JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2011, **44,** 327–340 NUMBER 2 (SUMMER 2011)

COMPARISON OF PROGRESSIVE PROMPT DELAY WITH AND WITHOUT INSTRUCTIVE FEEDBACK

BRIAN REICHOW AND MARK WOLERY

VANDERBILT UNIVERSITY

We examined the effectiveness and efficiency of 2 instructional arrangements using progressive prompt delay (PPD) with 3 young children with autism and 1 child with developmental delays. Specifically, we compared PPD with instructive feedback (IF) to PPD without IF in an adapted alternating treatment design. The results suggested that (a) children with autism and developmental delays can learn when PPD is used with IF, (b) IF can be an effective method of instruction for young children with autism and developmental delays, and (c) the combination of PPD and IF can increase the efficiency of instruction. Data collected 8 to 9 weeks after instruction ended showed that participants maintained mastery of 58% to 92% of the acquired behaviors. We discuss these results within the constraints and limitations of the data and recommend areas for future research.

Key words: autism spectrum disorders, progressive prompt delay, instructive feedback, instructional efficiency, response prompting

Children with autism spectrum disorders (ASD) benefit from organized programs of direct instruction (Eldevik et al., 2009; Reichow & Wolery, 2009), which often include prompting and prompt-removal procedures (e.g., MacDuff, Krantz, & McClannahan, 2001). Researchers define response-prompting procedures by the method of presenting or removing the prompts and have used these procedures to teach individuals of all ages and levels of functioning (e.g., severe mental retardation to typical development), including individuals with ASD (Walker, 2008). Variations of these procedures include constant prompt delay, progressive prompt delay (PPD), system of least prompts, simultaneous prompting, and most-to-least prompts (M. Wolery, Ault, & Doyle, 1992).

Educational researchers can evaluate instructional procedures based on their effectiveness and their efficiency (M. Wolery, in press), with *effectiveness* referring to whether students acquire the instructed behaviors and *efficiency* referring to the rapidity with which learning occurs (M. Wolery et al., 1992; e.g., a comparison of time required by each procedure to teach an equivalent number of skills). For example, Ault, Wolery, Gast, Doyle, and Eizenstat (1988) compared the acquisition of numeral naming with two children with autism by simultaneously teaching numeral pairs using either prompt delay or the system of least prompts. Both teaching procedures resulted in acquisition of numeral naming (i.e., both were effective), but the prompt delay was associated with fewer trials and thus was more efficient.

A different way to evaluate the efficiency of instructional methods is to provide equal intensities of instruction (i.e., the same number of trials) and to measure whether one teaching condition results in a different quantity of acquired skills. A procedure called *instructive feedback* (IF) may result in additional learning efficiency given this conceptualization. IF

This study was conducted by the first author in partial fulfillment of the requirements for the PhD degree from Vanderbilt University.

We thank Ann Kaiser, Craig Kennedy, and Patricia Snyder for their involvement on Brian Reichow's dissertation committee and Hunter Gast and the Franklin Public Schools for allowing us to use their classroom to conduct the study. This project was supported by the U.S. Department of Education Office of Special Education through an ESCE Doctoral Leadership Training Grant (H325D030012).

Address correspondence to Brian Reichow, Yale Child Study Center, 230 South Frontage Road, New Haven, Connecticut 06520 (e-mail: brian.reichow@yale.edu).

doi: 10.1901/jaba.2011.44-327

typically involves adding extra nontarget stimuli to the consequent events of direct instructional trials. Students do not necessarily respond to these extra stimuli, and they do not receive programmed reinforcement for providing a response (Werts, Wolery, Holcombe, & Gast, 1995). For example, a teacher may show a child a picture of a cat and say, ''What's this?'' When the child answers ''cat,'' the teacher might say, ''Great; and cat is spelled c, a, t.'' The teacher's statement, ''Cat is spelled c, a, t,'' is the IF stimulus.

In a review of 23 studies including IF, Werts et al. (1995) concluded that delivering IF stimuli to learners with disabilities, from preschool-aged to adulthood, resulted in acquisition of a majority of the responses to IF stimuli without additional instruction. This finding occurred in one-to-one (M. Wolery et al., 1991) and small-group instruction (Ledford, Gast, Luscre, & Ayres, 2008) as well as when instructional trials were embedded into independent seat work (Caldwell, Wolery, Werts, & Caldwell, 1996). The finding occurred when the instructor presented two IF stimuli for each target stimulus (M. Wolery, Werts, Holcombe, Billings, & Vassilaros, 1993) and when the instructor presented the IF on intermittent trials (Griffen, Schuster, & Morse, 1998).

Researchers have described three variations of IF stimuli in terms of their relation to target stimuli (those taught directly): expansion, novel, and parallel. With expansion, the target and IF stimuli and responses are different from one another, but they are related conceptually. For example, if the target behavior involves reading sight words, the IF may be presenting the spelling of those words (Gast, Doyle, Wolery, Ault, & Baklarz, 1991). In a variation of expansion, target responses and IF stimuli are different from one another and are not conceptually related, but they are in the same curricular domain. For example, if the targets are reading sight words (e.g., ''cat,'' ''dog,'' ''goldfish''), the IF may be reading other sight words that are not related conceptually to the target words but are in the same curricular domain (e.g., ''cow,'' ''pig,'' ''chicken''; T. D. Wolery, Schuster, & Collins, 2000). With novel or unrelated IF, the target and IF stimuli and responses are different from one another, they are not conceptually related, and they are not drawn from the same curricular domain. For example, if the target responses are naming fractions given a shaded portion of a figure, the IF may be naming a state when shown an outline of the state's borders (Werts, Wolery, Holcombe, & Frederick, 1993). With parallel IF, the target and IF stimuli are different from one another, but the target and IF responses are the same. For example, if the target responses are naming Arabic numerals (e.g., 3 and 4), then the IF may be naming the corresponding Roman numerals (e.g., III and IV; Holcombe, Wolery, Werts, & Hrenkevich, 1993).

Although considerable research has demonstrated the utility of IF for participants with disabilities, few studies have evaluated these procedures with individuals with ASD (cf. Ledford et al., 2008; M. Wolery et al., 1991). IF has led to increased efficiency of learning when examined with individuals who have ASD. These preliminary findings warrant additional research on IF with children with ASD. The current study sought to extend the work on IF for individuals with ASD by examining the effectiveness and efficiency of expansion IF. Specifically, we compared the effects of IF to a condition with no IF.

METHOD

Participants and Setting

We selected four participants for inclusion in this study who (a) were 3 to 7 years old, (b) were in attendance at a public preschool program for children with ASD during at least 80% of days over the past 4 months, (c) imitated single words, (d) had identifiable reinforcers, and (e) had no previous exposure

to IF procedures. The public preschool program from which these participants were drawn served children with autism 5 hr per day, 5 days per week. We conducted sessions for all participants in a one-to-one arrangement in an individual work area of the classroom.

Sally was a 3-year-old girl who had been diagnosed by a neurologist as having a developmental delay with characteristics of pervasive developmental disorder. When tested at 29 months of age, her Developmental Assessment of Young Children (DAYC; Voress & Maddox, 1998) age-equivalent scores were 6 to 14 months below her chronological age. She had limited expressive language skills that often consisted of echoing adults' statements or one- to three-word combinations of attributes and nouns.

Amanda was a 5-year-old girl with a diagnosis of autism, visual impairment, and albinism. Results from her diagnostic assessment showed scores in the range of autism for all domains of the Autism Diagnostic Observation Schedule (Lord, Rutter, DiLavore, & Lisi, 1999). When tested at 46 months of age, results of the DAYC showed delays of at least 22 months across subtests. Her expressive language consisted of one- and two-word requests and comments. Prior to inclusion in the study, she could name three sight words (''Amanda,'' ''lunch,'' and ''snack'').

Chris was a 4-year-old boy with a diagnosis of autism. When assessed at 35 months of age, he had delays of at least 12 months on all subtests of the DAYC. His expressive language consisted of one- to two-word combinations that he typically used to request items. He had received one-on-one intervention services based on applied behavior analysis in his home for 18 months at the beginning of the study. He was the only participant with a history of such instruction.

Paul was a 5-year-old boy with a diagnosis of autism. Diagnostic assessment results showed an autism quotient of 96 on the Gilliam Autism Rating Scale (Gilliam, 1995), and results of a Vineland Adaptive Behavior Scale (Sparrow, Balla, & Cicchetti, 1984) conducted when he was 34 months old showed delays of at least 12 months on communication, social, and adaptive behavior. His expressive language consisted of two- to five-word combinations, and he was beginning to form complete sentences. He could read at least 100 sight words and could identify the Arabic numerals. Immediately prior to this study, Paul completed a similar evaluation related to IF; however, we terminated that evaluation due to threats to its internal validity and began the evaluation presented in this paper with novel stimuli.

Materials

We used target stimuli, reinforcers, stopwatches, and data-collection sheets during this evaluation. We presented target stimuli as two-dimensional pictures, symbols, or words. The specific stimuli and target responses varied across participants (Table 1). For Sally and Chris, we used photographs printed on heavy-stock paper (14.6 cm by 11.4 cm) with a glossy finish. No printed words were shown on the stimuli. For Amanda, we printed three- to five-letter words in black ink using 100-point Century Gothic font on flash cards (9.6 by 4.6 cm). For Paul, we printed numerals and colors in 100-point Century Gothic font on flash cards.

We determined initial reinforcers through teacher nomination, and, to prevent satiation, we rotated reinforcer delivery throughout the study according to children's choices. We included bubbles, balloons, and spinning light-up toys and delivered descriptive verbal praise to all participants. We used all reinforcers in each condition to minimize any systematic differences in reinforcer quality across conditions.

Response Measurement, Interobserver Agreement, and Procedural Fidelity

Observers scored unprompted responses as either correct or in error. We defined an

Table 1
Behaviors for Each Participant by Set

Set	Condition	Sally: name pictures	Amanda: name sight words	Chris: name pictures	Paul: name Spanish color or numeral
1	PPD no IF	sink	rat	ticket	negro
		dress	taxi	olive	siete
2	PPD with IF	bowl	fox	magnet	blanco
		gloves	train	lobster	cinco
3	IF	glass	lion	ruler	azul
		coat	bus	pickle	doce
4	PPD no IF	range	fish	ladder	rojo
		tie	car	rooster	uno
5	PPD with IF	pot	dog	easel	marron
		yarn	ship	lettuce	diez
6	IF	fridge	cow	toolbox	verde
		belt	van	gecko	ocho
7	Control	fan	kite	compass	gris
		quilt	ant	garlic	cuatro

unprompted correct response as the child saying the correct answer to the task direction (e.g., "What's this word?") during the delay interval before the delivery of the controlling prompt. We defined an *unprompted error* as the child saying anything other than the correct answer to the task direction during the delay interval before the delivery of the controlling prompt. We used unprompted correct responses, which illustrate the transfer of stimulus control when using the PPD procedure (M. Wolery et al., 1992), to make all data-based decisions.

Observers also scored student behavior after the delivery of the controlling prompt. Prompted responses were those that occurred after the controlling prompt (instructional sessions only) and included (a) prompted correct, (b) prompted error, and (c) no response. We defined a *prompted correct* response as a correct imitation produced by the child within 5 s of an instructor model (i.e., controlling prompt); a *prompted error* response was defined as a child saying anything other than the instructor's model within 5 s. *No response* was defined as the participant not saying anything within 5 s of the controlling prompt.

A graduate student in special education collected reliability data simultaneously but independently to assess interobserver agreement and procedural fidelity (Billingsley, White, & Munson, 1980) for at least 20% of sessions for each participant and for each phase of the study. We assessed interobserver agreement on the participant's responding on each trial, which was calculated by dividing the number of agreements (i.e., the number of trials coded the same by the raters) by the number of agreements plus disagreements (i.e., trials coded differently by the raters) and converting the quotient to a percentage (Kennedy, 2005). Agreement averaged 99.6% across participants (range across all sessions by participant, 91% to 100%).

We scored eight experimenter responses on each trial to assess procedural fidelity: (a) securing the child's attention, (b) using the correct stimulus, (c) delivering the task direction, (d) providing the appropriate delay interval, (e) delivering the controlling prompt, (f) providing contingent positive reinforcement on the correct schedule, (g) delivering the IF, and (h) maintaining a 2- to 5-s intertrial interval. We calculated the percentage of correct implementation by dividing the number of observed experimenter behaviors by the number of planned experimenter behaviors and converting the quotient to a percentage, resulting in a mean fidelity of 99.9% (range across all behaviors by session and participant, 91% to 100%). No consistent procedural errors oc-

curred for any given experimenter behavior or any participant in any condition.

Design

We used an adapted alternating treatments design (Sindelar, Rosenberg, & Wilson, 1985), which permitted the comparison between different procedures through the rapid alternation of PPD no IF and PPD with IF. The design included two comparison phases (i.e., an initial demonstration and subsequent replication within each participant) and four probes.

Procedure

Initial assessment. We conducted three to six initial assessment sessions for each participant before assigning stimuli and beginning data collection. Each session occurred in a one-on-one arrangement in the child's classroom and was 5 to 8 min in duration. The purpose of these sessions was twofold. First, these sessions informed us as to what type of stimuli would be appropriate (e.g., sight words, pictures) for inclusion (e.g., students were able to name multiple stimuli of a certain type, e.g., pictures, but not the specific stimuli included in this study). We selected 12 to 14 stimuli from two or three curricular domains (e.g., Spanish numerals, household appliances, vehicle names) that the participant did not name vocally when shown the card and asked, "What is this?"

We randomly assigned these unlearned stimuli to six or seven stimulus sets for each participant. Each set consisted of two stimuli from different curricular domains (e.g., reading sight words for animals and reading sight words for vehicles). We randomly assigned each set to one of four conditions: (a) target stimuli for PPD no IF, (b) target stimuli for PPD with IF, (c) IF stimuli for PPD with IF, or (d) control. Six stimulus sets were used to allow a within-participant replication of the first three conditions (i.e., PPD no IF, PPD with IF, and IF), and the seventh set was used for the control condition. For participants who began the evaluation with six sets (Sally, Amanda, and Chris), we added a seventh set after the first comparison to serve as the control set. We assessed the difficulty of each stimulus set through a logical analysis of the characteristics of each target behavior (e.g., number of syllables, number of letters, number and type of letter blends) and expert opinion (i.e., teacher, therapist, professor) to approximate equal difficulty across conditions.

Instruction. Each instructional session consisted of 12 trials (six trials of each target in one stimulus set). During each trial, the experimenter first delivered an attending cue (e.g., said the child's name) and then waited until the participant responded (e.g., looked at the experimenter). The experimenter presented the target stimulus (e.g., showed the child a flash card on which the sight word [cat] was printed) and delivered the task direction (e.g., said, "What is this?"). The experimenter waited for a response for a designated interval at the conclusion of the task direction (i.e., prompt delay). If the child produced a correct unprompted response, the experimenter delivered descriptive praise (e.g., said, "That's right; this spells 'cat.'") followed by positive reinforcement (e.g., blowing bubbles, activating light-up toys). If the child produced an incorrect unprompted response or did not respond, the experimenter ignored the response and delivered the controlling prompt (e.g., verbal model ["cat"]) at the conclusion of the interval. The experimenter waited for 5 s after the delivery of the controlling prompt. If the participant provided a correct prompted response, the experimenter delivered descriptive praise and an identified reinforcer. If the child provided an incorrect prompted response or did not respond, the experimenter ignored the response; this ended the trial. The experimenter initiated the next trial 2 to 5 s after completion of the trial. Trials for the PPD no IF and PPD with IF conditions differed only with respect to the presentation of the IF stimulus. In the PPD with IF condition, the experimenter presented IF (e.g., showed the

child a flash card with a printed sight word and provided a verbal model of the response [''this says dog'']) after delivering reinforcement for all unprompted and prompted correct responses. The experimenter did not provide IF following an incorrect response. The experimenter did not wait for the participant to respond to the IF stimulus and provided no programmed reinforcers for responding correctly or incorrectly to the IF stimulus. The PPD instructional procedure in each condition included five different delay durations: 0, 1, 2, 3, and 4 s. The 0-s delay remained in effect until the participant completed two consecutive sessions with 100% prompted correct responses. Thereafter, we used each delay interval for two sessions until reaching the 4-s delay. Following a criterion of 100% unprompted correct responding for three consecutive sessions at the 4-s delay, we then altered the schedule of reinforcement to a variable-ratio (VR) 3 schedule (range, 2 to 4) to help to ensure fluency and promote maintenance. This condition remained in effect until the participant met the terminal criterion of two consecutive sessions with 100% unprompted correct responses on the VR 3 schedule.

We conducted up to four instructional sessions per day that were separated by at least 1 hr. If the participant had not reached the criterion level during any given day, we conducted an equal number of PPD no IF and PPD with IF sessions on that day. When a participant met criterion in one condition but not the other (e.g., Stimulus Set 2 for Amanda), we conducted multiple sessions (two to four) of the condition that had not reached the criterion level. We also conducted four review trials of the set at criterion at the conclusion of every other instructional session for the set not at criterion. Review trials had the same format as instructional trials.

Probes. We conducted four probes (Probe 1 [baseline], Probe 2, Probe 3, and maintenance), which differed only in terms of their temporal relation in the study and by the number of stimulus sets assessed. We assessed baseline performance across stimulus sets during Probe 1. After Probe 1, we taught participants in daily sessions to name the stimuli from two target sets with one set taught with PPD no IF and one set taught with PPD with IF, as described previously. We presented a third stimulus set as the IF in the PPD with IF procedure. After participants demonstrated criterion-level performance on the two target stimulus sets, we conducted a second probe to ascertain acquisition of the third (IF) set and continued baseline levels of the other stimulus sets. After the second probe, we replicated these procedures with the fourth, fifth, and sixth stimulus sets. Again, when children demonstrated criterion-level performance with these target sets, we conducted a third probe. We conducted a final probe 8 to 9 weeks after completion of this evaluation to assess maintenance. All probes assessed a control stimulus set, which permitted the detection of maturation and history effects.

During probes, we presented the stimulus sets in a predetermined random order such that each session had 12 trials (three trials of each stimulus from two sets). We presented the stimuli within a set together, probed each stimulus set during three nonconsecutive sessions, and assessed different combinations of stimulus sets in each session. Probe trials had a similar format as the training trials with a 5-s response interval described previously; however, the experimenter did not provide any controlling prompts or IF. In addition, we presented trials of previously mastered stimuli approximately every third trial (range, once every two to four trials) during probe sessions in which it was unlikely for the participant to name any stimuli correctly (e.g., during the baseline probes or during subsequent probes with stimulus sets not yet presented in instructional trials). Trials in which previously mastered stimuli were presented had the same format as probe trials and included stimuli that the

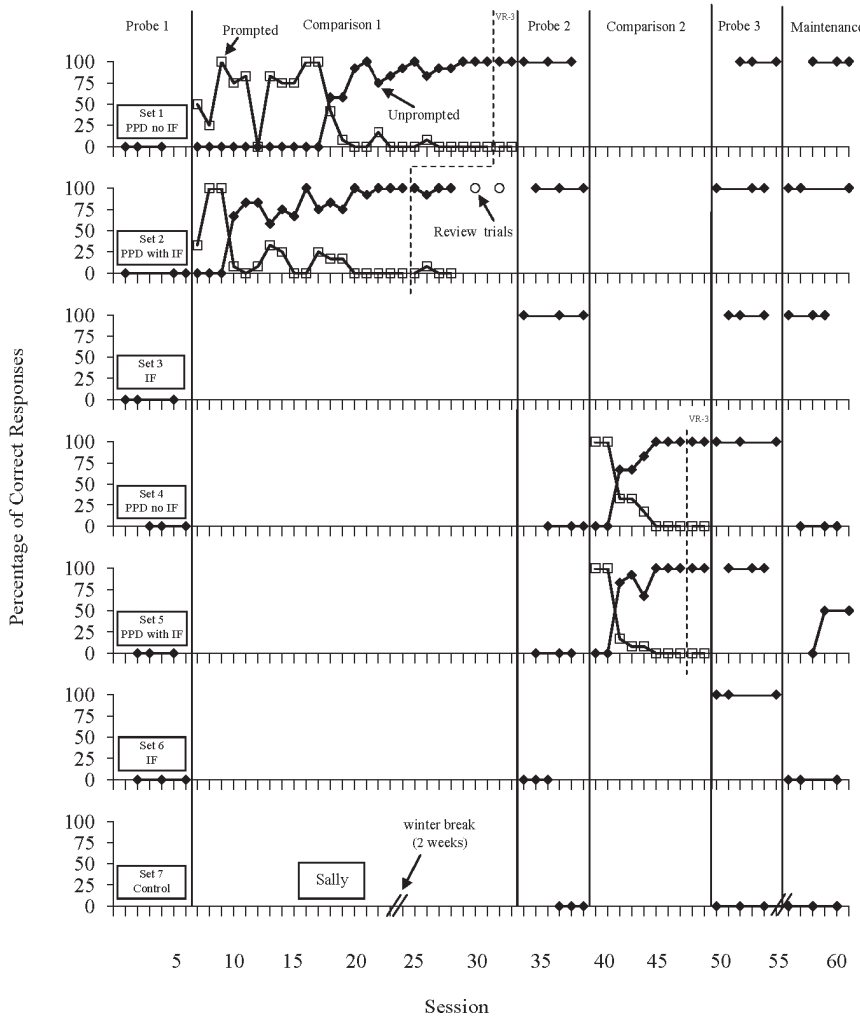

Figure 1. Percentage of unprompted and prompted correct responses for each stimulus set and condition (open circles represent review trials) for Sally.

participant named correctly during the initial assessment sessions. We used these trials to provide reinforcement opportunities, thereby minimizing deflated probe performance.

RESULTS

Figure 1 shows the results for Sally. Each panel shows responding with a different stimulus set. In the first phase (Probe 1), Sally was tested on Stimulus Sets 1 through 6 under baseline conditions. Correct responding was at zero across all sessions and stimulus sets. In the second phase, Stimulus Set 1 was exposed to PPD no IF, and criterion-level performance (three consecutive sessions with 100% correct responding) was reached in 24 sessions, for a total of 108 min. Stimulus Set 2 was exposed to PPD with IF, and criterion-level performance was reached in 16 sessions that lasted a total of 96 min 24 s. In the third phase (Probe 2), Stimulus Sets 1 through 7 were presented under the baseline condition, and responding was at criterion level for Stimulus Sets 1 through 3,

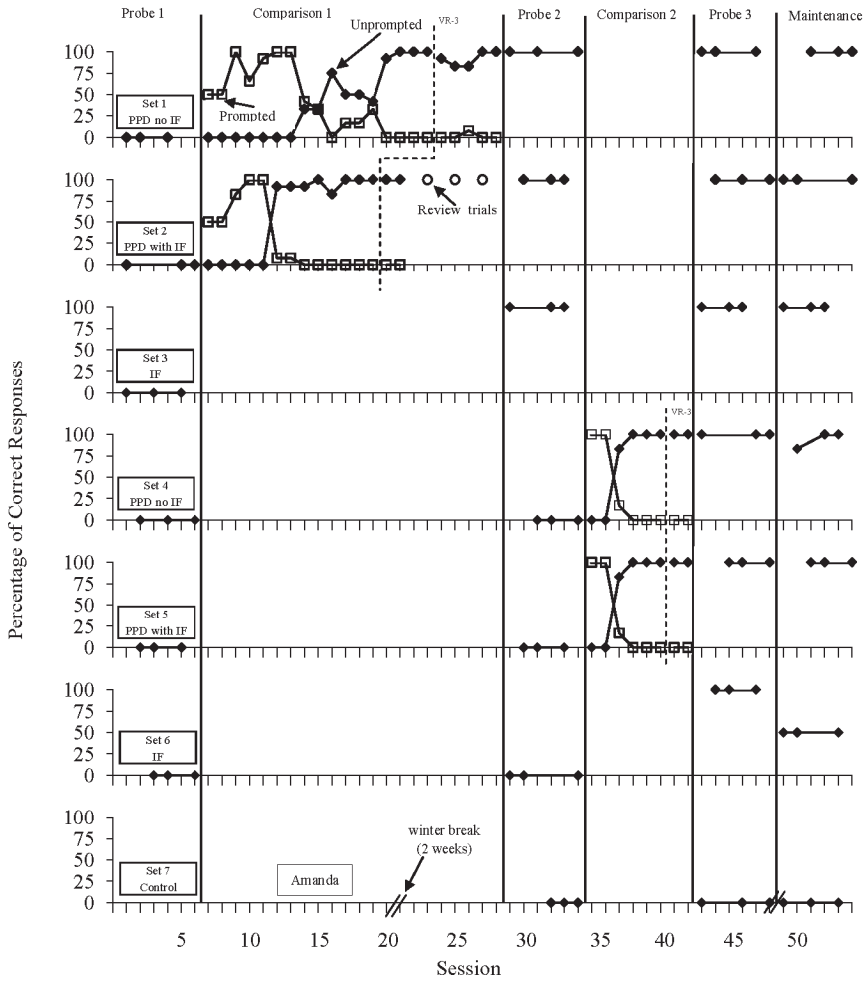

Figure 2. Percentage of unprompted and prompted correct responses for each stimulus set and condition (open circles represent review trials) for Amanda.

which were exposed to PPD no IF, PPD with IF, and IF, respectively. In contrast, correct responding was at zero for Stimulus Sets 4 through 7, which had not been not exposed to any training. These results indicated that PPD was equally effective regardless of whether it was combined with IF (Stimulus Set 2) or not (Stimulus Set 1). Furthermore, adding IF resulted in mastery of Stimulus Set 3. These results were replicated in the fourth (Comparison 2) and fifth (Probe 3) phases. During the maintenance phase, Sally displayed criterion-level performance for the stimulus sets trained first (Stimulus Sets 1 through 3). However, correct responding was much lower on the stimulus sets trained in Comparison 2 (0% for Stimulus Sets 4 and 6; 33% for Stimulus Set 5). Correct responding remained at zero for Stimulus Set 7, which was not exposed to training.

Results for Amanda (Figure 2) were similar to those obtained with Sally. Amanda reached criterion-level performance with PPD no IF (Stimulus Set 1) in 22 sessions that lasted 71 min 12 s. She reached this criterion with PPD with IF (Stimulus Set 2) in 15 sessions that lasted 54 min 48 s. She displayed criterion-level performance for Stimulus Sets 1 and 4

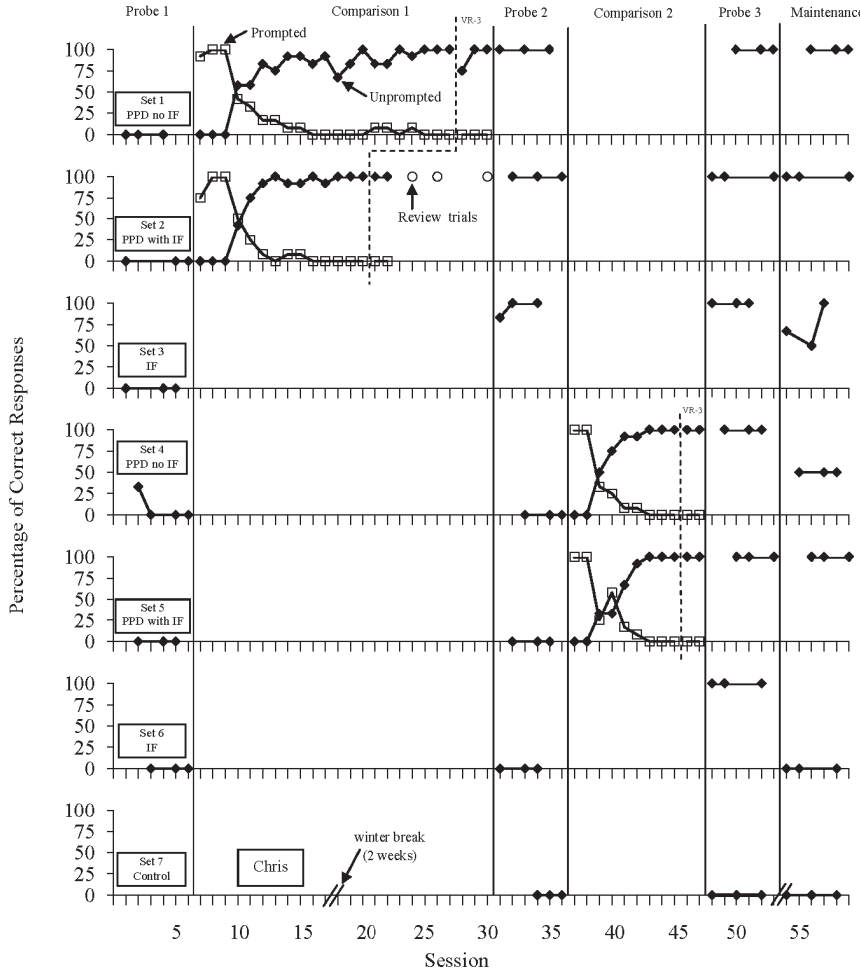

Figure 3. Percentage of unprompted and prompted correct responses for each stimulus set and condition (open circles represent review trials) for Chris.

during the probe phases that immediately followed training with PPD no IF, and she similarly showed criterion-level performance for Stimulus Sets 2, 3, 5, and 6 after exposure to PPD with IF. One difference (from Sally's results) was that Amanda responded at or near the criterion level during the maintenance phase for Stimulus Sets 1 through 5 and showed a decrement in performance only for Stimulus Set 6, which was exposed to IF. Correct responding remained at zero for Stimulus Set 7, which was not exposed to training.

Results for Chris (Figure 3) were similar to those obtained with Sally and Amanda. Chris reached criterion-level performance with PPD no IF (Stimulus Set 1) in 27 sessions that lasted 80 min 24 s. He reached this criterion with PPD with IF (Stimulus Set 2) in 22 sessions that lasted 77 min 36 s. He displayed criterion-level performance for Stimulus Sets 1 and 4 during the probe phases that immediately followed training with PPD no IF, and he similarly showed criterion-level performance for Stimulus Sets 2, 3, 5, and 6 after exposure to PPD with IF. During the maintenance phase, he displayed criterion-level performance for the first stimulus sets trained first (Stimulus Sets 1 through 3). However, like Sally, correct re-

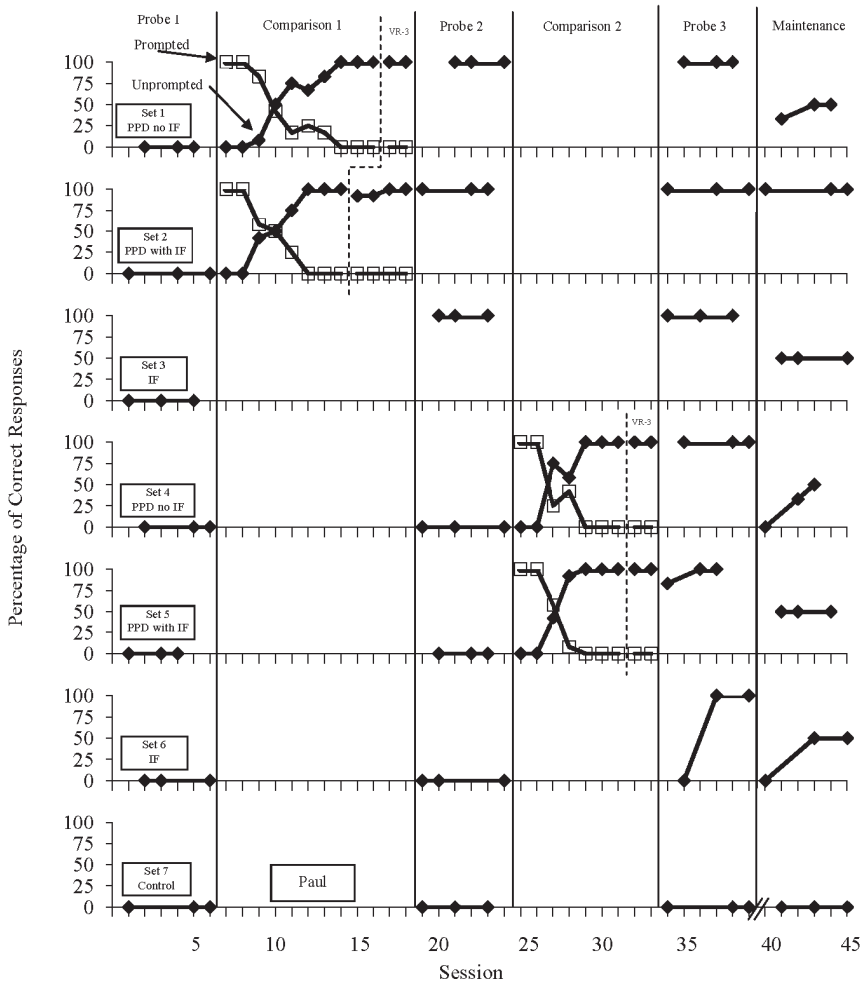

Figure 4. Percentage of unprompted and prompted correct responses for each stimulus set and condition (open circles represent review trials) for Paul.

sponding was much lower for the stimulus sets trained in Comparison 2 (50% for Stimulus Set 4, 100% for Stimulus Set 5, and 0% for Stimulus Set 6). Correct responding remained at zero for Stimulus Set 7, which was not exposed to training.

Results for Paul (Figure 4) were similar to those obtained for other participants. Paul reached criterion-level performance with PPD no IF (Stimulus Set 1) in 12 sessions that lasted 46 min 36 s. He reached this criterion with PPD with IF (Stimulus Set 2) in 12 sessions that lasted 57 min 12 s. He displayed criterion-level performance for Stimulus Sets 1 and 4 during the probe phases that immediately followed training with PPD no IF, and he similarly showed criterion-level performance for Stimulus Sets 2, 3, 5, and 6 after exposure to PPD with IF. His maintenance data were variable. Stimulus Set 2 demonstrated maintenance with 100% correct responding. All other stimulus sets that were taught (1, 3, 4, 5, and 6) were associated with maintenance at or below 50%. Correct responding remained at zero for Stimulus Set 7, which was not exposed to training.

To determine the efficiency of PPD with IF relative to PPD no IF, we compared the two

conditions in terms of the number of sessions and amount of time that the participants required to reach criterion-level performance. During Comparison 1, participants reached this criterion in fewer sessions for PPD with IF (M = 16.3; range, 12 to 22) than for PPD no IF (M = 21.3; range, 12 to 27). However, the mean duration of the sessions was longer for PPD with IF (M = 4 min 30 s; range, 3 min 30 s to 6 min) than for PPD no IF (M = 3 min 24 s; range, 2 min 54 s to 4 min 30 s). To reconcile these differences, we calculated the total length of time allocated to teaching with PPD and IF (mean of 16.3 sessions multiplied by 4.5 min per session) divided by the number of acquired skills (four) relative to the length of time committed to teaching during PPD no IF (mean of 21.3 sessions multiplied by 3.4 min per session) divided by the number of skills acquired during that time (two). In this regard, the mean number of minutes of teaching per acquired skill was 18.4 min during PPD with IF and 36.2 min during PPD no IF. Thus, the IF component resulted in much greater instructional efficiency.

For the second comparison, all participants reached the criterion level in the same number of sessions for both conditions (i.e., M = 9.8 sessions). When we examined efficiency in terms of the number of minutes per condition, PPD with IF had a slightly longer mean duration (37.2 min) than PPD no IF (31.2 min). Although these two measures did not favor PPD with IF, more targets were acquired during PPD with IF. Thus, it is prudent to examine the number of sessions and minutes per target. Again, when we compared the mean sessions per target (PPD with IF = 2.5 vs. PPD no IF = 4.9) and the mean minutes per target (PPD with IF = 9.3 vs. PPD no IF = 15.6), PPD with IF resulted in more efficient learning.

DISCUSSION

Previous research has documented the utility of PPD for individuals with ASD (Walker, 2008), but the use of IF in instructional trials with children with ASDs has not been thoroughly examined (Werts et al., 1995). The current results showed that (a) PPD with IF was as effective as PPD no IF, (b) PPD with IF doubled the number of words learned when compared with PPD no IF, and (c) PPD with IF was about twice as efficient as PPD no IF in terms of the amount of time required per skill acquired. These results are similar to previous research on IF conducted with other populations (Werts et al., 1995) and provide generality to the conclusion that IF is an effective and efficient instructional practice when educating young children with ASD.

We examined the efficiency of IF by comparing the acquisition of target behaviors when IF was not presented (PPD no IF) and when IF was presented during the consequent event of correct responses (PPD with IF). Because the behaviors associated with the stimuli presented as IF (Sets 3 and 6) demonstrated mastery during the subsequent probes, the participants likely acquired the behaviors through their presentation during the preceding comparisons. Thus, learners could have potentially acquired four behaviors in PPD with IF and two behaviors in PPD no IF.

We can extrapolate these data with reference to two measures to illustrate how this increased efficiency might affect learning in other situations. With reference to the average number of sessions needed to acquire target behaviors, 10 sessions (one session per day for 2 school weeks) would be needed to result in the acquisition of two skills using PPD no IF, but four skills could be taught over the same number of sessions using PPD with IF. With reference to the amount of instructional time needed to teach four behaviors, approximately 145 min would be required when using PPD no IF, and approximately 74 min would be necessary when using PPD with IF.

Although the results from the present study are robust, one must take into consideration certain limitations when interpreting the find-

ings. We conducted this study under one set of conditions using a small sample of participants with developed verbal and imitative repertoires. These factors limit the generality of the findings to the larger population of children with autism, and it is unclear if one could achieve replication under different circumstances (e.g., group instructional format, older children, nonimitative children, children with more severe autism). Future research should examine participant attributes and environmental conditions to determine the necessary elements for the greatest probability of success and to whom and what conditions generalizations can be made.

The data from the third probe showed maintenance of previously acquired behaviors across the 2-week period during which they were not presented (i.e., during the second comparison). Data from the maintenance probe (fourth probe) showed that previously acquired behaviors were maintained after 2 months at a much lower level (an average of 58% to 92% across participants). These levels of maintenance are lower than those obtained in previous studies on IF (e.g., Holcombe et al., 1993); thus, further work is needed to identify methods that would promote response maintenance over extended periods.

Research should continue to examine the efficiency of IF. In this study, PPD with IF provided the opportunity to acquire four target behaviors per condition (two directly instructed and two through IF), whereas PPD no IF provided the opportunity to acquire only two behaviors. Thus, the outcomes of this comparison may have been influenced by a ceiling effect of PPD no IF (i.e., we provided fewer opportunities). Comparisons of teaching PPD with IF (two targets and two IF) to PPD no IF with four target behaviors may offer a more balanced comparison of the efficiency of these procedures.

Future studies should also consider examining IF as part of a standardized curriculum to identify target behaviors. A curriculum would allow the researcher to identify conceptually and sequentially related stimulus classes and relations between stimulus classes and would allow the manipulation of relations to examine conditions under which the procedure is most effective.

One interesting finding of the current study was the more rapid acquisition of the stimuli during the second comparison condition (i.e., mastery in 8 to 11 sessions vs. 15 to 27 sessions during the first comparison). Future research should examine this phenomenon of more rapid learning following the initial exposure to teaching procedures, which is consistent with the theory of learning sets proposed by Harlow (1949). Research is needed to determine whether priming events or other preintervention techniques can be used to provide similar gains in efficiency.

Given that the responding of individuals with ASD often fails to generalize across important contexts, PPD with IF might be a useful instructional arrangement for promoting generalized responding. For example, IF could be used to present multiple exemplars of the target stimuli (e.g., if the target response is to name a picture of a bear [a black bear], the IF stimuli might be pictures of other bears [grizzly, polar, and panda bears]). Similarly, if the behaviors being targeted follow a pattern (e.g., naming numerals 41 though 49, as in Paul's first exposure to the study), then directly teaching some numbers (e.g., 41, 44, and 48) and presenting others as IF (e.g., 42, 43, 47) may allow the student to generate a prescriptive rule (i.e., the 4 should be said "forty," and then the name of the second numeral should be said). This arrangement may produce generalization to those numerals not taught directly or presented through IF (i.e., 45, 46, and 49). Thus, it is possible that individuals with autism could increase the efficiency of learning, conceptualized as greater generalization (M. Wolery et al., 1992), in addition to more rapid learning when using PPD with IF when the target behaviors have similar patterns.

Our procedures did not permit us to determine when the participants acquired the IF responses. It is clear when to introduce a new skill during direct instruction; however, it is unclear when to introduce novel IF. Future research might include probes of the IF stimuli during instruction to examine the acquisition process and thus potentially increase teaching efficiency by allowing a teacher to include additional IF after a student has mastered earlier relations. One previous study included daily probes of the IF responses, and results suggested that students acquired IF stimuli prior to the target responses that received direct instruction (Anthony, Wolery, Werts, Caldwell, & Snyder, 1996).

Finally, the behavioral process by which IF results in acquisition is unclear. In this and nearly all previous studies, each target stimulus was associated with a specific IF stimulus (Werts et al., 1995). One possibility is that the pairing of a specific IF stimulus with a specific target behavior or stimulus resulted in some form of associative learning. Results of a study by Werts, Caldwell, and Wolery (2003), however, are inconsistent with this hypothesis. Multiple IF stimuli were presented after each target response, eliminating a one-to-one correspondence between a given target behavior and a given IF stimulus; nonetheless, students acquired the responses to these IF stimuli. It is also possible that the participants in this study acquired the IF responses due to a repertoire of generalized imitation (Baer, Peterson, & Sherman, 1967) or through observational learning (Bandura, 1977). The evaluation of IF with more novice learners and those with less developed imitative repertoires may help to determine whether imitative repertoires are necessary for the acquisition of IF responses.

REFERENCES

Anthony, L., Wolery, M., Werts, M. G., Caldwell, N. K., & Snyder, E. D. (1996). Effects of daily probing on acquisition of instructive feedback responses. *Journal of Behavioral Education*, *6*, 111–133.

Ault, M. J., Wolery, M., Gast, D. L., Doyle, P. M., & Eizenstat, V. (1988). Comparison of response prompting procedures in teaching numeral identification to autistic subjects. *Journal of Autism and Developmental Disorders*, *18*, 627–636.

Baer, D. M., Peterson, R. F., & Sherman, J. A. (1967). The development of imitation by reinforcing behavioral similarity to a model. *Journal of the Experimental Analysis of Behavior*, *10*, 405–416.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Billingsley, F. F., White, O. R., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment*, *2*, 229–241.

Caldwell, N. K., Wolery, M., Werts, M. G., & Caldwell, Y. (1996). Embedding instructive feedback into teacher-student interactions during independent seat work. *Journal of Behavior Education*, *6*, 459–480.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology*, *38*, 439–450.

Gast, D. L., Doyle, P. M., Wolery, M., Ault, M. J., & Baklarz, J. L. (1991). Acquisition of incidental information during small group instruction. *Education and Treatment of Children*, *14*, 1–18.

Gilliam, J. E. (1995). *Gilliam autism rating scale*. Austin, TX: Pro-Ed.

Griffen, A. K., Schuster, J. W., & Morse, T. E. (1998). The acquisition of instructive feedback: A comparison of continuous versus intermittent presentation schedules. *Education and Training in Mental Retardation and Developmental Disabilities*, *33*, 42–61.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*, 51–65.

Holcombe, A., Wolery, M., Werts, M. G., & Hrenkevich, P. (1993). Effects of instructive feedback on future learning. *Journal of Behavioral Education*, *3*, 259–285.

Kennedy, C. H. (2005). *Single-case designs for educational research*. New York: Allyn & Bacon.

Ledford, J. R., Gast, D. L., Luscre, D., & Ayres, K. M. (2008). Observational and incidental learning by children with autism during small group instruction. *Journal of Autism and Developmental Disorders*, *38*, 86–103.

Lord, C., Rutter, M., DiLavore, P. C., & Lisi, S. (1999). *Autism diagnostic observation schedule*. Los Angeles: Western Psychological Services.

MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (2001). Prompts and prompt-fading strategies for people with autism. In C. Maurice, G. Green, & R. M. Foxx (Eds.), *Making a difference: Behavioral intervention for autism* (pp. 37–50). Austin, TX: Pro-Ed.

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA Young Autism Project Model. *Journal of Autism and Developmental Disorders*, *39*, 23–41.

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children*, *8*, 67–76.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales*. Circle Pines, MN: American Guidance Service.

Voress, J. K., & Maddox, T. (1998). *Developmental assessment of young children*. Los Angeles: Western Psychological Services.

Walker, G. (2008). Constant and progressive time delay procedures for teaching children with autism: A literature review. *Journal of Autism and Developmental Disorders*, *38*, 261–275.

Werts, M. G., Caldwell, N. K., & Wolery, M. (2003). Instructive feedback: Effects of a presentation variable. *Journal of Special Education*, *37*, 124–133.

Werts, M. G., Wolery, M., Holcombe, A., & Frederick, C. (1993). Effects of instructive feedback related and unrelated to the target behaviors. *Exceptionality*, *4*, 81–95.

Werts, M. G., Wolery, M., Holcombe, A., & Gast, D. L. (1995). Instructive feedback: Review of parameters and effects. *Journal of Behavioral Education*, *5*, 55–75.

Wolery, M. (in press). Studying response prompting strategies. In C. H. Kennedy, C. G. Breen, & C. A. Peck (Eds.), *Context, intervention, and disability: Essays in honor of Thomas G. Haring*. New York: Routledge.

Wolery, M., Ault, M. J., & Doyle, P. M. (1992). *Teaching students with moderate to severe disabilities: Use of response prompting strategies*. New York: Longmann.

Wolery, M., Doyle, P. M., Ault, M. J., Gast, D. L., Meyer, S., & Stinson, D. (1991). Effects of presenting incidental information in consequent events on future learning. *Journal of Behavioral Education*, *1*, 79–104.

Wolery, M., Werts, M. G., Holcombe, A., Billings, S. S., & Vassilaros, M. A. (1993). Instructive feedback: A comparison of simultaneous and alternating presentation of non-target stimuli. *Journal of Behavioral Education*, *3*, 187–204.

Wolery, T. D., Schuster, J. W., & Collins, B. C. (2000). Effects on future learning of presenting non-target stimuli in antecedent and consequent conditions. *Journal of Behavioral Education*, *10*, 77–94.

Received August 20, 2009
Final acceptance July 21, 2010
Action Editor, Jeffrey Tiger