BRIEF REPORT

A COMPARISON OF EXCLUSION AND TRIAL-AND-ERROR PROCEDURES: PRIMARY AND SECONDARY EFFECTS

Cristiana Ferrari, Júlio C. de Rose, and William J. McIlvane

UNIVERSIDADE ESTADUAL PAULISTA AT MARÍLIA, MARÍLIA, BRAZIL, UNIVERSIDADE FEDERAL DE SÃO CARLOS, SÃO CARLOS, BRAZIL, AND E. K. SHRIVER CENTER, UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL

The arbitrary matching to sample procedure is widely used to teach conditional relations between stimuli: In the presence of a sample stimulus, the student learns to select a particular comparison stimulus from an array of two or more stimuli. Such procedures are used not only in basic laboratory studies (e.g., Sidman & Tailby, 1982) but also in special education, e.g. in "fixed trial" training procedures used with children with autism (Maurice, Green, & Luce, 1996). Research has shown that many children with developmental disabilities have considerable difficulties in learning elementary discrimination performances and may need special procedures to do so (e.g., Schilmoeller, Schilmoeller, Etzel, & LeBlanc, 1979; Dube, Iennaco, & McIlvane, 1993; McIlvane, Kledaras, Iennaco, & Stoddard, 1990; Saunders & Spradlin, 1989; Zygmont, Lazar, Dube, & McIlvane, 1992). One such procedure is "learning by exclusion" (Dixon, 1977; McIlvane & Stoddard, 1981).

The exclusion procedure uses a defined comparison stimulus as a prompt to teach a relation between an undefined sample and an undefined comparison. The undefined comparison stimulus is displayed together with the one already defined, and the undefined sample is presented. The term *defined* is used to designate stimuli that have already been related to a sample or comparison in the participant's matching-to-sample history (i.e., defined within the operative reinforcement contingencies). The term *undefined* is used to designate stimuli that do not yet have such a history. Human participants virtually always respond to the undefined comparison in the presence of the undefined sample, and that experience often results in very rapid learning. For example, if the sample is a dictated name, a small number of trials (even one) may be sufficient to teach the participant to produce the dictated word as a name for the comparison stimulus (de Rose, de Souza, & Hanna, 1996; Ferrari, de Rose, & McIlvane, 1993; McIlvane, Bass, O`Brien, Gerovac, & Stoddard, 1984).

The conditions under which exclusion produces learning of new conditional and naming relations are incompletely understood and few studies have endeavored to assess the relative advantages of the procedure. Ferrari et al. (1993) compared learning by exclusion with trial-and-error learning in otherwise typically developing

These data were presented in a dissertation submitted by the first author to Universidade de São Paulo in partial fulfillment of the requirements for the doctoral degree. We gratefully acknowledge the support from Regina Miura and Centro de Orientação Educacional from Universidade Estadual Paulista–Marilia.

This research was supported in part by a grant from the Brazilian Ministry of Science and Technology, MCT/PRONEX/FINEP # 7697105600. The second author was also supported by a research productivity grant from the Brazilian National Research Council, CNPq. Manuscript preparation was supported in part by NICHD grant HD25995 and FAPESP, Grant #03/09928-4. Statements made in this article reflect the authors' own views and do not constitute policy statements from their universities or funding agencies.

Correspondence concerning this article should be addressed to Cristiana Ferrari, Departamento de Fonoaudiologia, Universidade Estadual Paulista at Marilia - Av. Hygino Muzzi Filho 737, Marília, São Paulo, Brasil, 17.525-900.

children that had histories of persistent school failure; they showed that exclusion produced faster and more reliable learning of new conditional discriminations and new naming relations. The superiority of the exclusion procedure in that study was all the more impressive because the investigators had sought to teach four new conditional relations simultaneously in each conditional discrimination problem; this procedural feature permitted a direct assessment of the relative contribution of exclusion per se as distinct from other variables that are inherent in trial-and-error procedures (e.g., the need to introduce more than one new relation at a time; see Ferrari et al. 1993 for further details).

As yet, there has not been a successful comparison of exclusion and trial-and-error procedures in children with intellectual disabilities. Cameron, Stoddard and McIlvane (1993) implemented a design similar to that of Ferrari et al. (1993) using children with autism and severe intellectual disabilities. Although all of the children showed perfect exclusion, learning outcomes were poor with both procedures, likely due to introducing too many new conditional relations simultaneously (cf. Wilkinson & Albert, 2001). The present study implemented a design similar to that of Ferrari et al. (1993) with two children with Down Syndrome. These children were higher functioning than those of Cameron et al. (1993), and it was hoped that floor effects could be avoided. The design also allowed certain other procedural comparisons that had not been accomplished in the earlier studies.

METHOD

Participants

Two female teenagers with Down Syndrome participated. Deb's age was 14:7 and Mari's was 15:3. Both participants spoke Portuguese. The Illinois Test of Psycholinguistics Abilities (Kirk, McCarthy, & Kirk, 1968), adapted to Portuguese (Bogossian & Santos, 1977) yielded psycholinguistic-age scores of 5:3 and 6:6, respectively, thus suggesting moderate-to-severe intellectual disabilities.

Setting and Equipment

The experiment was conducted in a quiet room at the Center for Educational Orientation, which is a unit for educational services belonging to Universidade Estadual Paulista at Marilia. The participants attended this Center four days each week. The room contained an IBM-compatible microcomputer with a multimedia card. On the table and within reach of the participants were a 14-inch monitor and keyboard. Software controlled presentation of visual stimuli on the monitor and auditory stimuli through headphones. Visual stimuli were line drawings of approximately 2 x 2 cm in white over a dark background; drawings appeared in five locations on the screen. Auditory stimuli were 1-3 syllable nonsense words, phonologically similar to Portuguese words (see Figure 1 for examples of stimuli and schematic representation of the procedures).

Participants responded to visual stimuli sometimes by selecting them and sometimes by naming them. They selected stimuli by pressing keys on the numeric keypad of the computer's keyboard. The software recognized input from keys 8, 2, 4, and 6, which corresponded to defined positions on the computer screen; these keys were covered with tape and had arrows pointing up, down, left, and right, respectively.

Participants sat on a chair, facing the keyboard. The experimenter sat to the right and made a written transcript of naming responses. For reliability scoring, an independent observer also recorded naming responses during 30% of the sessions.

Procedure

The experimental tasks were auditory-visual matching to sample and picture naming. Auditory-visual matching trials began with the presentation of four pictures as comparison stimuli. After an interval of 3 s, the computer's headphones presented the sample. Selections of one of the comparison stimuli then produced differential consequences for correct or incorrect responses and a 2-s intertrial interval. The experimenter verbally instructed the participant to select pictures only after a word had been dictated. Had participants responded prematurely, the sample would have been delayed for 3 s (cf. McIlvane, Kledaras et al, 1990), but they never did so. Naming trials presented a picture on the center of the screen. The experimenter asked the participant to name the picture.

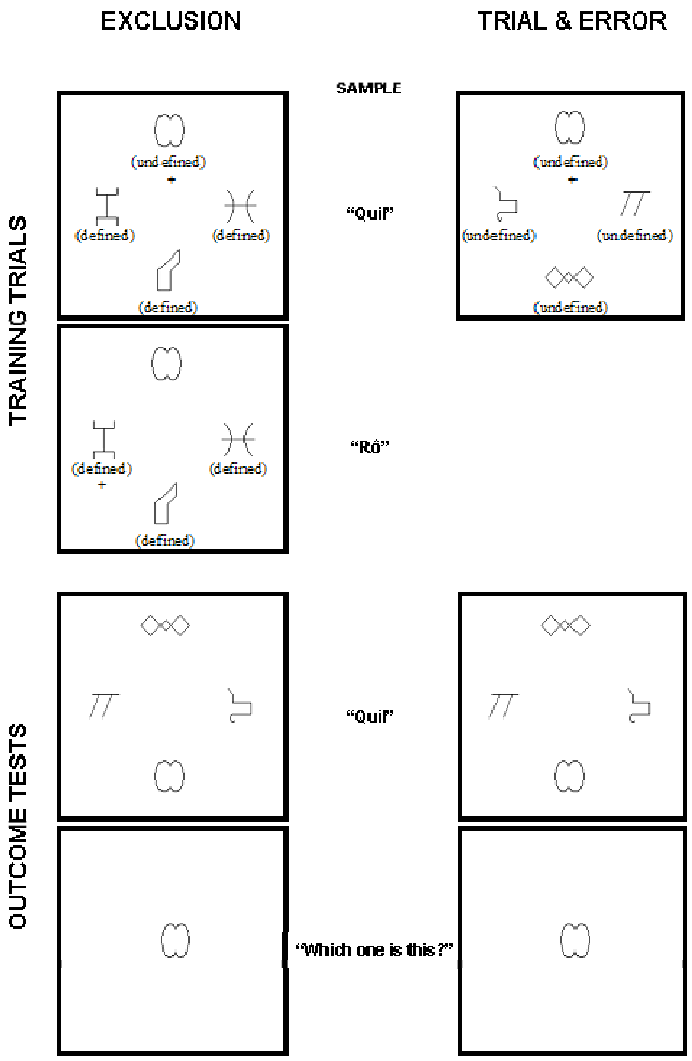

Figure 1

Pretraining. During an initial pretraining phase, participants were taught a matching-to-sample task with familiar stimuli. They learned to match four common pictures (girl, dog, house and fish) to their dictated Portuguese names. Blocks of eight trials, two with each sample, repeated until participants selected the correct picture on all trials of a block. Correct selections were always followed by reinforcing consequences during this and a subsequent pretraining phase; incorrect selections were followed only by the next trial. The reinforcing consequences were points, later exchanged for money (the equivalent of one US cent per point). Participants could earn a maximum of two US dollars per session. Some sessions contained trials without differential consequences (see below). Such sessions concluded with trials of a well-learned task in which participants could rapidly accumulate points (cf. Sidman & Tailby, 1982).

In a second pretraining phase, participants were taught a baseline of three arbitrary relations between nonsense dictated words and arbitrary pictures. These would serve as defined samples and comparisons for later exclusion problems. The teaching procedure, adapted from Saunders and Spradlin (1989), presented blocks of trials, alternating several consecutive trials with one sample ("quita") with the same number of consecutive trials with a second sample ("rô"). Only the comparisons corresponding to "quita" and "rô" were presented on these trials. The

number of consecutive trials with each sample decreased gradually from six to two, and then trials with "quita" and "rô" irregularly alternated. This same procedure was repeated with "rô" and a third sample, "chivata," and then with "quita" and "chivata," and their corresponding comparisons. The last trial block presented 12 trials with all three samples and comparisons, with the correct comparison varying unpredictably from trial to trial. Throughout this pretraining, the criterion to advance from one block to the next was 100% correct.

In a third pretraining phase, participants were adapted to the absence of differential consequences for performance: The experimenter told them that the computer would no longer "tell" them whether responses were correct or incorrect. Six matching trials followed – two with each sample presented in a quasi-randomized sequence. The next six trials tested naming of the pictures. Each picture appeared on two trials, also presented in a quasi-randomized sequence. Any incorrect responses in these pretraining blocks resulted in repetition of a similar block with differential consequences. There followed a similar block without differential consequences. This cycle repeated until participants scored 100% correct on matching and naming trials without differential consequences.

Experimental problems. For ten subsequent experimental problems, participants were exposed to a series of conditional discriminations with novel sample and comparison stimuli. During each problem, the training sought to teach participants to relate four undefined nonsense-words samples to four undefined arbitrary-picture comparisons. The teaching procedure was exclusion for five of these problems and trial-and-error training for the other five problems. The order of exclusion and trial-and-error problems varied between participants and is reported with the results.

Each of these problems had three training sessions. Training sessions always had 32 training trials followed by 15 outcome test trials (see below). During trial-and-error training, all problems presented the four undefined comparisons, and each undefined nonsense word appeared as a sample on eight trials. The sample and correct comparison varied unsystematically over trials. Prompts for correct selections were given in the first session of the first problem with trial-and-error training. On the first trial with each sample, the experimenter pointed to correct comparison and said the Portuguese equivalent of "this one." First-trial prompts were given also with certain other trial-and-error problems.

Exclusion training consisted of 16 exclusion trials intermixed with 16 control trials. Exclusion trials displayed the three defined comparisons taught in pretraining together with an undefined comparison (one of four to be taught in that particular phase). The sample was a corresponding undefined nonsense word. Control trials also presented three defined comparisons and an undefined one. The sample on control trials was always one of the three defined nonsense words from the initial baseline. Control trials were a necessary part of training because they ruled out the possibility of correct responding based solely on the novelty of stimuli. Therefore, exclusion sessions presented undefined stimuli four times each, whereas trial-and-error problems presented each undefined stimulus eight times. However, the number of training trials in both conditions was the same.

Outcome tests were similar for exclusion and trial-and-error sessions. Eight matching outcome test trials followed the training trials. These outcome trials ascertained whether participants had learned to select previously undefined pictures conditionally upon undefined words without the availability of defined comparison stimuli to exclude. On four trials, the four undefined comparison stimuli were presented; the four undefined sample stimuli were presented on one trial each. The remaining four trials were like control trials in the exclusion training. These trials presented the three initially defined comparison stimuli together with one initially undefined, and the sample stimulus was one of the three initially defined words.

Seven naming outcome trials followed. The four initially undefined pictures appeared on one trial each, intermixed with three other naming trials displaying the defined baseline stimuli.

RESULTS

Reliability on naming trials was calculated as follows: agreements/ (agreements+ disagreements) x 100. Agreements were scored when both observers recorded (1) the same word, (2) an unintelligible response, or (3) no response. The mean reliability score for sessions with Mari was

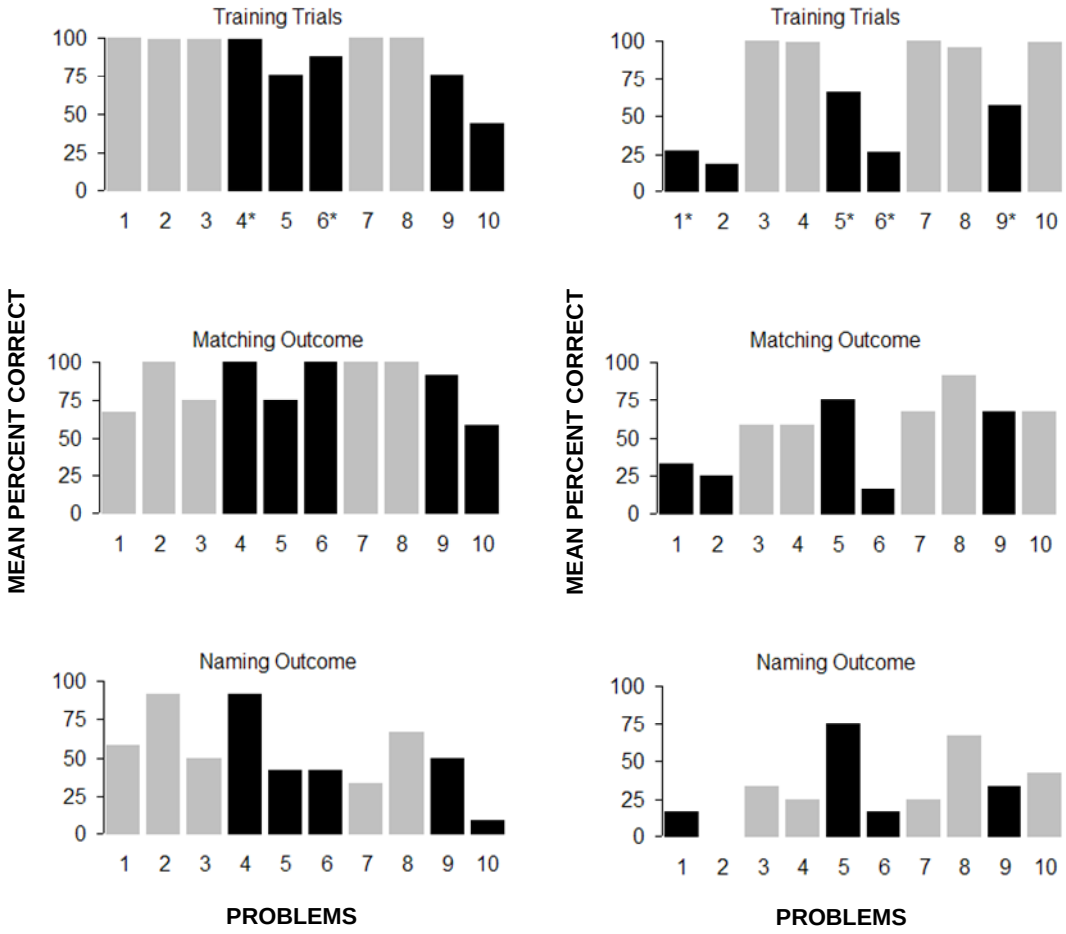

Figure 2

Figure 3

96.4 (four-session sample, range 85.7 to 100). Corresponding scores for Deb, whose speech was less clear, was 85.7 (five-session sample, range 42.9 to 100).

Figures 2 (Mari) and 3 (Deb) show scores on training trials and on matching and naming outcome trials for each of the ten experimental problems. Gray bars show performances during exclusion problems, and black bars show performances during trial-and-error problems. Asterisks indicate problems for which first-trial prompts were given. Both participants responded highly accurately on exclusion training trials. By contrast, both made numerous errors on most of the trial-and-error training trials. These findings were predicted: Exclusion training is a potentially errorless training method whereas errors are expected in trial-and-error procedures.

Concerning the matching and naming outcome tests, there was substantial variability with both procedures. Neither procedure reliably led to impressive learning outcomes. As suggested in the Introduction, we had anticipated that finding. Earlier work had shown that teaching four new relations at a time led to less than optimal learning outcomes even with children who were typically developing (Ferrari et al., 1993). In general, however, exclusion appeared to produce better matching outcome test performances (68.3% vs. 43.3% for Deb, and 88.3% vs. 85% for Mari; chance score: 25%). This was also true on the naming outcome tests (38.3% vs. 28.3% for Deb, and 60% vs. 46.7% for Mari). Thus, the exclusion procedure permitted Mari and Deb to achieve comparable or better learning outcomes

than with trial- and-error procedures while making a smaller number of errors overall.

Another finding seems noteworthy. The two procedures were not equivalent in their effects on the learning of subsequent conditional relations. The left portion of Figure 4 shows that when exclusion procedures were used with a given problem, matching outcome scores on subsequent trial-and-error problems tended to be higher. This finding was particularly marked with Deb. A similar result was observed on naming outcome tests for both participants. The right portion of Figure 4 presents parallel findings for exclusion problems. In this case also, matching and naming outcome scores tended to be higher when a given problem was preceded by a problem trained via the exclusion procedure.

DISCUSSION

Two participants with Down Syndrome excluded defined picture-comparisons on training trials that presented an undefined picture-comparison stimulus and an undefined spoken word as sample. The participants virtually always chose the undefined picture conditionally upon the undefined word. Unlike participants with intellectual disabilities in a previous study (Cameron et al., 1993), those in the present study displayed learning outcomes that were substantially above chance, particularly on the matching outcome tests. Thus, we were successful in replicating systematically the results of Ferrari et al. (1993) with participants with intellectual disabilities. As in that study, the intermediate accuracy scores were likely due to the substantial learning challenges presented (i.e., learning four new conditional relations simultaneously). Had the number of new conditional relations to be learned been reduced, the learning outcomes would undoubtedly have been better, but we might also have reduced the capability of the experimental design to conduct a valid comparison of exclusion with trial-and-error procedures.

Another finding pointing to the superiority of exclusion procedures was the observed order effect: Whatever the method of training for a given problem, performance was generally better if the preceding one had been an exclusion problem. Why did we observe such order effects? It seems likely that these effects occurred because each training condition established behavior that

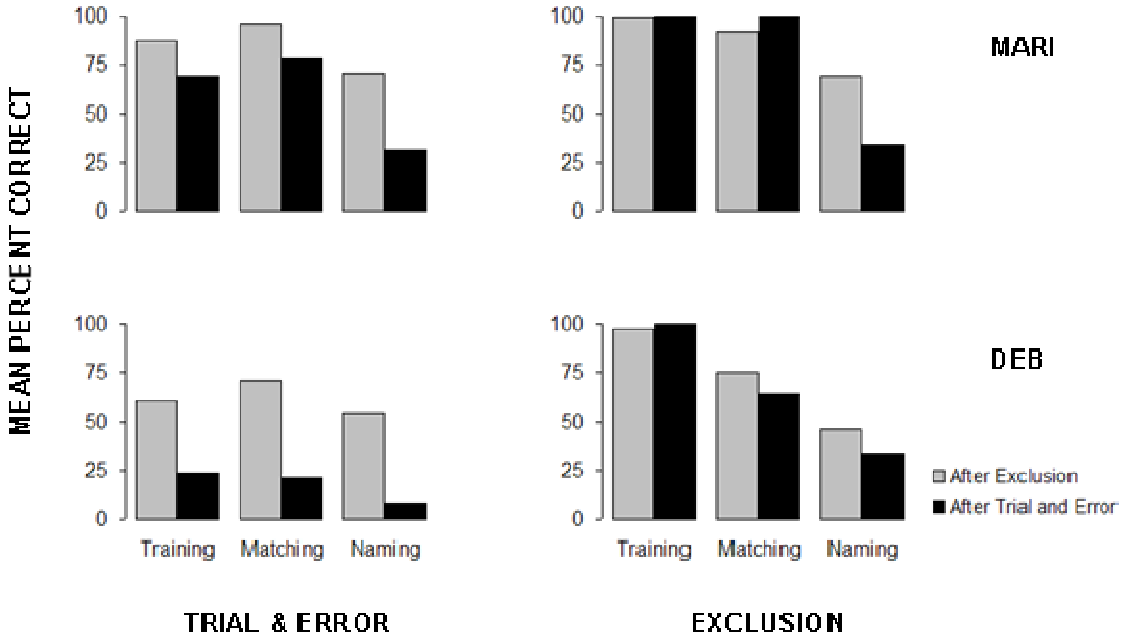

Figure 4

persisted in the subsequent phase. Perhaps the most important behavior may have been continued attending to the sample stimuli and comparison stimuli. Reliable exclusion demonstrates reliable stimulus control by the sample, and such control was evident on all exclusion training trials. By contrast, trial-and-error training procedures do not necessarily encourage attending to the sample and/or comparison stimuli. In the absence of stimulus control by the sample, control by irrelevant stimuli such as position may be inadvertently encouraged, and such irrelevant control may have led to subsequent errors. In other words, undesired topographies of stimulus control (Dube & McIlvane, 1996) may have competed with desired topographies (i.e., attending to relevant sample and comparison stimuli). Perhaps such competition may help to explain why "errors create more errors," as Sidman and Stoddard (1966) suggested long ago.

It is not clear, however, whether the order effects were due to beneficial effects of exclusion or interfering effects of trial-and-error procedures. Either effect would be sufficient to account for the data, or both may have occurred. Further research is needed to clarify this issue, with more variation in the order of conditions and with techniques to observe attending behaviors and to assess stimulus control topographies (e.g., Serna, Wilkinson, & McIlvane, 1998). The study does confirm, however, that exclusion has advantages over trial-and-error procedures for teaching. Moreover, it demonstrates that particular teaching procedures are not independent of one another; stimulus control established via one procedure, whether beneficial or detrimental, may transfer substantially to another.

REFERENCES

Bogossian, M. A. D., & Santos, M. J. (1977). *Manual do examinador: teste de habilidades psicolinguísticas*. [Examiner's manual: test of psycholinguistics abilities.] Rio de Janeiro: EMPSI.

Cameron, M., Stoddard, L. T., & McIlvane, W. J. (1993). A comparison of exclusion vs. selection training in children with severe intellectual disabilities. *Experimental Analysis of Human Behavior Bulletin, 11*, 50-51.

de Rose, J. C., Souza, D. G., & Hanna, H. S. (1996). Teaching reading and spelling: exclusion and stimulus equivalence. *Journal of Applied Behavior Analysis, 29,* 451-469.

Dixon, L. (1977). The nature of control by spoken words over visual stimulus selection. *Journal of the Experimental Analysis of Behavior, 29,* 433-442.

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. (1996). Some implications of a stimulus control topography analysis for emergent behavior and stimulus class. In T. R. Zentall & P. M. Smeets (Eds.) *Stimulus class formation in human and animals* (pp. 197-218). New York: Elsevier.

Ferrari, C., de Rose, J. C., & McIlvane, W. J. (1993). Exclusion vs. selection training of auditory-visual conditional relations. *Journal of Experimental Child Psychology, 56,* 49-63.

Kirk, S. S., McCarthy, J. J., & Kirk, W. D. (1968). *The Illinois test of psycholinguistic abilities*. Urbana: University of Illinois Press.

Maurice, C. , Green, G., & Luce, S. (Eds.) (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals.* Austin, TX: PRO- ED.

McIlvane, W. J., Bass, R. W., O'Brien, J. M., Gerovac, B. J., & Stoddard L. T. (1984). Spoken and signed naming of foods after receptive exclusion training in severe retardation. *Applied Research in Mental Retardation, 5,* 1-27.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., Iennaco, F. M., & Stoddard, L. T. (1990). Teaching relational discrimination to individuals with mental retardation: some problems and possible solutions. *American Journal on Mental Retardation, 95,* 283-296.

McIlvane, W. J., Kledaras, J. B., Lowry, M. J. & Stoddard, L. T. (1992). Studies of exclusion in individuals with severe mental retardation. *Research in Developmental Disabilities, 13,* 509-532.

McIlvane, W. J., Kledaras, J. B., Stoddard, L.T., & Dube, W. V. (1990). Delayed sample presentation in MTS: Some possible advantages for teaching individuals with developmental limitations. *Experimental Analysis of Human Behavior Bulletin, 8,* 31-33.

McIlvane, W. J., Serna, R. W., Dube, W. V., & Stromer, R. (2000). Stimulus control topography coherence and stimulus equivalence: Reconciling test outcomes with theory. In J. Leslie & D. E. Blackman (Eds.) *Issues in experimental and applied analyses of human behavior* (pp. 85-110). Context Press: Reno.

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

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

Schilmoeller, G. L., Schilmoeller, K. J., Etzel, B. C., & LeBlanc, J. M. (1979). Conditional discrimination after errorless and trial-and-error training. *Journal of the Experimental Analysis of Behavior, 31,* 405-420.

Serna, R. W., Wilkinson, K. M., & McIlvane, W. J. (1998). Blank comparison assessment of stimulus-stimulus relations in individuals with mental retardation: A methodological note. *American Journal on Mental Retardation, 103,* 60-74.

Sidman, M., & Stoddard, L. T. (1966). Programming perception and learning for retarded children. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 2, pp. 151-208). New York: Academic Press.

Sidman, M., & Tailby, W. (1982). Conditional discriminations vs. matching-to-sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22.

Wilkinson, K. M., & Albert, A. (2001). Adaptations of "fast mapping" for vocabulary intervention with augmented language users. *Augmentative & Alternative Communication, 17,* 120-132.

Zygmont, D. M., Lazar, R. M, Dube, W. V., McIlvane, W. J. (1992). Teaching arbitrary matching via sample stimulus-control shaping to young children and mentally retarded individuals: a methodological note. *Journal of the Experimental Analysis of Behavior, 57,* 109-117.