stimulus classes based on common members. Much additional research may be needed to determine the best way to integrate these phenomena into broadly effective teaching techniques that are easy to apply in classroom settings. The present study might be viewed as a helpful step in the technology transfer process. Such a process could result in teaching techniques that benefit both persons with developmental disabilities and service providers who face practical constraints that limit the often-extensive teaching time necessary to build complex repertoires.

REFERENCES

Allen, K. D., & Fuqua, R. W. (1985). Eliminating selective stimulus control: A comparison of two

procedures for teaching mentally retarded children to respond to compound stimuli. *Journal of Experimental Child Psychology, 39,* 55–71.

Baron, A., & Menich, S. R. (1985). Reaction times of older and younger: Effects of compound samples and a prechoice signal on delayed matching-to-sample performances. *Journal of the Experimental Analysis of Behavior, 44,* 1–14.

Barsalou, L. W. (1992). *Cognitive psychology: An overview for cognitive scientists.* Hillsdale, NJ: Erlbaum.

Buckley, S. (1995). Teaching children with Down syndrome to read and write. In L. Nadel & D. Rosenthal (Eds.), *Down syndrome: Living and learning in the community* (pp. 158–169). New York: Wiley-Liss.

Burke, J. C., & Cerniglia, L. (1990). Stimulus complexity and autistic children's responsivity: Assessing and training a pivotal behavior. *Journal of Autism and Developmental Disorders, 20,* 233–253.

Carrigan, P. F., Jr., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58,* 459–504.

Dube, W. V., Green, G., & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus classes. *Journal of the Experimental Analysis of Behavior, 59,* 103–114.

Dugdale, N. A., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 115–138). Hillsdale, NJ: Erlbaum.

Ehri, L. C., Wilce, L. S., & Taylor, B. B. (1987). Children's categorization of short vowels in words and the influence of spellings. *Merrill-Palmer Quarterly, 33,* 393–421.

Folk, M. C., & Campbell, J. (1978). Teaching functional reading to the TMR. *Education and Training of the Mentally Retarded, 13,* 322–326.

Goswami, U. (1993). Toward an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology, 56,* 443–475.

Green, G., Mackay, H. A., McIlvane, W. J., Saunders, R. R., & Soraci, S. A. (1990). Perspectives on relational learning in mental retardation. *American Journal on Mental Retardation, 95,* 249–259.

Halle, J. W., & Holt, B. (1991). Assessing stimulus control in natural settings: An analysis of stimuli that acquire control during training. *Journal of Applied Behavior Analysis, 24,* 579–589.

Harris, F. C., & Lahey, B. B. (1982). Recording system bias in direct observational methodology. *Clinical Psychology Review, 2,* 539–556.

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

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 185–242.

Johansson, I. (1993). Teaching prereading skills to disabled children. 9th World Conference of the International Association for the Scientific Study of Mental Deficiency. *Journal of Intellectual Disability Research, 37,* 413–417.

Johnson, C., & Sidman, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59,* 333–347.

Kennedy, G., & Serna, L. A. (1995). Emergent relational responding based upon quantity and equivalence. *The Psychological Record, 45,* 241–260.

Lovaas, O. I., & Schreibman, L. (1971). Stimulus overselectivity of autistic children in a two stimulus situation. *Behavior Research and Therapy, 9,* 305–310.

Mackay, H. A. (1985). Stimulus equivalence in rudimentary reading and spelling. *Analysis and Intervention in Developmental Disabilities, 5,* 373–387.

Maguire, R. W., Stromer, R., Mackay, H. A., & Demis, C. A. (1994). Matching to complex samples in adults with autism and young children. *Journal of Autism and Developmental Disabilities, 24,* 753–772.

Markham, M. R., & Dougher, M. J. (1993). Compound stimuli in emergent stimulus relations: Extending the scope of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 60,* 529–542.

Reid, J. B. (1982). Observer training in naturalistic research. In D. P. Hartman (Ed.), *Using observers to study behavior* (pp. 37–50). San Francisco: Jossey-Bass.

Reynolds, G. S. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior, 4,* 203–208.

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., & Spradlin, J. E. (1990). Long-term stability of equivalence relations in the absence of training or practice. *American Journal on Mental Retardation, 95,* 297–303.

Saunders, R. R., Wachter, J., & Spradlin, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior, 49,* 95–115.

Schenk, J. J. (1993). Emergent conditional discrimination in children: Matching to compound stimuli. *Quarterly Journal of Experimental Psychology, 46B(4),* 345–365.

Schenk, J. J. (1995). Complex stimuli in nonreinforced simple and conditional discriminations. *The Psychological Record, 45,* 477–494.

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

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

Sidman, M., & Kirk, B. (1974). Letter reversals in naming, writing, and matching to sample. *Child Development, 45,* 616–625.

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.

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

Smeets, P. M., & Striefel, S. (1994). Matching to complex stimuli under non-reinforced conditions: Errorless transfer from identity to arbitrary matching tasks. *Quarterly Journal of Experimental Psychology, 47B,* 39–62.

Stoddard, L. T., Brown, J., Hurlbert, B., Manoli, C., & McIlvane, W. J. (1989). Teaching money skills through stimulus class formation, exclusion, and component matching methods: Three case studies. *Research in Developmental Disabilities, 10,* 413–439.

Stromer, R. (1996). On the experimental analysis of naming and the formation of stimulus classes. *Journal of the Experimental Analysis of Behavior, 65,* 250–251.

Stromer, R., & Mackay, H. A. (1992). Spelling and emergent picture-printed word relations established with delayed identity matching to complex samples. *Journal of Applied Behavior Analysis, 25,* 893–904.

Stromer, R., & Mackay, H. A. (1993). Delayed identity matching to complex samples: Teaching students with mental retardation spelling and the prerequisites for equivalence classes. *Research in Developmental Disabilities, 14,* 19–38.

Stromer, R., McIlvane, W. J., Dube, W. V., & Mackay, H. A. (1993). Assessing control by elements of complex stimuli in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 59,* 83–102.

Stromer, R., McIlvane, W. J., & Serna, R. W. (1993). Complex stimulus control and equivalence. *The Psychological Record, 43,* 585–598.

Stromer, R., & Stromer, J. B. (1990a). The formation of arbitrary stimulus classes in matching to complex samples. *The Psychological Record, 40,* 51–66.

Stromer, R., & Stromer, J. B. (1990b). Matching to complex samples: Further study of arbitrary stimulus classes. *The Psychological Record, 40,* 505–516.

STUDY QUESTIONS

1. What is the potential advantage of using stimulus equivalence procedures to teach new skills?

2. Illustrate the use of an identity-based variation of a multielement delayed match-to-sample procedure using the numbers 1, 2, 3, and 4 as stimuli.

3. What was the purpose of the current study?

4. What were the consequences of responding during training and during testing? What consequences were arranged to increase the likelihood of continued participation?

5. What basic responses were taught during the tutorial phase?

6. Briefly describe the various posttests that were conducted in Phase 3 and the results that were obtained.

7. Identify two potential sources of spurious effects in the acquisition of complex stimulus control and describe how the authors controlled for these two sources in the present study.

8. How might the methodology used in this study be extended to facilitate mergers in other instructional contexts?

Questions prepared by Gregory Hanley and Michele Wallace, The University of Florida