REINFORCER FREQUENCY AND RESTRICTED STIMULUS CONTROL

WILLIAM V. DUBE AND WILLIAM J. MCILVANE

E. K. SHRIVER CENTER FOR MENTAL RETARDATION
AND NORTHEASTERN UNIVERSITY

Stimulus control was evaluated in 3 individuals with moderate to severe mental retardation by delayed identity matching-to-sample procedures that presented either one or two discrete forms as sample stimuli on each trial. On pretests, accuracy scores on one-sample trials were uniformly high. On two-sample trials, the correct stimulus (i.e., the one that subsequently appeared in the comparison array) varied unpredictably, and accuracy scores were substantially lower, suggesting that both sample stimuli did not exert stimulus control on every trial. Subjects were then given training sessions with the one-sample task and with a new set of four stimuli. For two of the stimuli, correct matching responses were followed by reinforcers on a variable-ratio schedule that led to a high reinforcer rate. For the other two stimuli, correct responses were followed by reinforcers on a variable-ratio schedule that led to a substantially lower reinforcer rate. Results on two-sample tests that followed showed that (a) on trials in which comparison arrays consisted of one high reinforcer-rate and one low reinforcer-rate stimulus, subjects most often selected the high-rate stimulus; and (b) on trials in which the comparison arrays were either two high reinforcer-rate stimuli or two low reinforcer-rate stimuli and the samples were one high reinforcer- and one low reinforcer-rate stimulus, accuracy was higher on trials with the high-rate comparisons. These results indicate that the frequency of stimulus control by high reinforcer-rate samples was greater than that by low reinforcer-rate samples. Following more training with the one-sample task and reversed reinforcement schedules for all stimuli, the differences in stimulus control frequencies on two-sample tests also reversed. These results demonstrate experimental control by reinforcement contingencies of which of two sample stimuli controlled selections in the two-sample task. The procedures and results may prove to be relevant for understanding restricted stimulus control and stimulus overselectivity.

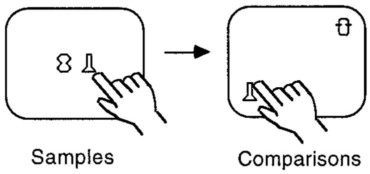
Key words: restricted stimulus control, stimulus overselectivity, reinforcer rate, matching to sample, pointing, humans with mental retardation

Discrimination training may result in stimulus control by only a subset of the potential controlling stimuli (e.g., Reynolds, 1961). When the number of controlling stimuli is atypically limited, the outcome has been termed *restricted stimulus control* (Litrownik, McInnis, Wetzel-Pritchard, & Filipelli, 1978). Restricted stimulus control is often observed in behavioral analyses of individuals with intellectual disabilities, and it is a continuing problem in their education and training (e.g., Allen & Fuqua, 1985; Bickel, Richmond, Bell, & Brown, 1986; Dunlap, Koegel, & Burke, 1981).

The present experiment follows from the studies of restricted stimulus control in individuals with intellectual disabilities described in the stimulus overselectivity literature (reviewed by Lovaas, Koegel, & Schreibman, 1979). The methods typically involved initial discrimination training with complex or multiple stimuli followed by testing with the individual elements or stimuli to determine how many of them controlled the target behavior (usually touching a response key or other manipulandum). For example, if initial training established the stimuli ABC, presented together, as positive and XYZ as negative, tests would present the various individual combinations A versus Y, B versus X, and so forth. When given such tests, individuals with developmental disabilities may respond appropriately to some stimuli but not others, and to fewer stimuli than nondisabled individuals. For example, with three-stimulus displays like those in the example above, Wil-

Data collection and manuscript preparation were supported by NICHD Grant HD 25995. The data were presented at the 1996 APA conference in Toronto, Ontario. For their help with implementation, data collection, or data analysis, we thank Lyn Balsamo, Kevin Farren, Paul Filkins, Fay Iennaco, Jason Langlois, Alison McVay, Jennifer Mlocek, Aimee Smith, and Susan Stanford. Thanks to Lynn Doucette for help with figure preparation, and to Richard Serna and two anonymous reviewers for very helpful comments on the manuscript. We also thank The Learning Center and The New England Center for Children (formerly New England Center for Autism).

Address correspondence to W. V. Dube, Behavioral Sciences Division, E. K. Shriver Center, 200 Trapelo Road, Waltham, Massachusetts 02254 (E-mail: wdube@shriver.org).



Samples	Comparison Displays			
AB	[A,X]	[A,Y]	[B,X]	[B,Y]
AX	[A,B]	[A,Y]	[X,B]	[X,Y]
AY	[A,B]	[A,X]	[Y,B]	[Y,X]
BX	[B,A]	[B,Y]	[X,A]	[X,Y]
BY	[B,A]	[B,X]	[Y,A]	[Y,X]
XY	[X,A]	[X,B]	[Y,A]	[Y,B]

Fig. 1. Upper portion: Delayed matching to sample with multiple sample stimuli. Lower portion: CSS trial types for a set of four stimuli, A, B, X, and Y. Sample stimuli are listed in the left column, and the four sets of comparison stimuli for each sample are shown in square brackets in the right column. Within each set of brackets, the comparison stimulus on the left is correct. In the actual trial displays, all stimuli appeared equally often in all sample and comparison positions (e.g., the top row represents Samples AB and BA).

helm and Lovaas (1976) reported reliable stimulus control by all three positive stimuli in typically developing children, two positive stimuli in children with moderate mental retardation, and only one or two positive stimuli in children with severe retardation (*M* = 1.6). Restricted stimulus control has been documented with multiple stimuli from the same stimulus dimension (e.g., arrays of discrete forms; Koegel & Wilhelm, 1973; Wilhelm & Lovaas, 1976) and with different dimensions (differences in color, form, etc.; Kovattana & Kraemer, 1974). The finding has been replicated in contexts that verified discrimination of all stimuli presented individually (Dube, Kledaras, Iennaco, Stoddard, & McIlvane, 1990) and discrimination of individual stimuli when presented in multiple-stimulus arrays (Stromer, McIlvane, Dube, & Mackay, 1993).

In the present experiment, restricted stimulus control was evaluated with a delayed matching-to-sample procedure with multiple sample stimuli, illustrated in the upper portion of Figure 1. Experimental stimuli were abstract forms. On each trial, two sample stimuli were presented side by side. The samples remained available for observation until the subject touched the sample display area, and then they disappeared and the comparison stimuli were presented immediately (0-s delay). The comparisons were two individual stimuli, one of which was identical to one of the sample stimuli. Touching the identical comparison was a correct response. This task was termed *CSS* by Cox and D'Amato (1982), where the first letter (C) indicates a complex (two-stimulus) sample and the second and third letters (SS) indicate two single comparison stimuli.

The lower portion of Figure 1 shows stimulus combinations for a CSS task with a set of four stimuli, designated A, B, X, and Y. During the sample observation period, the subject cannot predict which one of the sample stimuli will appear as the correct comparison to select; every stimulus is correct equally often. Maximizing reinforcement, therefore, involves prerequisite behavior that is sufficient for stimulus control by sample-comparison identity with both of the sample stimuli; for example, the subject must observe both samples before they are removed from the display. When CSS matching is highly accurate (scores near 100%), one can infer that such behavior has occurred on nearly every trial. By contrast, when CSS accuracy scores are at or near chance levels (50%), that result may indicate missing behavioral prerequisites that are necessary to establish identity control by any samples.

Of interest for the present experiment are intermediate CSS accuracy scores of approximately 75%. In a two-sample two-comparison procedure like the one we used, such scores could result from averaging together responses based on sample-comparison identity on approximately half of the trials and responses under other forms of stimulus control on the remaining trials (Sidman, 1980). There are a number of stimulus control topographies that could result in such intermediate accuracy scores.

One possibility is prerequisite behavior sufficient for identity control by both sample stimuli on approximately half of the trials in the session and a general failure of identity control by any sample stimuli on the remaining trials (e.g., because of failure to observe either of the sample stimuli). The subject's responses would be scored as correct on all of the former trials and on half of the latter trials, resulting in the overall 75% score. A general failure of the prerequisite behavior for identity control at some point during the session would produce intermediate accuracy scores regardless of the specific sample stimuli or the number of sample stimuli presented. For example, accuracy scores would be intermediate on trials with either one or two samples. Although such performance is certainly relevant to discrimination learning problems encountered in clinical populations, it also seems to be distinct from the problem of restricted stimulus control because it is independent of stimulus complexity. For this reason, the subjects selected for the present experiment were individuals with mental retardation who had very high pretest scores on trials with one sample stimulus and intermediate scores on trials with two samples.

The remaining stimulus control possibilities for intermediate CSS accuracy scores involve identity control by some subset of the experimental stimuli: (a) identity control by two specific stimuli and no identity control by the other two stimuli, for example, reliable identity control by Stimuli A and B and some other form of stimulus control when A or B do not appear as both sample and comparison (e.g., control by comparison locations); (b) idiosyncratic control by three specific combinations of sample stimuli only, for example, perfect identity control by Samples AB, AX, and AY (trial types shown in the top three rows of Figure 1) but no identity control by BX, BY, or XY (bottom three rows of Figure 1; cf. Markham & Dougher, 1993; Stromer, McIlvane, & Serna, 1993); or (c) prerequisite behavior sufficient to establish identity control by only one sample stimulus on each trial. For example, if the subject were to observe only one of the two samples on each trial, then on those trials in which the observed stimulus also appeared in the comparison array (half of the total trials), accuracy would be very high. On the remaining trials, when the observed stimulus did not appear as a comparison, accuracy would be at chance levels, and the overall accuracy score would be intermediate. A wide range of stimulus features could conceivably control such selective observing behavior; examples include the left (or right) sample position, the sample position on the same (or opposite) side as the previous comparison selection, relative features of stimulus structure such as greater (or lesser) height, width, pixel density, or degree of symmetry, and so forth (cf. Bickel, Stella, & Etzel, 1984). Distinguishing among these possibilities may be complicated if the contingencies of reinforcement maintain multiple stimulus control topographies or if there are shifts from one form of stimulus control to another during sessions. Because these three possibilities (a, b, and c) all involve stimulus control by subsets of the experimental stimuli, they all seem to be aptly characterized as restricted stimulus control and are potentially relevant to the problem that Lovaas et al. (1979) described as stimulus overselectivity.

Although the occurrence of restricted stimulus control has been more than amply documented, its controlling variables have been analyzed only incompletely. The present experiment examined the relationship between reinforcer rate and restricted stimulus control. With subjects whose CSS accuracy scores were intermediate, we asked whether control by specific subsets of the experimental stimuli could be predicted from known differences in reinforcement history. Further, we asked whether changes in reinforcement contingencies would be reflected in corresponding changes in the specific stimuli that controlled responding.

METHOD

Subjects

Subjects were 3 individuals with moderate to severe mental retardation. Their chronological ages (and Peabody Picture Vocabulary Test mental age-equivalent scores in parentheses) were: Subject KRK, 17 years (3 years 3 months); Subject JPW, 20 years (3 years 2 months); and Subject PAW, 17 years (6 years 3 months). They were selected for participation on the basis of preliminary tests (see below).

Apparatus and Setting

Subjects sat before a Macintosh Plus® computer with a 9-in. black-and-white video display and a touch-sensitive screen. Sample stimuli were displayed in the center of the

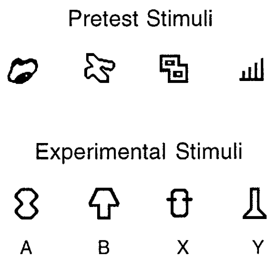


Fig. 2. Pretest and experimental stimuli. Stimulus designations A, B, X, and Y did not appear on the computer displays.

screen, and comparison stimuli were displayed in any two of the four corners of the screen. The stimuli were black forms approximately 1 cm by 1.5 cm, displayed on a white background; the stimuli are shown in Figure 2. The computer controlled all experimental events and data collection, with one exception: Tokens following correct responses were presented by an experimenter who was seated behind and to one side of the subject. Experimental sessions of approximately 10 min duration were conducted 3 or 4 days per week in a small, quiet room at the subjects' school.

Preliminary Training and Testing

Reinforcer function test. Subjects were taught to exchange plastic poker-chip tokens for snack foods or soft drinks. Then, they were given one or more sessions in which the contingency between pressing a button and receiving a token was manipulated on a mixed schedule. Reinforcer function was demonstrated in all subjects by increases in the button-pressing rate when presses produced tokens and subsequent decreases when pressing no longer produced tokens. Training and evaluation required three to five sessions.

Delayed identity-matching pretests. Two types of delayed identity-matching tasks were presented. The pretest stimuli are shown in the top portion of Figure 2. Subjects received six to nine pretest sessions.

The SSS task presented one sample stimulus on each trial; in the task designation, the first letter (S) indicates a single-stimulus sample, and the second and third letters (SS) indicate two single comparison stimuli (after Cox & D'Amato, 1982). Trials began when a sample stimulus appeared in the center of the screen. When the subject touched the sample, it disappeared and two comparison stimuli appeared immediately (0-s delay) in two corners of the screen. One comparison was identical to the sample, and the other was nonidentical. Touching the identical comparison was defined as a correct response and was followed by presentation of a brief auditory-visual computer display and a token. Touching the nonidentical comparison was an error, and it was followed only by a 5-s intertrial interval (ITI) with a blank display screen. A session consisted of 48 trials. Over trials, all stimuli appeared equally often as samples, correct and incorrect comparisons, and in each comparison position.

The CSS task, described above and shown in Figure 1, was the same as the SSS task with one exception: The sample consisted of two stimuli displayed side by side, 1.75 cm center to center. Each stimulus appeared equally often in the left and right sample positions. The comparisons were single stimuli, and the correct comparison was identical to one of the samples. All sessions that included CSS trials began with a block of 24 SSS trials followed by a block of 24 CSS trials.

Experimental Procedure

Subjects received alternating baseline and test conditions. Subjects KRK and JPW each received 39 experimental sessions, and PAW received 83 experimental sessions. The experimental stimuli are shown in the lower portion of Figure 2. In baseline conditions, correct SSS matching with two of the stimuli had a relatively high probability of reinforcement, and correct matching with the other two stimuli had a relatively low probability of reinforcement. Then, CSS trials in test conditions were used to evaluate control on trials with two samples. Table 1 shows the programmed schedules of reinforcement for identity matching in each experimental condition.

Baseline 1. Subjects received 12 sessions of 48 SSS trials each. The reinforcement schedule was continuous (CRF) in the first session, and then intermittent reinforcement was introduced gradually in the second session. Thereafter, for Subjects KRK and JPW, the reinforcement schedule for correct responses when Stimulus A or B was the sample was variable-ratio (VR) 1.3 (three of four correct re-

Table 1

Experimental conditions and programmed variable-ratio (VR) and continuous (CRF) reinforcement schedules for identity matching with Stimuli A, B, X, and Y.

Subject		SSS trials		CSS trials
		A,B	X,Y	A,B,X,Y
KRK and JPW	Baseline 1	VR 1.3	VR 4	
	Test 1	VR 1.3	VR 4	CRF
	Baseline 2	VR 4	VR 1.3	
	Test 2	VR 4	VR 1.3	CRF
PAW	Baseline 1	VR 4	VR 1.3	
	Test 1	VR 4	VR 1.3	CRF
	Baseline 1b	VR 8	VR 1.14	
	Test 1b	VR 8	VR 1.14	CRF
	Baseline 2	VR 1.14	VR 8	
	Test 2	VR 1.14	VR 8	CRF
	Baseline 2b	VR 1.04	VR 24	
	Test 2b	VR 1.04	VR 24	CRF

sponses followed by a token), and the schedule for correct matching with Sample X or Y was VR 4 (one of four correct responses followed by a token). For Subject PAW, the schedule for Stimuli A and B was VR 4 and the schedule for X and Y was VR 1.3. Subject PAW received six additional Baseline 1 sessions because of a school transfer and relocation to a different testing room after his first six sessions.

Test 1. Subjects received six test sessions. Each session began with a block of 16 SSS trials with reinforcement schedules as in Baseline 1. The SSS trials were followed immediately by a block of 24 CSS trials. Throughout the experiment, CSS tests were always conducted with CRF.

Baseline 1b and Test 1b (Subject PAW only). Baseline (10 sessions only) and Test 1 conditions were repeated with Subject PAW. On SSS trials, the schedule for Stimuli A and B was VR 8 and the schedule for X and Y was VR 1.14 (seven of eight correct responses followed by a token).

Baseline 2. Subjects received 15 sessions of SSS trials with the reinforcement schedules reversed relative to Baseline 1. For Subjects KRK and JPW, the schedule for correct responses to Stimuli A and B was VR 4, and the schedule for correct responses to Stimuli X and Y was VR 1.3. For Subject PAW, the schedule for Stimuli A and B was VR 1.14 and the schedule for X and Y was VR 8.

Test 2. This test repeated Test 1, but each session began with 16 SSS trials with the reinforcement schedules as in Baseline 2. The CSS trials in Test 2 were exactly the same as those in Test 1.

Baseline 2b and Test 2b (Subject PAW only). The reversed Baseline 2 and Test 2 conditions were repeated with Subject PAW. On SSS trials, the schedule for Stimuli A and B was VR 1.04 (23 of 24 correct responses were followed by a token) and the schedule for X and Y was VR 24.

RESULTS

Preliminary SSS and CSS Tests

In sessions that presented SSS trials only, all subjects' accuracy scores were 98% to 100% with continuous reinforcement and with intermittent reinforcement on a VR 2 schedule. The leftmost sets of bars in Figure 3 show the results of pretest sessions that presented both SSS and CSS trials. All subjects had SSS accuracy scores near 100% and CSS scores of approximately 75%. As noted in the introduction, this combination of highly accurate delayed identity matching with single sample stimuli and intermediate scores with two samples is consistent with restricted stimulus control. Subjects were selected for the present experiment on the basis of these results.

Accuracy Scores

Figure 3 shows mean accuracy scores for each condition. For Subjects KRK and JPW, the high accuracy scores for SSS trials during

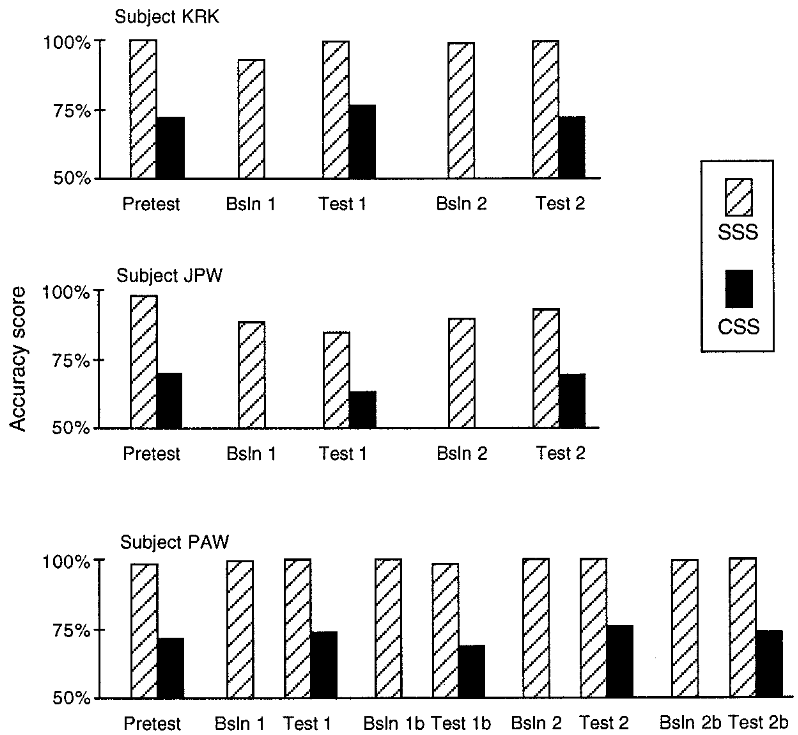


Fig. 3. Mean accuracy scores in pretests and each experimental condition for SSS trials (striped bars) and CSS trials (black bars).

pretests declined somewhat in Baseline 1 when reinforcement schedules were manipulated. The SSS scores remained well above chance levels for all subjects, however. Most baseline errors occurred on trials in which the comparison array presented a choice between stimuli that had different schedules of reinforcement and the correct comparison was the one with the leaner schedule. For example, Subject JPW made a total of 68 errors in Baseline 1 after intermittent reinforcement was introduced, and 50 of these were selections of Comparison Stimulus A or B (VR 1.3) on trials in which the correct comparison was X or Y (VR 4). Accuracy scores for all CSS tests remained at intermediate levels throughout the experiment.

Baseline Reinforcement Data

Figure 4 shows obtained reinforcer rates on SSS trials averaged over the last three sessions of each baseline condition. Rates were calculated by dividing the total number of reinforcers obtained for correct selections of each stimulus by the session duration. These data confirm that reinforcer frequencies were higher for the stimuli with the richer schedule. For convenience in presenting the rest of the results, stimuli will be referred to as high rate or low rate, terms that refer to reinforcer rate on baseline trials. Presentation rates were approximately equal for all stimuli.

Obtained reinforcer ratios were consistent with what would be expected from the programmed VR schedules. In the last three sessions of each baseline condition, the only deviation from programmed ratios greater than 1% was for Subject JPW in Baseline 1, when 25% of the reinforcers would be expected for responses to X and Y, but only 19% were obtained. This discrepancy was due to missed reinforcers because of errors on trials in which the correct comparison stimulus was X or Y (low rate) and the incorrect comparison was A or B (high rate).

CSS Tests

The CSS tests were conducted to determine the effects of the different reinforcer rates on stimulus control. Two related issues

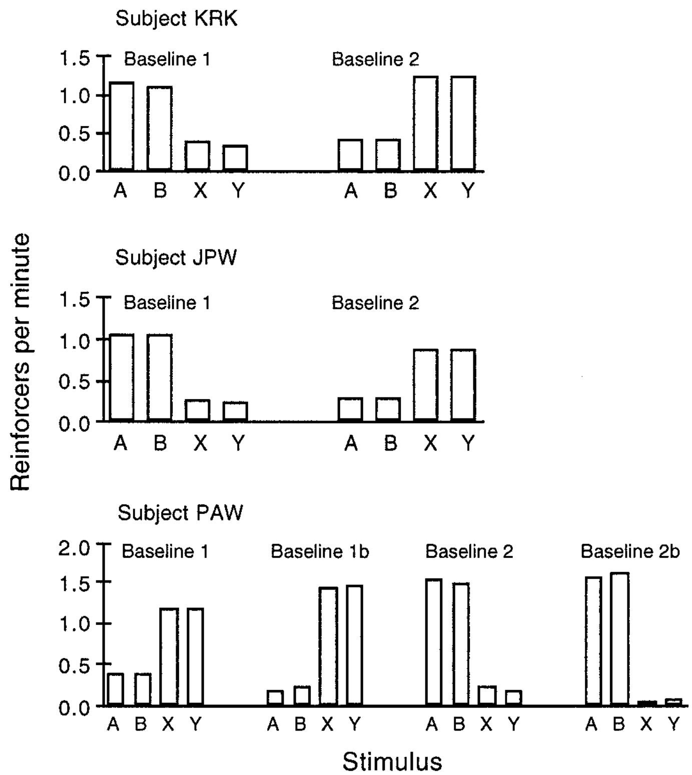


Fig. 4. Average obtained reinforcer rates for responses to Stimuli A, B, X, and Y on SSS trials during the last three sessions of each baseline condition.

bear on the choice of a technique for presenting the data. One issue concerns the heterogeneous nature of the comparison displays. The top portion of Figure 5 shows the 24 CSS trial types that were presented in the tests. Each trial type was presented six times per test, for a total of 144 CSS trials per test (24 trials per sessions for six sessions). The lower portions of Figure 5 show the six categories of CSS trial types that can be derived by grouping the stimuli according to high and low reinforcer rates on the baseline SSS trials that preceded the tests. For example, the upper left category (HH $\rightarrow$ H vs. L) is two high-rate sample stimuli, one high-rate comparison stimulus, and one low-rate comparison; in Test 1, this category included the trial types shown in the top row for Subjects KRK and JPW and the trial types in the bottom row for Subject PAW (samples were both VR 1.3 on baseline trials, one comparison was VR 1.3, and one comparison was VR 4).

The three trial-type categories in the center of Figure 5 (labeled "Different Comparison History") required the subject to choose between one high-rate and one low-rate comparison stimulus. These trial types are shown in the upper portion of Figure 5 by unshaded comparison displays. On these trials, the unequal rates of reinforcement on the preceding SSS trials may be expected to result in a preference for the high-rate comparison (e.g., Hartl & Fantino, 1996; Wixted, 1989). In contrast, the two test-trial categories in the lower portion of Figure 5 (labeled "Similar Comparison History") required a choice between either two high-rate or two low-rate comparison stimuli. These trial types are shown in the upper portion of Figure 5 by shaded comparison displays. On these trials, the reinforcer rates on the preceding SSS trials provided no basis for a comparison preference; the comparison stimuli were either both high rate or both low rate. Thus, one source of stimulus control that could compete with identity control on different-history trials was not present on similar-history trials. Because of this difference in the potential for

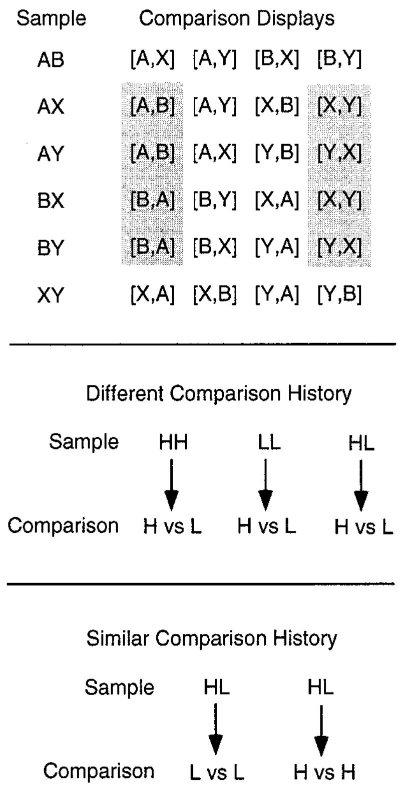


Fig. 5. Upper portion: CSS trial types. Sample stimuli are listed in the left column, and the four sets of comparison stimuli for each sample are shown in square brackets in the right column. Within each set of brackets, the comparison stimulus on the left is correct. In the actual trial displays, all stimuli appeared equally often in all sample and comparison positions. Shaded areas show test trials with similar comparison histories. Lower portion: CSS test-trial categories derived by grouping stimuli according to high (H) or low (L) rates of reinforcement on the SSS baseline trials that preceded the CSS tests.

stimulus control to be exerted by the comparison arrays, the CSS data will be presented separately for different- and similar-history trials.

The second issue for CSS data presentation is an appropriate way to measure the frequency of identity control by each of the experimental stimuli. Conditional stimulus control by individual stimuli cannot be determined merely by calculating separate accuracy scores for each sample stimulus because conditionality is lost if positive and negative stimulus functions are separated. For example, a subject who selected Stimulus A every time it was displayed as a comparison, regardless of the sample, would have a 100% accuracy score for trials with Sample A. It would be misleading, however, to conclude that such a score indicates perfect identity control; it may indicate no control at all by Sample A, and, instead, nonconditional control by Comparison A (Sidman, 1980). An analysis of identity control by individual stimuli must also take into consideration the frequency of nonconditional control by comparison stimuli.

Sidman (1992) used signal-detection space to present identity matching-to-sample data in a way that showed both conditional control by sample stimuli and competing control by comparison arrays. The signal-detection space is shown in Figure 6. In the present application to CSS identity matching, hit rates and false-alarm rates were calculated separately for each of the four stimuli. The hit rate, plotted on the ordinate, is the percentage of selections of the stimulus on trials in which it was the correct comparison; that is, it had also appeared as one of the sample stimuli. The false-alarm rate, plotted on the abscissa, is the percentage of selections of the stimulus on trials in which it was the incorrect comparison; that is, it had not appeared as one of the samples.

Performances indicative of CSS identity control, which is both a high hit rate and a low false-alarm rate, are plotted in the upper left corner of the signal-detection space (see Figure 6). The shaded area in this corner indicates performances equivalent to 90% or greater conditional discrimination accuracy scores. For example, the lower edge of the gray area touches the ordinate at a point at which the hit rate is 0.8 and the false-alarm rate is 0. The hit rate indicates 80% correct on trials in which the stimulus under consideration was the correct comparison, and the false-alarm rate of 0 indicates 100% correct on an equal number of trials in which the stimulus was the incorrect comparison; together, these rates are equivalent to a conditional discrimination accuracy score of 90%. Greater distance from the upper left corner indicates lower frequency of identity control. Performances plotted along the major diagonal (solid line in Figure 6) indicate a complete failure of identity control by the stimulus as a sample. Performances plotted along the minor diagonal (dashed line in Figure 6) indicate no bias for the stimulus as a com-

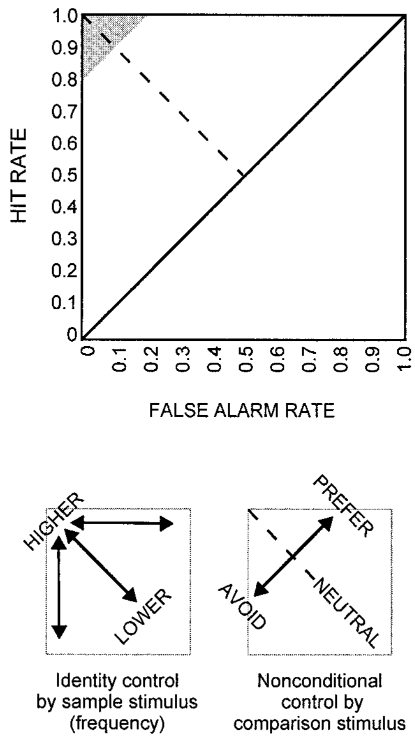


Fig. 6. Signal-detection space. Data are plotted for individual stimuli. The hit rate on the ordinate is the percentage of selections of the stimulus when it was the correct comparison. The false-alarm rate on the abscissa is the percentage of selections of the stimulus when it was the incorrect comparison. The shaded area in the upper left corner indicates identity control equivalent to 90% or greater conditional discrimination accuracy scores (see text for details). Greater distance from the upper left corner indicates lower frequency of identity control, as illustrated in the lower left portion. Performances plotted along the solid-line major diagonal indicate a complete failure of identity control. Performances plotted along the dashed-line minor diagonal indicate no bias for the stimulus as a comparison. Deviations from the minor diagonal indicate preference for or avoidance of the comparison stimulus, as illustrated in the lower right portion.

parison. Deviations above or below the minor diagonal indicate preference for or avoidance of the stimulus, respectively, as a comparison.

Different comparison history trials. Each CSS test included a total of 96 different-history trials (16 trials per session for six sessions). The left portion of Figure 7 shows initial performances with each stimulus plotted in signal-detection space. There was evidence for identity control at approximately 90% accuracy in one instance only: Stimulus Y for KRK. Identity control was lower in all other cases. All subjects had preferences for the high-rate comparison stimuli, as shown by plotted locations above the minor diagonal in Test 1 for KRK and JPW and in Test 1b for PAW.

The right portion of Figure 7 shows the results of Tests 2 and 2b, which followed baseline conditions with reversed high and low reinforcer rates. There was no evidence of identity control equivalent to 90% accuracy, and the comparison preferences reversed to varying degrees: two of four stimuli for JPW; three of four stimuli for KRK, and, in Test 2b, all stimuli for PAW. Taken together, the data for different-history trials shown in Figure 7 indicate poor stimulus control by sample-comparison identity and comparison-stimulus preferences that were related to the reinforcer frequency disparities on SSS baseline trials.

Similar comparison history trials. The results of the CSS test trials in which comparison stimuli had similar reinforcement histories were of major interest for analyzing restricted stimulus control. Each test included a total of 48 similar-history trials (eight trials per session for six sessions). As noted above, the reinforcer rates on SSS baseline trials provided no basis for a preference for either comparison stimulus on similar-history trials. Further, on all such trials the samples consisted of one high- and one low-rate stimulus (see Figure 5). If relative reinforcer rate determined the topography of restricted stimulus control, then identity control would be restricted to the high-rate samples. That is, identity control would occur on trials in which the high-rate sample was correct and the choice was between two high-rate comparisons, and identity control would not occur on trials in which the low-rate sample was correct and the choice was between two low-rate comparisons.

The left portion of Figure 8 shows the results of the initial similar-history tests. In Test 1 for KRK and JPW and in Test 1b for PAW, there were high frequencies of identity control by the high-rate sample stimuli, as shown by the points in the shaded area of the signal-detection space, and little or no identity control by low-rate stimuli. There is little evidence of comparison stimulus preferences for KRK and JPW, as shown by the proximity of the points to the minor diagonal. The data for PAW in Test 1 indicate comparison-stimulus preferences within both the high- and low-rate stimulus categories. In contrast to

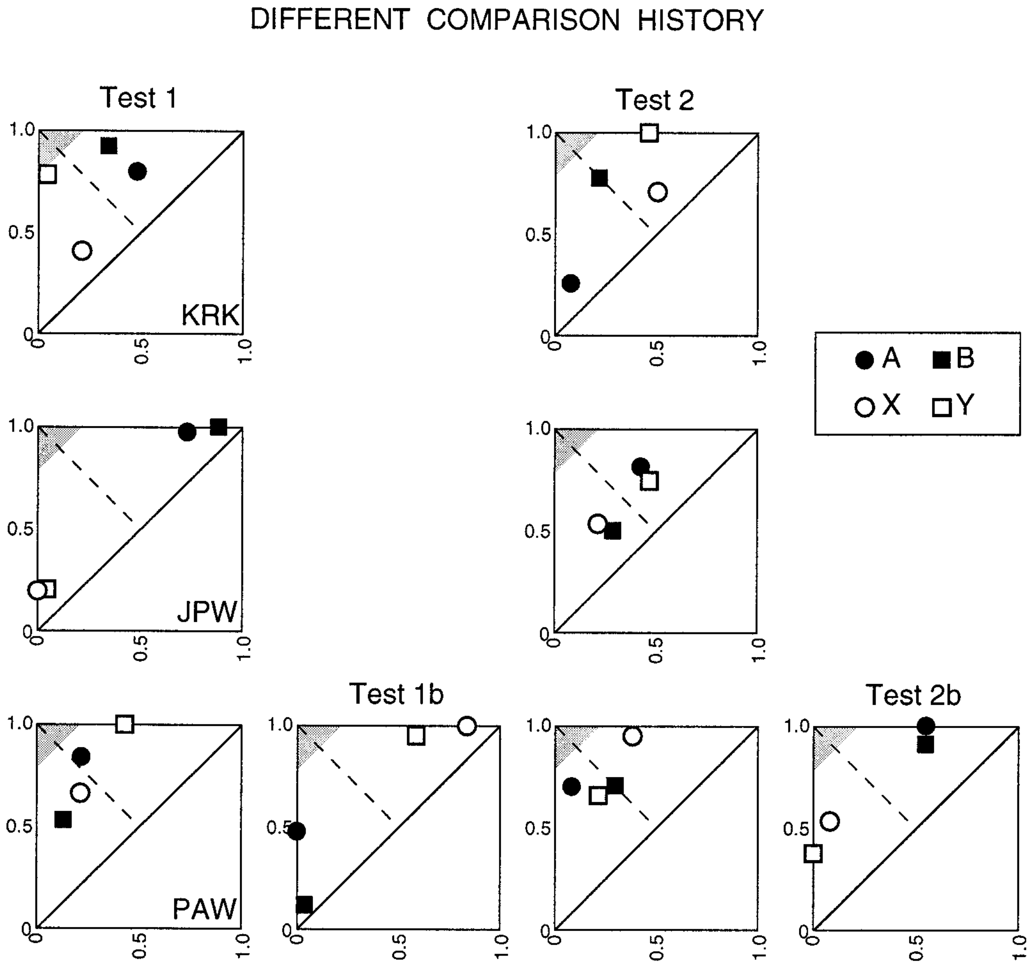


Fig. 7. Performances on CSS test trials in which comparison stimuli had different reinforcement histories. See Figure 6 for axis labels and other details. For Subjects KRK and JPW (top two rows): In Test 1, A and B were the high-rate stimuli, and X and Y were the low-rate stimuli; in Test 2, X and Y were the high-rate stimuli, and A and B were the low-rate stimuli. For Subject PAW (bottom row): In Tests 1 and 1b, X and Y were the high-rate stimuli, and A and B were the low-rate stimuli; in Tests 2 and 2b, A and B were the high-rate stimuli, and X and Y were the low-rate stimuli.

the data in Figure 7, these preferences could not have resulted from the differences in reinforcer rates on baseline trials, because both of the comparisons had approximately equal reinforcement histories. Following the increased disparity in reinforcer rates on SSS trials in Baseline 1b, PAW's Test 1b results show that the comparison preference disappeared for the high-rate stimuli and was replaced by perfect identity control, but that the preference with low-rate comparison displays became extreme.

The right portion of Figure 8 shows the results of Tests 2 and 2b, which followed baseline training with reversed high and low reinforcer rates. For Subjects KRK and JPW, the stimuli exerting identity control reversed in Test 2. There was a comparison-stimulus preference within the low-rate stimulus set for JPW but not for KRK. Subject PAW's results for Test 2 showed continued identity control by Stimuli X and Y, the low-rate stimuli in Baseline 2. The change from low to high reinforcer rates for Stimuli A and B did not result in identity control in Test 2, but it was accompanied by the elimination of the preference for comparison A (cf. Tests 1b and 2). In Test 2b, after exposure to an additional

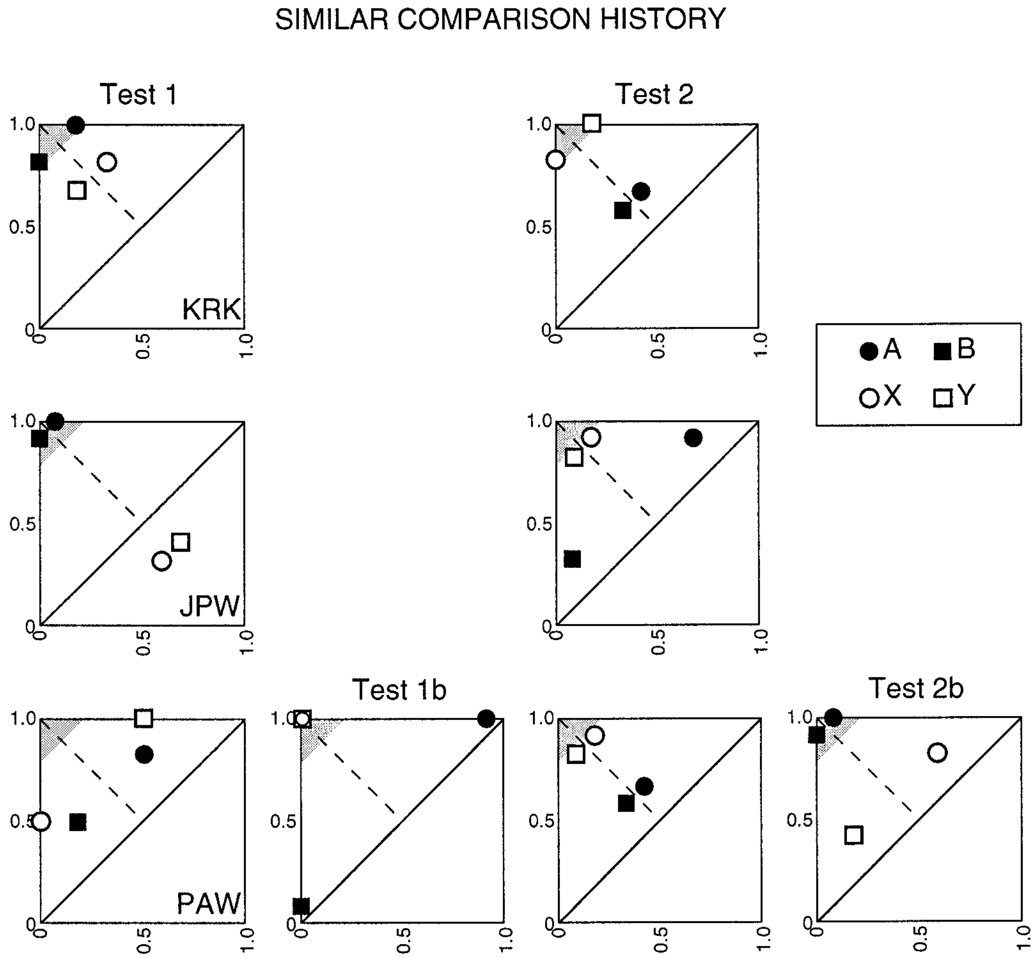


Fig. 8. Performances on CSS test trials in which comparison stimuli had similar reinforcement histories. See Figure 6 for axis labels and other details. For Subjects KRK and JPW (top two rows): In Test 1, A and B were the high-rate stimuli, and X and Y were the low-rate stimuli; in Test 2, X and Y were the high-rate stimuli, and A and B were the low-rate stimuli. For Subject PAW (bottom row): In Tests 1 and 1b, X and Y were the high-rate stimuli, and A and B were the low-rate stimuli; in Tests 2 and 2b, A and B were the high-rate stimuli, and X and Y were the low-rate stimuli.

baseline condition in which the disparity in reinforcer rates was again increased, identity control shifted to the high-rate Stimuli A and B. In sum, the data for similar-history trials in Figure 8 indicate that reinforcer rate can determine the topography of restricted stimulus control.

DISCUSSION

Our findings are consistent with those from previous studies of restricted stimulus control in individuals with intellectual disabilities. This study directly replicates the demonstration by Stromer, McIlvane, Dube, and Mackay (1993) of the SSS-CSS accuracy disparity in humans. The study goes further by initiating formal experimental analysis of how reinforcement variables may influence the nature of stimulus control by multiple sample stimuli in the CSS task. An explicit goal of our program has been to forge a link between clinically oriented research on restricted stimulus control in humans (e.g., Lovaas et al., 1979) and basic research on stimulus control, including that conducted with nonhuman animals (e.g., D'Amato & Salmon, 1984). To that end, the data presented in Figure 8 demonstrate clear experimental control of which of two sample stimuli will exert stimulus control.

Although our results are in line with certain data that have been obtained with pigeons on single-sample delayed matching-to-sample tasks (e.g., Wixted, 1989), these findings are a first in research on restricted stimulus control in humans. Moreover, these are the first such data obtained from subjects who routinely demonstrated generalized identity matching.

One aspect of Subject PAW's data merits discussion. Identity control did not reverse in Test 2 after exposure to the reversed schedules in Baseline 2, but did so after Baseline 2b, when the schedule disparity was increased from 7:1 to 23:1. Schroeder (1975) reported similar instances of perseveration following schedule changes in a study of concurrent operants in humans with mental retardation. Schroeder found that behavior often changed to conform to schedule changes only after the implementation of procedural modifications that disrupted established response patterns (see also Joyce & Chase, 1990; Mace, Neef, Shade, & Mauro, 1994). Subject PAW's data (Figure 8) suggest a similar progression. PAW was the only subject whose data showed a prereversal stimulus preference within the low-rate stimulus set (Test 1b; because the reinforcement histories were similar for Stimuli A and B, the controlling variables for such preferences are not revealed in the present analysis; see also JPW, Test 2, and PAW, Test 2b). Test 2 showed that reversing the reinforcer rates had the initial effect of eliminating the previously established preference for Stimulus A. The reversal of identity control followed later in Test 2b. One question raised by these data is whether eliminating the preference was a necessary intermediate step in the reversal or whether it merely accompanied the reversal.

One possible explanation for our main findings (Figure 8) is suggested by time-allocation analyses of behavioral choice (Baum & Rachlin, 1969). According to these analyses, the proportion of time allocated to two concurrently available discriminated operants will be related to the proportion of reinforcers obtained for each option. Time-allocation matching has been applied successfully to human vigilance (e.g., Baum, 1975), and Schroeder and Holland (1969) have reported consistent findings for eye-movement rates. For the present experiment, a matching formulation predicts longer observation of the high-rate sample stimulus than the low-rate sample stimulus. Such an outcome would be consistent with findings from experiments with explicit observing responses that show a preference for observing stimuli better correlated with reinforcement (e.g., Dinsmoor, 1983; Mulvaney, Hughes, Jwaideh, & Dinsmoor, 1981). Disparities in stimulus control such as those found in the present experiment would result if longer observation times produced higher stimulus control frequencies.

Another possible account comes from analyses of behavioral resistance to change. Nevin and colleagues' research on behavioral momentum indicates that resistance to change is directly proportional to rate of reinforcement, and it appears to be determined by the stimulus–reinforcer relation of the reinforcement contingencies (Nevin, 1992; Nevin, Mandell, & Atak, 1983). The behavioral momentum analysis has been confirmed in studies with laboratory animals and also with typically and atypically developing humans (see Nevin, 1992, for an integrative summary). For the present experiment, this formulation suggests that observing behavior controlled by the high-rate sample stimuli may be more persistent than observing behavior controlled by low-rate stimuli; that is, observing high-rate stimuli may be more resistant to competing control of observing behavior by other stimuli. As with a time-allocation matching analysis, a momentum analysis predicts the identity-control disparity if longer observation durations increase stimulus control. In future research, direct measurements of observing behavior with multiple sample stimuli will be needed to determine whether there is a relation between observation durations and restricted stimulus control.

We used the signal-detection space to present data because it provided a clear and concise way to illustrate the two main types of stimulus control the procedures were likely to produce: sample-comparison identity and comparison-stimulus preference. As Sidman (1992) pointed out, using the signal-detection space to present conditional discrimination data ''does not serve the same function for stimulus-control as for signal-detection theory'' (p. 180). A signal-detection analysis is concerned with sample-stimulus discriminability per se, whereas our stimulus control analysis was concerned with isolating a specific form of control: sample-comparison identity. Continued study of restricted stimulus control with delayed matching-to-sample procedures may benefit from greater incorporation of behavioral detection analyses of stimulus and contingency discriminability (Davison & Jenkins, 1985; Davison & Tustin, 1978). Behavioral detection analyses have been extended from the classic two-stimulus two-response situation to those involving greater numbers of stimuli (e.g., Davison, 1991), and such extensions greatly increase the number of cells in the signal-detection matrices. Applications to the CSS task, with two simultaneously displayed samples and the relatively large number of trial types shown in Figure 5, seem likely to require much larger data sets than those of the present study.

Because our procedures explicitly arranged disparities in the reinforcement schedules associated with elements of complex discriminative stimuli, it is reasonable to ask whether our data could be relevant to more typical situations in which such disparities are not explicitly arranged by an experimenter or teacher. We speculate that such disparities can in fact be subtly present in the typical situation as well. For purposes of illustration, consider a CSS task in which the printed letters QX and SZ serve as complex discriminative stimuli. Suppose further that the subject does not observe all aspects of the individual letters initially, which is a reasonable supposition for subjects with intellectual disabilities (e.g., Touchette, 1969). Note that both X and Z share a physical feature (/). To the extent that / controls a matching selection, the subject would be as likely to match X with Z and vice versa as he or she would X to X or Z to Z. The former two matching performances would never be reinforced and the latter two would be reinforced continuously; the overall reinforcement schedule for matching the common feature would approximate VR 2, given typical trial balancing practices.

In the situation just outlined, our data suggest that the subject might tend to exhibit more frequent stimulus control by the elements Q and S, matching performances that would be reinforced continuously rather than intermittently. The terms *restricted stimulus control* or *stimulus overselectivity* would describe the resulting data, but the analysis would not be complete without an understanding of the role of the relevant reinforcement variables. Such accounts of restricted stimulus control in typical situations seem plausible, even though complex discriminative stimuli may not share obvious physical features. Ample research studying individuals with intellectual disabilities has demonstrated stimulus control by extremely subtle and initially unsuspected aspects of complex stimuli (e.g., McIlvane & Cataldo, 1996; Schreibman & Lovaas, 1973). Also potentially relevant is the role of extraexperimentally established stimulus control and associated reinforcement schedules, whereby primary stimulus generalization from stimuli encountered outside of experimental settings could introduce unsuspected reinforcer-related inequalities.

To conclude, it seems appropriate to consider what the present data and conceptual analysis might add to the development of effective procedures for managing or ameliorating restricted stimulus control in clinical populations. Although such problems have been appreciated for at least 25 years, there have been few reports of successful broadening of control, typically by arranging contingencies that require control by compound stimuli (Allen & Fuqua, 1985; Schreibman, Charlop, & Koegel, 1982). In our experience, such procedures can be effective, but often are not. The uneven success may help to explain why the literature on remediation of a significant clinical problem is so small. Results of the present study raise the question of whether manipulating specific reinforcement histories could be an effective remedial tactic.

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.

Baum, W. M. (1975). Time allocation in human vigilance. *Journal of the Experimental Analysis of Behavior, 23,* 45–53.

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior, 12,* 861–874.

Bickel, W. K., Richmond, G., Bell, J., & Brown, K. (1986). A microanalysis of the controlling stimulus-response relations engendered during the assessment of stimulus overselectivity. *The Psychological Record, 36,* 225–238.

Bickel, W. K., Stella, M. E., & Etzel, B. C. (1984). A reevaluation of stimulus overselectivity: Restricted stimulus control or stimulus control hierarchies. *Journal of Autism and Developmental Disorders, 14,* 137–157.

Cox, J. K., & D'Amato, M. R. (1982). Matching to compound samples by monkeys (*Cebus apella*): Shared attention or generalization decrement? *Journal of Experimental Psychology: Animal Behavior Processes, 8,* 209–225.

D'Amato, M. R., & Salmon, D. P. (1984). Cognitive processes in cebus monkeys. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 149–168). Hillsdale, NJ: Erlbaum.

Davison, M. C. (1991). Stimulus discriminability, contingency discriminability, and complex stimulus control. In M. L. Commons, J. A. Nevin, & M. C. Davison (Eds.), *Signal detection: Mechanisms, models, and applications* (pp. 57–78). Hillsdale, NJ: Erlbaum.

Davison, M., & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behavior, 13,* 77–84.

Davison, M. C., & Tustin, R. D. (1978). The relation between the generalized matching law and signal-detection theory. *Journal of the Experimental Analysis of Behavior, 29,* 331–336.

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences, 6,* 693–728.

Dube, W. V., Kledaras, J. B., Iennaco, F. M., Stoddard, L. T., & McIlvane, W. J. (1990). Observing complex visual stimuli: Effects of component pretraining. *Experimental Analysis of Human Behavior Bulletin, 8,* 7–11.

Dunlap, G., Koegel, R. L., & Burke, J. C. (1981). Educational implications of stimulus overselectivity in autistic children. *Exceptional Education Quarterly, 2(13),* 37–49.

Hartl, J. A., & Fantino, E. (1996). Choice as a function of reinforcement ratios in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 66,* 11–27.

Joyce, J. H., & Chase, P. N. (1990). Effects of response variability on the sensitivity of rule-governed behavior. *Journal of the Experimental Analysis of Behavior, 54,* 251–262.

Koegel, R. L., & Wilhelm, H. (1973). Selective responding to the components of multiple visual cues by autistic children. *Journal of Experimental Child Psychology, 15,* 442–453.

Kovattana, P. M., & Kraemer, H. C. (1974). Response to multiple visual cues of color, size, and form by autistic children. *Journal of Autism and Childhood Schizophrenia, 4,* 251–261.

Litrownik, A. J., McInnis, E. T., Wetzel-Pritchard, A. M., & Filipelli, D. L. (1978). Restricted stimulus control and inferred attentional deficits in autistic and retarded children. *Journal of Abnormal Psychology, 87,* 554–562.

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin, 86,* 1236–1254.

Mace, F. C., Neef, N. A., Shade, D., & Mauro, B. C. (1994). Limited matching on concurrent-schedule reinforcement of academic behavior. *Journal of Applied Behavior Analysis, 27,* 585–596.

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

McIlvane, W. J., & Cataldo, M. F. (1996). On the clinical relevance of animal models for the study of human mental retardation. *Mental Retardation and Developmental Disabilities Research Reviews, 2,* 188–196.

Mulvaney, D. E., Hughes, L. H., Jwaideh, A. R., & Dinsmoor, J. A. (1981). Differential production of positive and negative discriminative stimuli by normal and retarded children. *Journal of Experimental Child Psychology, 32,* 389–400.

Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior, 57,* 301–316.

Nevin, J. A., Mandell, C., & Atak, J. R. (1983). The analysis of behavioral momentum. *Journal of the Experimental Analysis of Behavior, 39,* 49–59.

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

Schreibman, L., Charlop, M. H., & Koegel, R. L. (1982). Teaching autistic children to use extra-stimulus prompts. *Journal of Experimental Child Psychology, 33,* 475–491.

Schreibman, L., & Lovaas, O. I. (1973). Overselective response to social stimuli by autistic children. *Journal of Abnormal Child Psychology, 1,* 152–168.

Schroeder, S. (1975). Perseveration in concurrent performances by the developmentally retarded. *The Psychological Record, 25,* 51–64.

Schroeder, S. R., & Holland, J. G. (1969). Reinforcement of eye movement with concurrent schedules. *Journal of the Experimental Analysis of Behavior, 12,* 897–903.

Sidman, M. (1980). A note on the measurement of conditional discrimination. *Journal of the Experimental Analysis of Behavior, 33,* 285–289.

Sidman, M. (1992). Adventitious control by the location of comparison stimuli in conditional discriminations. *Journal of the Experimental Analysis of Behavior, 58,* 173–182.

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. *Psychological Record, 43,* 585–598.

Touchette, P. E. (1969). Tilted lines as complex stimuli. *Journal of the Experimental Analysis of Behavior, 12,* 211–214.

Wilhelm, H., & Lovaas, O. I. (1976). Stimulus overselectivity: A common feature in autism and mental retardation. *American Journal of Mental Deficiency, 81,* 26–31.

Wixted, J. T. (1989). Nonhuman short-term memory: A quantitative reanalysis of selected findings. *Journal of the Experimental Analysis of Behavior, 52,* 409–426.

Received September 4, 1996
Final acceptance July 8, 1997