JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2002, **77,** 189–198 NUMBER 2 (MARCH)

HIGH-PROBABILITY STIMULUS CONTROL TOPOGRAPHIES WITH DELAYED S+ ONSET IN A SIMULTANEOUS DISCRIMINATION PROCEDURE

WILLIAM J. MCILVANE, JOANNE B. KLEDARAS,
T. C. CALLAHAN, AND WILLIAM V. DUBE

E. K. SHRIVER CENTER,
UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL,
PROTESTANT GUILD LEARNING CENTER,
AND SHRIVER CLINICAL SERVICES CORPORATION

Experimenters and teachers use discrimination learning procedures to encourage reliable attending to stimulus differences defined as relevant for their purposes. Put another way, the goal of discrimination training is to establish high-probability stimulus control topographies that are coherent with experimenter or teacher specifications. The present research was conducted to investigate a novel procedure for encouraging stimulus control topography coherence. Participants were 13 adolescents with severe intellectual handicaps. During an initial Condition A, all were exposed to a simultaneous discrimination procedure. Participants could select a form alternating with a black field (S+) or an identical form that did not alternate (S−). Accuracy scores were typically low, and there was little evidence of coherent stimulus control topographies. Subsequently, the procedure was changed. During Condition B, every trial initially presented two identical nonalternating S− forms (Trial State 1). If the participant made no selection for 5 s, one of the forms began to alternate with the black field, and he or she could make the S+/S− discrimination (Trial State 2). Selections during Trial State 1 prolonged the delay to Trial State 2 until there had been no response for 5 s. During Condition B, S+/S− discrimination accuracy scores improved rapidly and markedly for most participants. Reinstating Condition A often resulted in diminished accuracy scores. This study thus (a) demonstrated a novel procedure for encouraging stimulus control topography coherence and (b) provided support for the interpretation that intermediate accuracy scores may be due to different topographies of stimulus control that co-occur in the same discriminative baseline.

Key words: discrimination learning, stimulus control topographies, key press, humans with mental retardation

---

The term *stimulus control* specifies a relation involving a class of behavior and a class of environmental events. Such a controlling relation requires merely that a given behavior be more (or less) probable when a given stimulus is present. It does not require that the behavior always (or never) occur in response to the controlling stimulus. There are circumstances both inside and outside the laboratory, however, in which high-probability controlling relations are essential. When driving, for example, stopping in the presence of red traffic lights and proceeding in the presence of green ones must occur reliably; discrimination failures may have fatal consequences. The present studies investigated circumstances under which highly reliable (i.e., high-probability) controlling relations might be established in a simple simultaneous discrimination task.

Discrimination training is a traditional procedure for establishing stimulus control. For example, consider a simple simultaneous discrimination procedure that displays a red light (S+) on one response key and a white light (S−) on another; key positions of the colors vary unsystematically across trials. A concurrent reinforcement schedule is in effect. Responses to the S+ are followed by reinforcers (fixed-ratio 1) and those to the S− are not (extinction). Early in training, the participant may respond to both colors equally often. As training progresses, however, one may observe more frequent responses to the S+ and less frequent responses to the S−. If the participant responds to the red light on 70% of trials, for example, one has evidence of stimulus control by some aspect of the red

This research was supported by NICHD Grants HD 25995 and HD 28141 and by a contract from the Commonwealth of Massachusetts (100220023SC). Statements made in this article reflect the authors' own opinions and do not constitute policy statements from their various affiliated institutions and funding agencies. We thank Richard Serna for his comments on an earlier draft of the manuscript.

Reprint requests can be addressed to William McIlvane, E. K. Shriver Center, 200 Trapelo Road, Waltham, Massachusetts 02452 (e-mail: william.mcilvane@umassmed.edu).

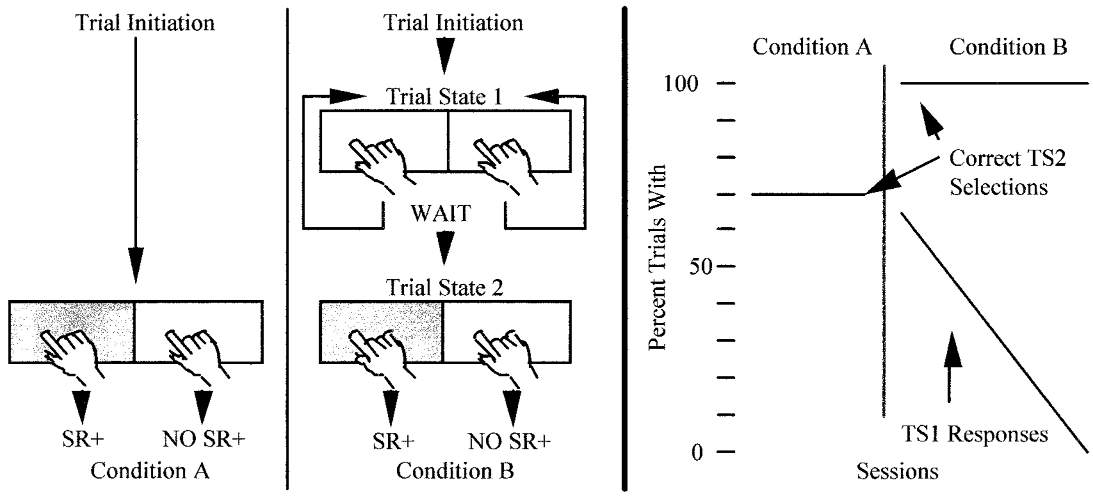

Fig. 1. Diagram of the delayed S+ procedure and hypothetical data suggesting its potential effects on a baseline of intermediate accuracy. TS1 and TS2 refer to Trial States 1 and 2, respectively.

versus white light difference. Control is unreliable, however, given that the participant responds to the white light on 30% of trials.

Increases in accuracy scores do not always result from continued discrimination training. Studies of humans with mental retardation, for example, have reported little or no departure from chance discrimination accuracy despite hundreds or thousands of training trials (e.g., Zeaman & House, 1963). Other discrimination training studies with this population have reported apparently asymptotic levels of discrimination (i.e., 65% to 80% correct), particularly on tasks that require the participant to observe multiple stimulus elements or features before responding (Ellis, Girardeau, & Pryer, 1962; House, Hanley, & Magid, 1979). Such findings, particularly those of the latter type, have never been satisfactorily explained.

Concerning persistently low accuracy scores, one suggestion has been that typical laboratory procedures may capture and maintain topographies of stimulus control that do not cohere with the experimenter-specified topographies (see McIlvane & Dube, 1992, for a discussion of the stimulus control topography concept). For example, whereas the experimenter may want the participant to attend to form differences, the participant may respond instead to position stimuli. Concerning persistently intermediate scores, Dube and McIlvane (1996) suggested that these scores may result when reinforcement contingencies capture and maintain multiple topographies of stimulus control within the same discrimination baseline. For example, they suggested that intermediate accuracy could result if behavior were occasioned on some trials by stimuli that were consistent with the experimenter-defined controlling relations and on other trials by stimuli that were not (see Sidman, 1969, 1980, for another perspective on intermediate accuracy scores).

Although there has been much speculation about the variables that are responsible for persistently low or intermediate discrimination accuracy scores, the problem has attracted little experimental study. Also little studied are circumstances under which discrimination training might establish high-probability stimulus control topographies. In earlier studies with individuals with severe mental retardation, we sought to analyze persistently imperfect discrimination accuracy scores and to establish reliable stimulus control (McIlvane, Kledaras, Dube, & Stoddard, 1989). We used a simultaneous discrimination procedure like that suggested in the leftmost portion of Figure 1 (labeled Condition A). After a trial-initiation response, two identical food items were displayed in side-by-side compartments. On each trial, the participant was to select the compartment that was illuminated with a red flashing light (S+) and to reject one that was lit with steady white light (S−);

the S+ position varied unsystematically across trials. With this procedure, a number of participants achieved discrimination accuracy scores that were significantly above 50% correct, suggesting some stimulus control by the red light. Consistently high accuracy scores were never achieved, however. The intermediate scores are suggested by the horizontal line under Condition A on the right portion of Figure 1.

After many failures to remediate our participants' learning problems, we developed the method labeled Condition B in Figure 1. This "delayed S+" method separated the trial into two "states." During the first, "Trial State 1," there was no difference in the lighting of the compartments and no basis for a red versus white discrimination. The programmed Trial State 1 duration was brief (i.e., 3 to 5 s), but any Trial State 1 responses lengthened its duration by the programmed value (i.e., 3 to 5 s were added). The second state, Trial State 2, commenced at the end of Trial State 1; the red light S+ was presented in one compartment, and a red versus white discrimination could be made.

The outcome of this delayed S+ procedure is shown schematically in the rightmost portion of Figure 1 (Condition B). Participants initially exhibited high levels of Trial State 1 responding, but such responding decreased to negligible levels as delayed S+ training progressed. During Trial State 2, however, participants immediately displayed high discrimination accuracy scores. We interpreted our data as follows. Prior to the delayed S+ procedure, the discrimination training had established competing stimulus control topographies. The target topography involving the red light occurred on many trials. Discrimination training had also established other stimulus control topographies (e.g., involving position stimuli, constant environmental features, etc.). Because any competing topographies occurred while the S+ was displayed, they were followed by reinforcers on about half of the trials of our two-choice task. Competing stimulus control topographies were thus reinforced intermittently and maintained. When the delayed S+ procedure was implemented, however, the competing topographies occurred during Trial State 1 and were no longer reinforced; their occurrence became less probable. The target stimulus control topography was not subject to extinction, however, and it became relatively more probable.

Our findings with the delayed S+ procedure were clear and consistent, but their generality remained to be determined. Generality was an important issue, because the study involved a unique setting and atypical procedures. The work was conducted as we implemented a larger program that sought to establish useful forms of instructional control with participants who were severely mentally retarded (McIlvane et al., 1989; Stoddard, 1982). The setting was an automated teaching laboratory that was designed especially for that purpose. During initial phases of the program, one teaching goal was to establish red lights as discriminative stimuli for responses to several different locations and manipulanda. One question, therefore, was whether our findings of principal interest—the immediately reliable control of compartment selection by the red light during Trial State 2—depended upon similar control established on previous tasks (e.g., initiating trials by responding to a red light in another location, discriminating a key lit with a red form from seven darker keys, etc.). The present study was undertaken to determine whether the delayed S+ procedure would produce comparable results when implemented with a more conventional experimental setting and discrimination learning procedure.

METHOD

Participants and Setting

Thirteen individuals participated. All were students at residential schools for individuals with mental retardation associated with autism, pervasive developmental disorders, or other diagnoses. Ages ranged from 13 to 22 years. The Peabody Picture Vocabulary Test was given to provide an estimate of intellectual functioning. Age-equivalent scores ranged from 23 to 60 months (M = 32 months). All but 3 participants were male, and all were considered to be moderately to severely or severely mentally retarded. Sessions were conducted two to four times per week in a quiet area of the participant's classroom or in a separate office at the participant's residence. The participant sat at a table in front of a microcomputer apparatus and the experimenter sat to the side and behind, such that the participant could not observe the apparatus and experimenter simultaneously.

Apparatus

The apparatus was a portable Macintosh® computer fitted with a touch-sensitive screen (Dube & McIlvane, 1989). The screen (19 cm by 14 cm) displayed experimental stimuli on a light gray background, and the participant responded to a stimulus by touching it. Responses were automatically recorded, and data were saved on disk.

Procedure

Pretraining. Sessions typically lasted about 10 to 15 min, consisted of 36 to 48 discrimination trials, and were separated by at least 24 hr. All participants had received pretraining via an introductory program designed to establish tokens as conditioned reinforcers and to familiarize them with the apparatus (Dube, Iennaco, & McIlvane, 1993; Dube & McIlvane, 1989). During pretraining, all participants learned to touch a form (a plus sign) when it was presented in any one of four different positions on the screen. Every trial began when stimuli to be discriminated were presented. Touches to stimuli defined as S+ were followed by a series of computer-generated musical notes and a complex animated display that filled the computer screen. These consequences were accompanied by the delivery of a token. During pretraining, the plus sign was the only form on the screen, and touches to other locations had no programmed consequences. During pretraining and all subsequent conditions, 1.5-s intertrial intervals were programmed automatically unless the participant touched the screen; intertrial responses delayed the onset of the trial for an additional 1.5 s.

Condition A. This condition was like the pretraining condition, except as follows. Stimuli to be discriminated were presented in two locations in the lower left and lower right corners of the screen. The S+ stimulus was a nonrepresentative black letterlike form that alternated every 0.5 s with a gray field that made the form appear to flash on and off. We will term this the *flashing* stimulus; the S− was the same form, but it did not alternate with the field. We will term this the *steady* stimulus. Selections of S+ and S− were followed by reinforcing consequences and a 3-s blackout, respectively. Across trials, the S+ appeared about equally often in each of the two positions. Nine participants received four Condition A sessions initially. The other 4 received 5, 7, 9, and 15 Condition A sessions initially.

Condition B. This condition implemented the delayed S+ procedure. Trial State 1 commenced with the presentation of the steady stimulus in both positions for 5 s. Any response to either steady stimulus (a Trial State 1 error) prolonged Trial State 1 by resetting the 5-s timer, so that Trial State 2 did not begin until there was no response for 5 s. During the subsequent Trial State 2, one of the forms began to alternate (S+) and the participant could make the flashing versus steady stimulus discrimination; selection of the steady stimulus was a Trial State 2 error. The Trial State 2 contingencies were identical to those in effect during Condition A. Selections of the flashing stimulus were followed by reinforcers, and selections of the steady stimulus produced the blackout. Condition B was in effect for 2 to 17 sessions, depending in part on performance.

Follow-up. For all but 1 participant, Condition B was followed by a return to Condition A. In follow-up work, some participants received further exposures to the two conditions.

RESULTS

Condition A. Data for the participants who received four initial Condition A sessions are shown in Figure 2. Condition A performance is shown in the leftmost portion of each plot. None of the participants made an accurate flashing versus steady discrimination during these sessions. Figure 3 shows that similar findings were obtained with the remaining participants, despite their greater number of Condition A sessions.

Both figures also show the percentage of trials on which the participant selected the left position in each session (position control scores). Given the position-balanced two-choice task, responding not controlled by po-

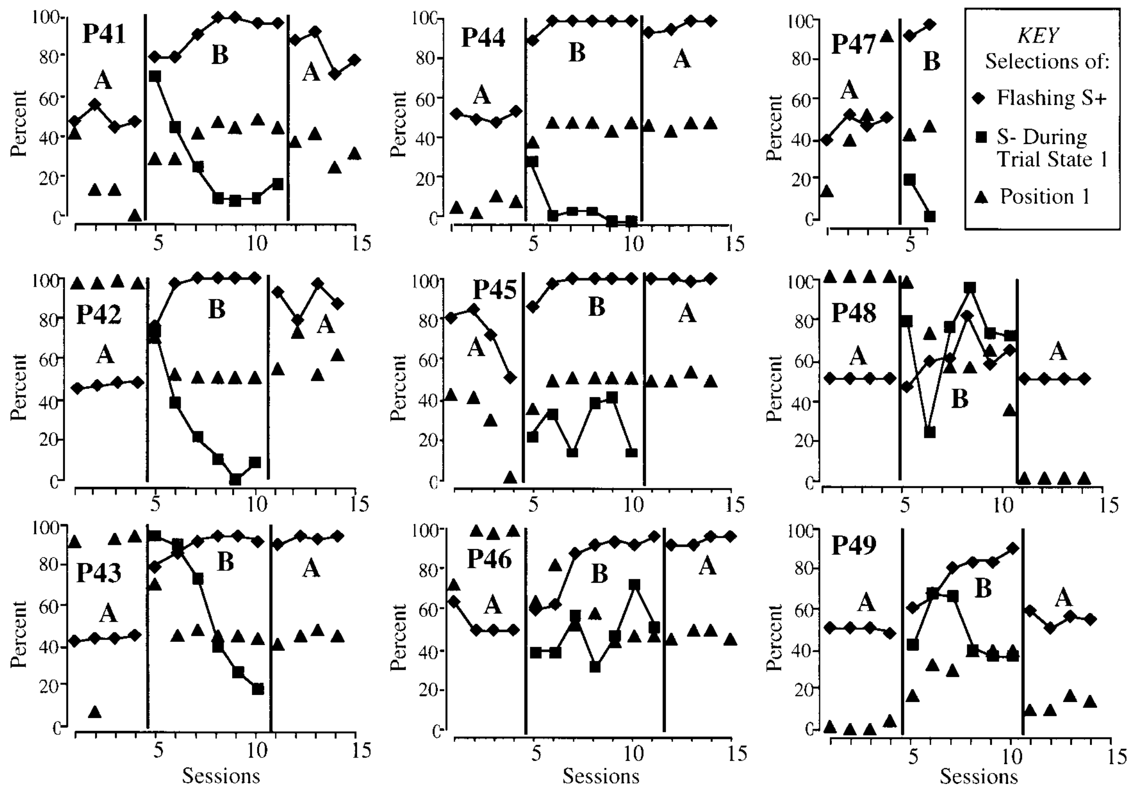

Fig. 2. Discrimination accuracy and position control scores for participants who had received four initial Condition A sessions. Filled diamonds show scores on the flashing versus steady discrimination. Filled triangles show the percentage of trials on which the participant selected the left position. Filled squares show the percentage of trials on which the participant made one or more selections of the forms displayed during Trial State 1 of Condition B sessions.

sition stimuli would produce position control scores in the vicinity of 50%. Position control scores of 100% or 0% indicate exclusive stimulus control by position, and intermediate scores suggest mixed stimulus control by position and other stimuli. During Condition A, position control scores were either very high or very low for most participants, particularly in the later sessions. These scores indicate that position control was the dominant stimulus control topography captured by the contingencies.

Condition B. Figures 2 and 3 show that flashing versus steady discrimination accuracy scores during Condition B became high, often immediately and markedly so. For 11 of 13 participants, accuracy scores in the first Condition B session were higher than during the last Condition A session. This consistency suggests that the improvement in flashing versus steady discrimination accuracy was due to the imposition of the delayed S+ procedure and not merely to continued discrimination training. Notably, 3 of 4 participants who received more than four initial Condition A sessions showed immediate substantial increases in flashing versus steady discrimination accuracy in their first Condition B session; this finding also suggests that it was the imposition of the delayed S+ procedure and not merely continued discrimination training that led to the higher accuracy scores.

Both figures also show the percentage of trials on which the participant made one or more responses to the forms displayed during Trial State 1. For all but a few participants, the highest level of such responding was observed in the first Condition B session and subsequently declined to low levels in the course of further sessions. Exceptions were P45, P46, and P49, who initially displayed and maintained low to intermediate levels of Trial State 1 responding, and P48 (see below).

Follow-up. For all but 1 participant, there

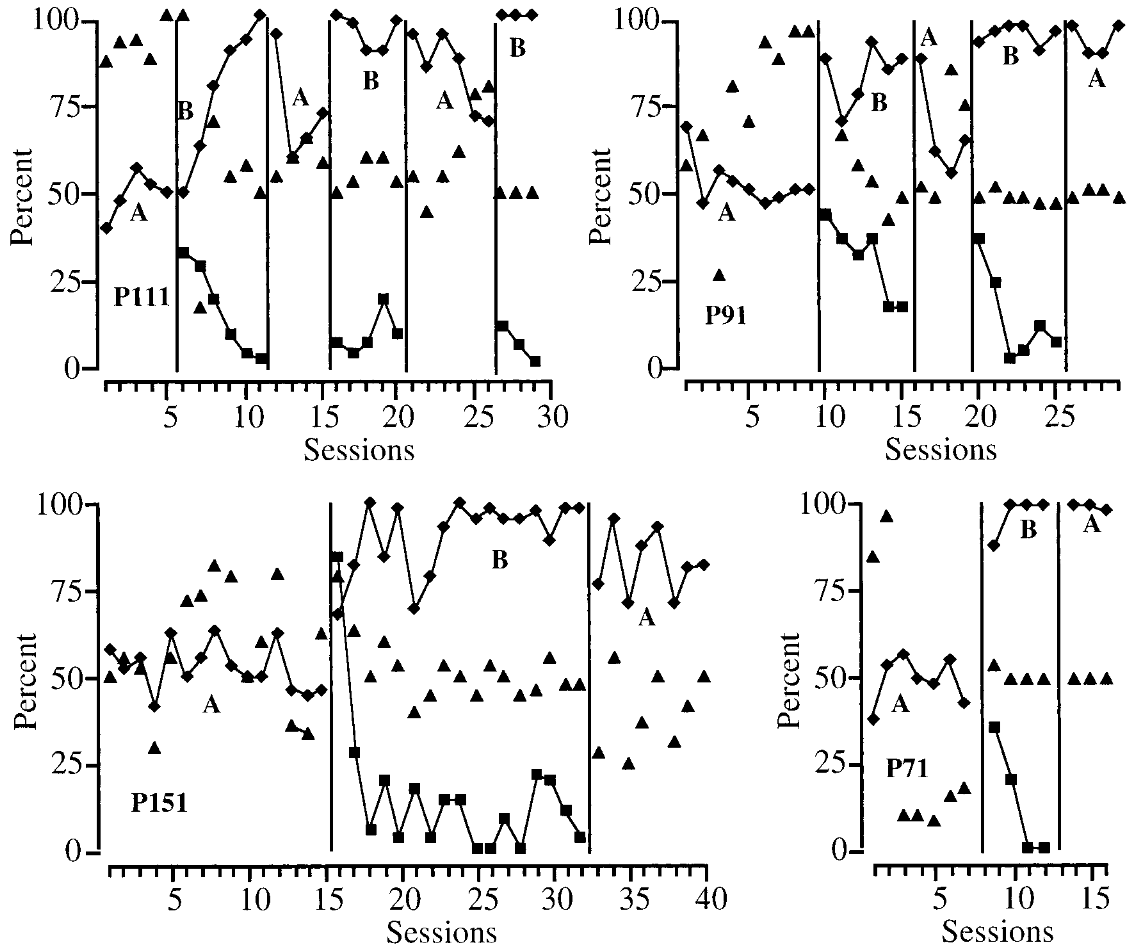

Fig. 3. Discrimination accuracy and position control scores for participants who had received more than four initial Condition A sessions. See Figure 2 for a description of each of the data-point types.

was a return to Condition A following Condition B. Six participants exhibited declines in accuracy of the flashing versus steady discrimination, thus providing further evidence that the delayed S+ procedure was an effective variable during Condition B. For several other participants, the contingency reversal design was not helpful in this regard. For these participants, the main evidence of the effectiveness of the procedure was the typically very rapid increase in flashing versus steady discrimination accuracy in the first Condition B session.

Extended testing was conducted with P48, who had displayed low to intermediate flashing versus steady discrimination accuracy during the initial Condition B sessions. We conducted 37 further sessions (data not shown). Blocks of Condition A sessions were alternated with blocks of Condition B sessions. During the third Condition B block (the 36th session), we observed a rapid increase in accuracy in the flashing versus steady discrimination, similar in character to those exhibited initially by the other participants. Scores in the next six sessions averaged 88% correct, and continued training resulted in further increases in accuracy. Accuracy scores during the last three sessions averaged 98% correct. Thus the delayed S+ procedure had a similar, albeit delayed, effect with this participant.

DISCUSSION

This study systematically replicates and complements research reported by McIlvane et al. (1989) in several ways: (a) For most participants, imposing the delayed S+ procedure produced an immediate, often marked, increase in discrimination accuracy (i.e., during Trial State 2). (b) For most, but not all, participants, these increased accuracy scores were accompanied by declines in responding during Trial State 1. (c) Eliminating the delayed S+ procedure was often, but not always, accompanied by diminished flashing versus steady discrimination accuracy; this diminished accuracy provides further evidence that the delayed S+ procedure was an effective variable during Condition B. Notably, all of these findings were obtained with a conventional discrimination procedure; they did not depend upon the stimulus conditions and training methods used in the automated teaching laboratory.

This study also obtained findings that were not reported by McIlvane et al. (1989). For example, the delayed S+ procedure did not appreciably alter Trial State 2 discrimination accuracy initially for 1 participant (P48). Protracted exposure to the procedures was necessary to establish the discrimination. Also different from the earlier results was the gradual increase in Trial State 2 accuracy shown by 4 participants (P46, P49, P111, and P151) following initial imposition of the delayed S+ procedure. In the earlier study, Trial State 2 accuracy increases were typically instantaneous or virtually so. Perhaps this outcome was due to the greater amount of pretraining provided in the teaching laboratory; that training might have rendered control by relevant stimulus differences more probable (see the introduction).

Taken together, the data demonstrate a frequently observed but poorly understood phenomenon: the highly variable response of participants to discrimination training procedures. Perhaps such variability must be expected. As Ray and Sidman (1970) wrote many years ago,

> All stimuli are [complex] in the sense that they have more than one dimension or aspect to which a participant might attend. To ask the experimenter to be aware of all possibilities is already, perhaps, an impossible demand. To ask further that the experimenter arrange conditions so that no undesired stimulus-response correlation is ever reinforced sets a truly impossible task. For these reasons, we may never have a generalizable formula for ''forcing'' participants to discriminate a specific stimulus aspect. We may have to settle, instead, for a combination of techniques, each of which is known to encourage stimulus control. (p. 199)

The delayed S+ procedure may prove to be one useful technique for encouraging stimulus control topographies that are consistent with experimenter or teacher definitions. As noted earlier, our interpretation is that Trial State 1 permits potentially competing topographies to occur and to be reduced in probability through extinction. When Trial State 2 commences, the probability of the target stimulus control topographies is correspondingly more probable. However, another contributing variable may be that the change from Trial State 1 to Trial State 2 results in the addition of a relatively novel (i.e., within the trial) stimulus characteristic (i.e., flashing) that defines the positive stimulus. Perhaps that readily detectable change in the stimulus array helps to direct attending to relevant stimulus differences.

The delayed S+ procedure is also of interest for analyzing the variables that are responsible for protracted low or intermediate accuracy scores on discrimination tasks. Our data emphasize the point that such scores might well be interpreted not as the absence of stimulus control but rather as the presence of stimulus control topographies that do not match those the experimenter wants to establish. This point is made especially well by the position control scores of many participants during the initial Condition A sessions; virtually all participants exhibited position-related topographies—not ''random'' responding. Related to this point is the suggestion of Dube and McIlvane (1996) that intermediate accuracy scores reflect mixtures of stimulus control topographies that do and do not match experimenter-specified topographies. For example, consider the Condition B performance of P91 and P41; the gradually increasing flashing versus steady discrimination accuracy directly tracked the decrease in preference for a particular position. Although the possibility of multiple and mixed sources of stimulus control within a discrimination baseline has been recognized for some time, there have been few direct empirical demonstrations of this phenomenon (e.g., Sidman, 1969, 1992). Our delayed S+ data are especially striking in this regard, particularly the very rapid increases in accuracy shown by many participants in the first Condition B session. To our knowledge, no previous study has reported such an immediate and clear separation of stimulus control topographies in individual participants.

The second purpose of the delayed S+ procedure was to provide a technique for rapidly isolating and selectively reinforcing desirable stimulus control topographies. However, if desirable topographies had never occurred in prior training (e.g., P48) or were extremely rare (as suggested in the data of P42, P46, and P49), then one would expect somewhat different results with the delayed S+ procedure. In such cases, the main effect of the procedure would be to extinguish the exclusive or highly predominant topography, increase behavioral variability, and perhaps encourage coherent topographies. We speculate that such processes are operative when the procedure leads to gradual improvement in accuracy. Unfortunately, not all stimulus control topographies are amenable to simple analyses such as calculating a position control score (e.g., frequent alternation between positions). Indeed, some initial data (e.g., Condition A for P41, P45, P91, and P151) can be interpreted as mixtures of competing irrelevant topographies. Further methodological development will be necessary to analyze fully such mixtures.

Encouraging high-probability stimulus control topographies. Our studies make several points that are not yet well represented in the literature of the experimental analysis of behavior. First, although there has been voluminous study of methods for establishing simultaneous and successive discrimination by differential reinforcement, little attention has been paid to the problem of producing high-probability stimulus control topographies (i.e., perfect or virtually perfect discrimination performances). Relative neglect of this problem is understandable given that the experimental questions of many types of studies do not require high-probability topographies. In behavioral pharmacology, for example, an 85% correct discrimination baseline may be more than adequate for assessing the impact of a given drug on discrimination. Nonetheless, there are circumstances under which even fairly infrequent discrimination failures may command attention and experimental analysis. In the past, for example, behavior analysts paid substantial attention to the vigilance decrements shown by radar operators on early warning systems (Holland, 1958). In those circumstances, it was necessary to ensure that operator discriminations were as close to perfection as possible.

Our laboratory has as a major focus building the capacity for teaching reliable discrimination skills to individuals with developmental limitations and disabilities. In our application, discrimination skills that are above chance but short of perfection have limited usefulness. For example, consider an individual who has learned to discriminate a $1 bill from a $20 bill with 90% accuracy. While that fairly high level of performance may reflect significant learning, the 10% error rate is clearly unacceptable for functional use of money. Technology must be developed or applied to ''purify'' the discrimination baseline; the problem is roughly analogous to that faced by the chemist who must achieve a high level of purity of a given chemical or chemical compound.

Within the experimental analysis of behavior, there appear to be two major approaches to producing high-probability stimulus control topographies. The first is protracted exposure to contingencies of differential reinforcement, sometimes supplemented by procedural variations that have been shown to enhance accuracy (e.g., trial correction procedures, blackouts following errors, etc.). This approach has acknowledged limitations (an often prohibitively long training course with no guarantee of success, emotional responses associated with protracted exposure to contingencies that produce many errors, etc.). The other approach, stimulus control shaping, begins with an already-established discrimination and uses programs of gradual stimulus change to shape new stimulus control topographies. Among the limitations of this approach is the poor current understanding of the variables that determine whether stimulus control transfers or not; stimulus control shaping remains more of an art than a science (Serna & Carlin, 2000).

In our view, both of these approaches help to illustrate the limits of our still incomplete understanding of how operant discriminations develop (or fail to do so) (cf. Dinsmoor, 1985). In pursuit of furthering that understanding, a number of studies have endeavored to accomplish so-called ''microanalyses'' of discrimination learning (Bickel, Richmond, Bell, & Brown, 1986; Dube & McIlvane, 1997; Stromer, McIlvane, Dube, & Mackay, 1993). The essence of the microanalytic approach is a ''quantal'' formulation of stimulus control (cf. Bickel & Etzel, 1985; Rilling, 1977). Briefly, stimulus control is conceptualized as a discrete rather than a continuous variable. A pattern of discrimination accuracy ranging from near chance through intermediate accuracy to ultimate perfection is not interpreted as the development of progressively ''stronger'' control by experimenter-specified stimulus differences. Rather, such a pattern is interpreted as a progressive increase in the frequency of coherent stimulus control topographies. Intermediate accuracy scores are interpreted as averaging the frequencies of multiple discrete stimulus control topographies, only some of which are desired by the experimenter or teacher (see Migler, 1964, for a similar analysis applied to the interpretation of generalization gradients).

If discrimination baselines of intermediate accuracy are in fact composed of multiple, competing stimulus control topographies, it should be possible to design analytical procedures (analogous to filters) to separate the desired stimulus control topographies from the undesired ones. The delayed S+ procedure used in this study might provide a model of such a procedure. Consider, for example, the performance of P44. This participant exhibited virtually instantaneous high-probability discrimination accuracy in the first Condition B session and subsequently maintained virtually perfect accuracy in all subsequent sessions. Several other participants (P43, P45, P46, and P71) also maintained high accuracy upon a return to Condition A. In these cases especially, it appeared that the delayed S+ procedure had acted as a very effective filter of stimulus control topographies.

Our study does not clarify, however, the individual differences that were observed in our participants' responses to the procedures. Why did the delayed S+ procedure act as an effective filter for some participants and not for others? Also, it remains a puzzle as to why Trial State 1 responding was highly persistent in some participants and relatively short-lived in others. Future research might explore whether protracted Trial State 1 responding might be correlated with more general difficulties in mastering successive discriminations that are exhibited by some individuals with mental retardation (cf. Saunders & Spradlin, 1989, 1990). Also of interest would be efforts to understand better why some participants adopt exclusive position habits but others exhibit more apparent variability in the stimulus control topographies that constitute their baselines. Yet to be determined is whether these differences reflect detectable neurological differences associated with the participants' developmental status, differences in preexperimental history, or both.

REFERENCES

Bickel, W. K., & Etzel, B. C. (1985). The quantal nature of controlling stimulus-response relations as measured in tests of stimulus generalization. *Journal of the Experimental Analysis of Behavior, 44,* 245–270.

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.

Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior, 43,* 365–382.

Dube, W. V., Iennaco, F. M., & McIlvane, W. J. (1993). Generalized identity matching-to-sample of two dimensional forms in individuals with intellectual disabilities. *Research in Developmental Disabilities, 14,* 457–477.

Dube, W. V., & McIlvane, W. J. (1989). Adapting a microcomputer for behavioral evaluation of mentally retarded individuals. In J. A. Mulick & R. F. Antonak (Eds.), *Transitions in mental retardation* (Vol. 4, pp. 104–127). Norwood, NJ: Ablex.

Dube, W. V., & McIlvane, W. J. (1996). Some implications of a stimulus control topography analysis for emergent stimulus classes. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 197–218). North Holland: Elsevier.

Dube, W. V., & McIlvane, W. J. (1997). Reinforcer frequency and restricted stimulus control. *Journal of the Experimental Analysis of Behavior, 68,* 303–316.

Ellis, N. R., Girardeau, F. L., & Pryer, M. W. (1962). Analysis of learning sets in normal and severely defective humans. *Journal of Comparative and Physiological Psychology, 55,* 860–865.

Holland, J. G. (1958). Human vigilance. *Science, 128,* 61–67.

House, B. J., Hanley, M. J., & Magid, D. F. (1979). A limitation on the law of effect. *American Journal of Mental Deficiency, 84,* 132–136.

McIlvane, W. J., & Dube, W. V. (1992). Stimulus control shaping and stimulus control topographies. *The Behavior Analyst, 15,* 89–94.

McIlvane, W. J., Kledaras, J. B., Dube, W. V., & Stoddard, L. T. (1989). Automated instruction of severely and profoundly retarded individuals. In J. A. Mulick & R. F. Antonak (Eds.), *Transitions in mental retardation* (Vol. 4, pp. 15–76). Norwood, NJ: Ablex.

Migler, B. (1964). Effects of averaging data during stimulus generalization. *Journal of the Experimental Analysis of Behavior, 7,* 303–307.

Ray, B. A., & Sidman, M. (1970). Reinforcement schedules and stimulus control. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 187–214). New York: Appleton-Century-Crofts.

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

Saunders, K. J., & Spradlin, J. E. (1989). Conditional discrimination in mentally retarded adults: The effects of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior, 52,* 1–12.

Saunders, K. J., & Spradlin, J. E. (1990). Conditional discrimination in mentally retarded adults: The development of generalized skills. *Journal of the Experimental Analysis of Behavior, 54,* 239–250.

Serna, R., & Carlin, M. T. (2000). Guiding visual attention in individuals with mental retardation. In L. M. Glidden (Ed.), *International review of research in mental retardation* (Vol. 24, pp. 321–357). New York: Academic Press.

Sidman, M. (1969). Generalization gradients and stimulus control in delayed matching-to-sample. *Journal of the Experimental Analysis of Behavior, 33,* 285–289.

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 discrimination. *Journal of the Experimental Analysis of Behavior, 58,* 173–182.

Stoddard, L. T. (1982). An investigation of automated methods for teaching severely retarded individuals. In N. R. Ellis (Ed.), *International review of research in mental retardation* (pp. 163–207). New York: Academic Press.

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.

Zeaman, D., & House, B. J. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency, psychological theory and research* (pp. 159–223). New York: McGraw-Hill.

Received October 10, 1995
Final acceptance November 1, 2001