

EVALUATION OF A SELF-INSTRUCTION PACKAGE FOR CONDUCTING STIMULUS PREFERENCE ASSESSMENTS

RICHARD B. GRAFF

NEW ENGLAND CENTER FOR CHILDREN AND
WESTERN NEW ENGLAND COLLEGE

AND

AMANDA M. KARSTEN

WESTERN NEW ENGLAND COLLEGE

Research suggests that inexperienced individuals cannot accurately implement stimulus preference assessments given written instructions alone. Training that includes written instructions supplemented with feedback from a professional with expertise in conducting preference assessments has proven effective; unfortunately, expert-facilitated direct training may not be widely available. In the current study, we used multiple baseline designs to evaluate the efficacy of an antecedent-only self-instructional package to train staff members to implement two methods of stimulus preference assessments. Accuracy was low when participants had access to written instructions alone. When access to enhanced written instructions was provided (i.e., technical jargon was minimized; instructions included pictures, diagrams, and step-by-step examples), inexperienced staff accurately implemented the assessments. Results are discussed in terms of opportunities to disseminate behavior-analytic technologies through self-instruction and print resources.

Key words: self-instruction, staff training, stimulus preference assessments

The discipline of applied behavior analysis involves systematic application of basic principles of behavior to a range of socially important issues, including education and treatment of individuals with special needs. The efficacy of programs designed to teach new skills and decrease problem behavior often relies on a clinician's ability to identify effective reinforcers for the individuals with whom they work. Research has demonstrated that reinforcers identified by conducting pretreatment stimulus preference assessments can be used to teach new skills (e.g., Karsten & Carr, 2009; Wallace, Iwata, & Hanley, 2006) and to

decrease problem behavior (e.g., Fisher et al., 1994). Given the size of the behavioral literature on assessing preferences of individuals with special needs (see Cannella, O'Reilly, & Lancioni, 2005, for a review), surprisingly little research has addressed how to train inexperienced individuals to conduct stimulus preference assessments.

Lavie and Sturmey (2002) trained three teaching assistants to conduct paired-stimulus (PS; Fisher et al., 1992) preference assessments. Teaching assistants had no background in behavior analysis, and all participants reported difficulty identifying reinforcers for students diagnosed with autism spectrum disorders. First, a task analysis of the steps required to implement the PS method of preference assessment was completed, and corresponding target skills were described briefly to the participants. Next, each trainee was given a checklist that described the individual skills that he or she was required to master, and the trainer verbally described each skill. Trainees then watched a videotape that

This study was completed by the first author in partial fulfillment of requirements for the PhD degree in behavior analysis at Western New England College. We express appreciation to Gregory Hanley, Rachel Thompson, and Eileen Roscoe for their helpful comments on an earlier version of this manuscript.

Address correspondence to Richard B. Graff, New England Center for Children, 33 Turnpike Rd., Southborough, Massachusetts 01772 (e-mail: rgraft@necc.org).
doi: 10.1901/jaba.2012.45-69

included demonstrations of assessment trials. Finally, trainees implemented preference assessment sessions with clients as the staff trainer observed and provided feedback on all target skills. If a trainee did not accurately implement at least 85% of the target skills on the checklist for two consecutive sessions, he or she watched the videotape again before implementing another series of trials. All trainees met the mastery criterion in two training sessions or approximately 80 min of training.

Although Lavie and Sturmey (2002) demonstrated that inexperienced staff could be taught to conduct PS preference assessments in about 80 min, training did require the presence of a behavior analyst or other experienced trainer. Live access to an expert may constitute the ideal training scenario, but this arrangement is also impractical (e.g., cost prohibitive) for agencies that do not employ a behavior analyst. In addition to concerns about the accessibility of an expert-facilitated training package for preference assessments, Lavie and Sturmey did not train the complete repertoire for summarizing and interpreting assessment outcomes (e.g., calculate preference hierarchies, select a stimulus for teaching) or evaluate which intervention components were necessary to produce mastery of preference assessment skills.

Roscoe, Fisher, Glover, and Volkert (2006) also evaluated a procedure to train inexperienced staff to conduct PS and multiple-stimulus without replacement (MSWO; DeLeon & Iwata, 1996) preference assessments (here again, trainees were not asked to generate preference hierarchies or interpret assessment outcomes). Participants included four trainees, simulated clients (trained staff who role played what clients might do during a preference assessment), and actual clients (individuals with developmental disabilities). During baseline sessions, trainees were given a blank piece of paper and were instructed to implement a preference assessment. All participants demonstrated low levels of accuracy. Next, trainees received written instructions describing

how to conduct the assessments. Instructions were taken directly from the method section of the seminal journal articles describing PS and MSWO assessments, with several slight modifications (e.g., some technical terms and irrelevant text were eliminated). Trainees had 30 min to read the instructions before conducting preference assessments with simulated clients. Written instructions were not available to trainees during preference assessment sessions, although the participants did have unlimited access to instructions outside the sessions. No trainee met the mastery criterion following instructions-only training.

In the next phase, participants observed their own videotaped performance and experienced either corrective feedback by the trainer (e.g., the trainer pointed out when the trainee presented the wrong stimuli on a trial or did not record data correctly) or contingent money (without corrective feedback) for correct responses. Roscoe et al. (2006) found that performance feedback from a professional who was well trained in conducting stimulus preference assessments was required by all trainees to achieve the mastery criterion.

Roscoe and Fisher (2008) examined whether brief exposure to training procedures would be sufficient for eight inexperienced staff to conduct PS and MSWO assessments. Results suggested that trainees correctly implemented PS and MSWO procedures following feedback and role-playing opportunities delivered in a single 15- to 20-min training session. No specific information was reported on trainees' ability to score or interpret completed preference assessments.

Despite advances in the efficiency of preference assessment training, additional research may be needed to refine and translate published findings for applied settings (Lerman, 2003). Previous research suggests the services of a professional with expertise in conducting preference assessments are required to train practitioners, but this level of direct training is not

widely available. It is not clear how many behavior analysts work in public schools and private agencies, but the number is undoubtedly much lower than the total number of educators and practitioners who serve individuals with special needs. For example, recent data suggest there are 6.7 million students with special needs in public school classrooms (National Center for Education Statistics, 2006), but there are fewer than 9,000 certified behavior analysts in the United States (Behavior Analyst Certification Board, 2011). Practicing behavior analysts and staff trainers may benefit from a set of low-cost, highly portable training resources they can use as a replacement or prelude to expert-facilitated staff training.

One strategy to minimize the need for expert staff trainers is to use antecedent-only training procedures. Previous research has evaluated the effectiveness of different antecedent training strategies, including written instructions and video modeling. Lerman, Swiezy, Perkins-Parks, and Roane (2000), for example, examined the interaction between procedure or skill type and instructional format while teaching three parents basic behavior-management strategies. During baseline sessions, problem behavior was evoked and the parents were instructed to "behave normally." Next, parents received written and spoken instructions about how to respond to their child's behavior. If a parent did not perform the targeted response with at least 75% accuracy during the following session, a feedback phase was implemented. The authors found that no parent implemented all skills correctly after written and spoken instructions alone. However, some combination of written and spoken instructions was sufficient to teach some behavior-management skills to all parents.

In a different approach to antecedent-only training, Moore and Fisher (2007) compared two types of video modeling (i.e., partial models vs. full models) on implementation of functional analysis procedures (FA; Iwata, Dorsey,

Slifer, Bauman, & Richman, 1982/1994). During baseline, three participants were instructed to implement FA procedures based on written instructions that they had previously received. Participants then were exposed to three different training procedures, each associated with one FA condition: spoken instructions, a partial modeling condition (videotapes depicted examples of 50% of all potential therapist behaviors), and a full modeling condition (videotapes contained multiple exemplars of every possible therapist behavior and many different client behaviors). The authors found that spoken instructions with or without partial modeling was insufficient to train staff to accurately implement FA procedures, but that full modeling, with a larger number of relevant discriminative stimuli, led to mastery of targeted skills in the absence of expert feedback.

Although video-modeling interventions might be effective for training staff to implement some behavior-analytic procedures, the feasibility of video-based interventions is limited by the degree to which the necessary technology (e.g., television, video player) is available. Further, video models may or may not be practical to take into the assessment environment when supplemental prompts are needed. Enhanced written instructions (i.e., materials that include minimal technical jargon and diagrams) may offer benefits similar to those described by Moore and Fisher (2007), but with fewer technological demands, greater portability, and lower cost.

Research is needed to identify effective methods for disseminating preference assessments to educators and practitioners who lack access to an expert trainer. The purpose of the current study was to evaluate the effects of a self-instruction package for inexperienced staff to implement, score, and interpret the results of preference assessments. In addition, we sought to isolate the relative contributions of enhanced written instructions and supplemental materials (i.e., data sheets) to training outcomes for a subset of participants.

METHOD

Participants

Participants were 11 teachers who worked at a school for individuals with autism and related developmental disabilities. Eight teachers were recent college graduates, and three (DS, GL, and NI) had worked in an integrated preschool and day-care classroom at the same school for 3 to 5 years. All teachers held a bachelor's degree or a master's degree and did not have previous experience observing or conducting stimulus preference assessments. After the completion of the study, the eight newly hired teachers were required to conduct preference assessments as part of their job responsibilities, but the preschool and day-care workers were not required to do so. Each teacher completed an informed consent process prior to enrollment in the investigation and confirmed that the investigators may use resulting data in presentations and publications.

All teachers completed a pretest to assess their knowledge of stimulus preference assessments. The pretest consisted of 10 questions that included true–false, matching, multiple choice, or fill-in-the-blank items. The criterion for inclusion in the study was a score of 50% or lower on the pretest. The mean pretest score across teachers was 28% correct (range, 0% to 50%).

In addition to teacher participants, three clinicians from the school served as simulated consumers. All simulated consumers held at least a master's degree in special education or applied behavior analysis and had 7 to 20 years of experience conducting preference assessments. All simulated consumers were supplied with scripts that prescribed atypical responses (i.e., responses other than selecting one stimulus) on specific assessment trials. The simulated consumers conducted these scripted trials with teachers to ensure that all teachers experienced a controlled number of assessment trials with atypical responses. Two simulated consumers and three students enrolled in a master's program in applied behavior analysis scored data live during sessions or from videotapes.

During generalization probes, teachers conducted preference assessment trials with actual consumers from day-treatment or residential programs at the school. Actual consumers ranged from 3 to 15 years of age and were included based on absence of serious physical (i.e., restricted mobility of the upper body) or sensory disabilities (e.g., visual impairments), and the absence of severe problem behavior (i.e., aggression, self-injury). The consumers' parents or legal guardians signed consent forms for their children to participate in the study and for experimenters to use their children's preference-related data for presentation and publications.

Setting and Materials

All teachers experienced training sessions and conducted generalization probes individually in classrooms, small treatment rooms, or conference rooms at the school. All rooms contained basic materials required to conduct assessments (e.g., tables and chairs). Prior to each session, the experimenter told teachers which assessment to conduct and supplied all necessary materials (e.g., data sheets, designated pool of eight toys or edible items, writing utensil). After the teacher and consumer or simulated consumer were seated, the experimenter instructed the teacher to begin implementing the assessment and to say when he or she was finished. The experimenter arranged materials such that the teacher always sat on one side of the table and the simulated or actual consumer sat across from the teacher. Eight edible stimuli were used throughout training with simulated consumers and most generalization probes. Toys were used during generalization probes with four actual consumers who did not typically receive food as part of their ongoing reinforcement programs.

Preference Assessments for Simulated and Actual Consumers

Teachers conducted, summarized data, and interpreted results from two types of preference assessments: the PS method and the MSWO method. We used the same eight edible items

(pretzel, cracker, candy corn, cookie, corn chip, marshmallow, raisin, chocolate candy) for both assessments. As in Roscoe et al. (2006), we made some additions to the published procedures (i.e., we specified how to respond when a consumer or simulated consumer made no response or tried to select multiple items simultaneously). Simulated consumers informed participants that they would not answer any questions or provide any feedback related to preference assessment procedures until after the completion of the study.

Session length varied for each method of preference assessment. For the PS assessment, each training session involved 10 to 13 consecutive trials. Most training sessions consisted of 10 trials. On occasion, teachers repeated trials after a simulated consumer made an atypical response, and data were collected on these trials as well. For the MSWO assessment, each session consisted of eight trials (or fewer, if the simulated or actual consumer stopped responding for 30 s before all eight trials were completed). After completion of the written instruction phase, teachers calculated selection percentages for data they had collected. After the generalization probe, teachers calculated selection percentages for a complete set of hypothetical data.

PS assessments. The PS preference assessment was based on procedures described by Fisher et al. (1992). For each simulated consumer, we assessed a fixed and predesignated pool of eight edible items. Written instructions directed teachers to engage in several responses for each assessment trial. First, two stimuli were to be placed approximately 0.3 m in front of the simulated consumer and 0.5 m apart. The teacher then was to prompt verbally the simulated consumer to select one item. If the simulated consumer selected an item within 5 s of the verbal prompt, the teacher was to remove the item that was not selected and record the name of the selected stimulus on his or her data sheet. If no selection response occurred within 5 s, the teacher was to remove both stimuli, to record “no response” for that trial, and to prepare stimuli for the next trial.

These procedures differed from Fisher et al.’s recommendation to present items individually and then re-present them as a pair on the subsequent trial. This modification served to save time, given that scripted no-response trials for simulated consumers were scheduled once or twice every 10 trials. Finally, if the simulated consumer attempted to select both stimuli at the same time, the teacher was instructed to block the response and repeat the trial.

MSWO assessments. The MSWO preference assessment was based on procedures described by DeLeon and Iwata (1996). We assessed eight edible items for each simulated consumer. Each session was comprised of a maximum of eight trials. At the beginning of each session, the teacher sat at a table across from the simulated consumer. Written instructions first directed the teacher to place the eight items in a straight line approximately 5 cm apart and 0.5 m in front of the consumer. Then, the teacher was to instruct the simulated consumer to “choose one.” The teacher was to block attempts to approach more than one item on a trial. After the simulated consumer approached one item, the teacher was not to replace the selected item in the array. The teacher was instructed to record the item that was selected, reorder the remaining stimuli by rotating the stimulus on the far left to the far right position, and reposition the array so that it was centered in front of the consumer. The assessment continued in this manner until the last item was approached or the simulated consumer did not approach any of the remaining items within 30 s.

Assessment scripts for simulated consumers. For each type of preference assessment, the simulated consumers used four different assessment scripts. Across scripts, we scheduled different consumer responses to occur on different trials. During each session of the PS assessment, we programmed atypical responses on 50% of trials. On the remaining trials, the simulated consumer simply selected one stimulus. Each eight-trial MSWO assessment included four

trials with programmed atypical responses. On the other four trials, the simulated consumer simply selected one stimulus. After the simulated consumer selected an edible item, he or she consumed it or pretended to place it in his or her mouth and later discarded the item. Atypical responses on the PS and MSWO assessments included attempts to select two items simultaneously, attempts to select two items in quick succession, not responding within the allotted time, and attempts to select an item that was not presented on the trial.

Design and Procedure

For each teacher, we used a multiple baseline design across assessment types to evaluate the training materials. We varied the order in which assessments were trained across teachers to detect potential sequence effects. The mastery criterion for each preference assessment was for teachers to conduct assessment trials and collect trial-by-trial data with 90% or greater accuracy for two consecutive sessions of 8 to 10 trials. We defined the mastery criterion for scoring and interpretation of preference assessment outcomes as 100% accurate calculation of selection percentages and an accurate verbal report of the most preferred stimulus for one complete set of preference assessment data.

To evaluate the contribution of specific treatment components to training outcomes, teachers experienced one of two training sequences. We conducted training sessions with one teacher at a time. Six teachers progressed from written instructions alone (baseline), to enhanced written instructions, to generalization probes. Five teachers progressed from written instructions alone, to written instructions plus data sheet, to enhanced written instructions, to generalization probes.

Written instructions alone (baseline). Prior to the first session, we provided teachers with written instructions for conducting PS or MSWO preference assessments (specific text available from the first author). The instructions were drawn from the methods sections of

previously published literature (DeLeon & Iwata, 1996; Fisher et al., 1992), and teachers had a maximum of 30 min to read the instructions immediately before the first session. Teachers informed the experimenter when they were ready to begin. Previous studies (e.g., Roscoe et al., 2006) found that providing participants with instructions prior to, but not during, sessions did not result in accurate performance. Thus, we allowed teachers in this study to take the written instructions with them and refer to them while conducting assessments during baseline and all subsequent phases of the investigation.

Enhanced written instructions. In this phase, we provided teachers with a detailed data sheet and step-by-step instructions written without technical jargon and supplemented with diagrams. Teachers had a maximum of 30 min to read the instructions immediately prior to the first session. We also permitted teachers to refer to the enhanced instructions while they conducted preference assessment sessions.

Written instructions plus data sheet. We conducted preference assessments with simulated consumers using the same written instructions from baseline. However, the data sheet from the enhanced written instructions phase was also available to 5 of 11 teachers during this phase. This allowed evaluation of the effects of the data sheet alone on performance.

Generalization probes. We conducted generalization probes with actual consumers (i.e., children with disabilities who received services at the teacher's place of employment) after each teacher achieved the mastery criterion for a specific method of preference assessment. For the eight recently hired teachers, probes occurred within 1 week of achieving the mastery criterion. Due to limited availability of day-care and integrated preschool teachers (DS, GL, and NI), generalization probes occurred from 1 week to 1 month after achieving mastery with simulated consumers. All actual consumers had previous exposure to stimuli assessed during the generalization probes. Each generalization

probe for the PS assessment consisted of 10 trials. For the MSWO assessment, each generalization probe was eight trials long.

Response Measurement and Interobserver Agreement

Dependent variables. Teacher target responses varied by assessment type. On each trial, the observer scored five specific target responses as correct or incorrect. Accuracy for implementing each specific target response was calculated by dividing the number of correct target responses by the total number of opportunities to engage in the target response each session and multiplying by 100%. Total accuracy for implementing each assessment procedure was calculated by dividing the number of trials with correct responses (i.e., all specific target responses on the trials were scored as correct) by the total number of trials implemented and multiplying by 100%.

On each trial of the assessment, observers recorded whether or not teachers accurately engaged in the following target responses: (a) Stimulus presentation (PS assessment): Observers recorded correct stimulus presentation on a trial if the teacher placed two different stimuli in front of the consumer. Observers scored an incorrect response if the teacher placed a smaller or larger number of stimuli on the tabletop. Also, observers scored a trial as incorrect if the teacher presented the same stimulus on three consecutive trials. (In typical PS assessments, across all trials, each item is paired with every other item the same number of times, and the positions of the stimuli are counterbalanced such that each item falls on the participant's right and left positions the same number of times. In this study, however, most sessions consisted of only 10 trials. Thus, these features were not included in the definition of correct stimulus presentation.) (b) Stimulus presentation (MSWO assessment): Observers scored correct stimulus presentation on the first trial of each session if the teacher placed eight different edible items in front of the consumer. Observers scored an incorrect response if the teacher did

not place eight different stimuli in front of the consumer. On subsequent trials, observers scored a correct stimulus presentation if the teacher did not replace the item consumed on the previous trial. Observers scored a trial as incorrect if the teacher replaced the selected item in the stimulus array or removed a nonselected item from the array. (c) Stimulus position: Observers scored correct stimulus position when the teacher placed stimuli approximately 0.3 m in front of the consumer and 0.5 m apart. (d) Postselection response: A correct postselection response was defined as the teacher immediately (i.e., before recording data) removing the item that was not selected by the consumer (PS) or the teacher rotating the positions of stimuli and respacing them so that the array was centered in front of the consumer (MSWO). (e) Response blocking: Observers scored a correct response on relevant trials if the teacher attempted to block approach responses to more than one stimulus. Observers scored an incorrect response if the teacher's hands did not move towards the hands of the consumer when he or she attempted to select more than one stimulus on a trial. (f) Trial termination: Observers scored a correct response if the teacher terminated a trial (i.e., removed stimuli from the tabletop) when no consumer response occurred within 5 s of stimulus presentation (PS) or removed all stimuli from the tabletop and moved to the next trial after no response occurred within 30 s (MSWO).

For both assessments, observers scored a trial as correct if the teacher accurately implemented the above-mentioned target responses. If the teacher incorrectly implemented any of these steps, the entire trial was scored as incorrect. Response blocking and trial termination could be scored only on trials in which the simulated consumer responded atypically; thus, for these trials to be scored as correct, the teacher had to implement an additional target response accurately. On trials in which the simulated consumer responded typically, the teacher had to respond to only three of five possible responses.

We also collected data on how accurately the teacher recorded data, summarized assessment results, and interpreted assessment outcomes. After completion of each trial, secondary observers recorded whether or not the teacher correctly wrote down which stimulus was selected or that no response occurred. After the final session of the written instructions alone condition, accuracy of summarizing assessment results (i.e., calculating selection percentages) was assessed by having the teacher calculate the percentage of approach responses for each stimulus from his or her assessment data. After the final session of the enhanced written instructions condition, the teacher calculated selection percentages (PS assessment) or total points (MSWO assessment) from a complete hypothetical set of data. The experimenter assessed accuracy of scoring by comparing the summary score associated with each stimulus to the corresponding score calculated by data recorders. For the PS assessment, the teacher calculated the percentage of trials on which each stimulus was approached. For the MSWO assessment, the teacher summed the number of points associated with each stimulus (Ciccone, Graff, & Ahearn, 2005). The point-based scoring method was selected because it produces more differentiated results than traditional selection percentages and may be easier for teachers to interpret. The point-based method was also standard protocol in the school in which these teachers worked.

After the teacher calculated the percentage of approach responses (PS) or the number of points obtained (MSWO), he or she interpreted those results. Specifically, he or she was asked to name which item he or she would use if he or she was trying to teach a student a new skill.

Social validity assessment. Three to 6 months after completion of the study, the experimenter provided teachers with copies of the written instructions, the enhanced written instructions, and a brief questionnaire. The questionnaire was designed to assess teacher reports of the

importance and acceptability of conducting preference assessments and which training method the teachers felt was more effective (i.e., written instructions alone or enhanced instructions). Finally, teachers were asked which method they would use if they were responsible for training new staff to conduct preference assessments.

Interobserver agreement and integrity of the independent variable. Two observers independently scored data (live and by viewing videotapes) for 33% of sessions in all experimental conditions for both assessment types. On each trial, an agreement was scored when two observers recorded an error or correct implementation for a specific target response. A disagreement was scored when one observer recorded a specific target response as correct and the other observer scored it as incorrect. Interobserver agreement for each target response was calculated by dividing the number of agreements by the number of agreements plus disagreements per session and multiplying by 100%.

We also calculated interobserver agreement for total accuracy of teacher responses. An agreement was scored when two observers recorded an entire trial as correctly implemented (i.e., all five specific target responses scored as correct). A disagreement was scored when one observer recorded a correct trial and the other observer scored an incorrect trial.

For the PS assessment, mean interobserver agreement across independent target responses was 94% (range, 88% to 98%). For the MSWO assessment, mean agreement across individual target responses was 96% (range, 92% to 99%). Across teachers, mean agreement for total accuracy was 95% (range, 90% to 99%) for the PS assessment and 96% (range, 91% to 99%) for the MSWO assessment.

We collected data on integrity of the independent variable on a trial-by-trial basis on the simulated consumers' implementation of the assessment scripts for 33% of sessions across teachers and phases. A trial was scored as correct

if the simulated consumer performed the actions prescribed on the script; a trial was scored as incorrect if the simulated consumer did not perform the action prescribed on the script. Mean integrity of the independent variable was 97% (range, 90% to 100%) for the PS assessment and 98% (range, 88% to 100%) for the MSWO assessment.

RESULTS

Figure 1 depicts the percentage of trials with accurate implementation of both assessment methods for the six teachers who experienced the written instructions alone (baseline) condition followed by the enhanced instructions condition and generalization probes. None of the teachers met the mastery criterion for either assessment using written instructions alone. When teachers had access to the methods section from the published study and a blank sheet of paper, mean total accuracy across participants for the PS assessment was 34% (range, 9% to 56%). Mean total accuracy across participants for the MSWO assessment was 46% (range, 13% to 61%). When teachers were provided with enhanced instructions, total accuracy increased immediately for all participants. Five of six participants (CE, CP, DS, GL, and NI) met the mastery criterion (at least 90% accuracy across two consecutive sessions) for both assessments in the minimum number of two sessions.

Five teachers experienced the experimental conditions in the following order: written instructions alone, written instructions plus data sheet, enhanced written instructions, and generalization probes. Accuracy data for these individuals are depicted in Figure 2. When provided with written instructions alone, mean total accuracy across teachers was 38% (range, 20% to 64%) for the PS assessment and 38% (range, 8% to 58%) for the MSWO assessment. When written instructions were supplemented with a data sheet, accuracy increased for four of five participants (MA, MS, BB, and TT) on the PS assessment and for three of five participants

(MS, BB, and TT) on the MSWO assessment. Mean accuracy of performance increased to 73% for the PS assessment (range, 65% to 89%) and 59% for the MSWO assessment (range, 12% to 79%) in the written instructions plus data sheet condition. Mean total accuracy increased further (98% for the PS assessment and 99% for the MSWO assessment) when enhanced written instructions were presented.

Mean percentage accuracy of specific target responses is depicted in Table 1. When provided with written instructions alone, teachers were more accurate in presenting stimuli on the MSWO assessment than on the PS assessment, but were more accurate on correct placement of stimuli for the PS assessment. Accuracy of postselection responses was similar across assessments when teachers were presented with written instructions alone. For the five teachers who experienced the traditional instructions plus data sheet condition, accuracy of stimulus presentation, stimulus placement, and postselection response increased to nearly 100% on both assessments. The addition of a data sheet to the written instructions (baseline) had no effect on accuracy of response blocking and trial termination. This finding may reflect the fact that the data sheet provided no additional information related to blocking or trial-termination responses. However, introduction of the enhanced instructions, which did include specific guidelines for blocking and trial termination, produced highly accurate responding in these areas for both assessment types. Generalization probes indicated that accuracy on all components of assessment implementation remained high when actual consumers participated.

To evaluate whether or not teachers could interpret preference assessment results accurately, each participant was asked to name the item that he or she would use if he or she were trying to teach a student a new skill. After the final session of the written instructions alone condition, 4 of 11 teachers named the item that was selected most frequently on the PS assessment.

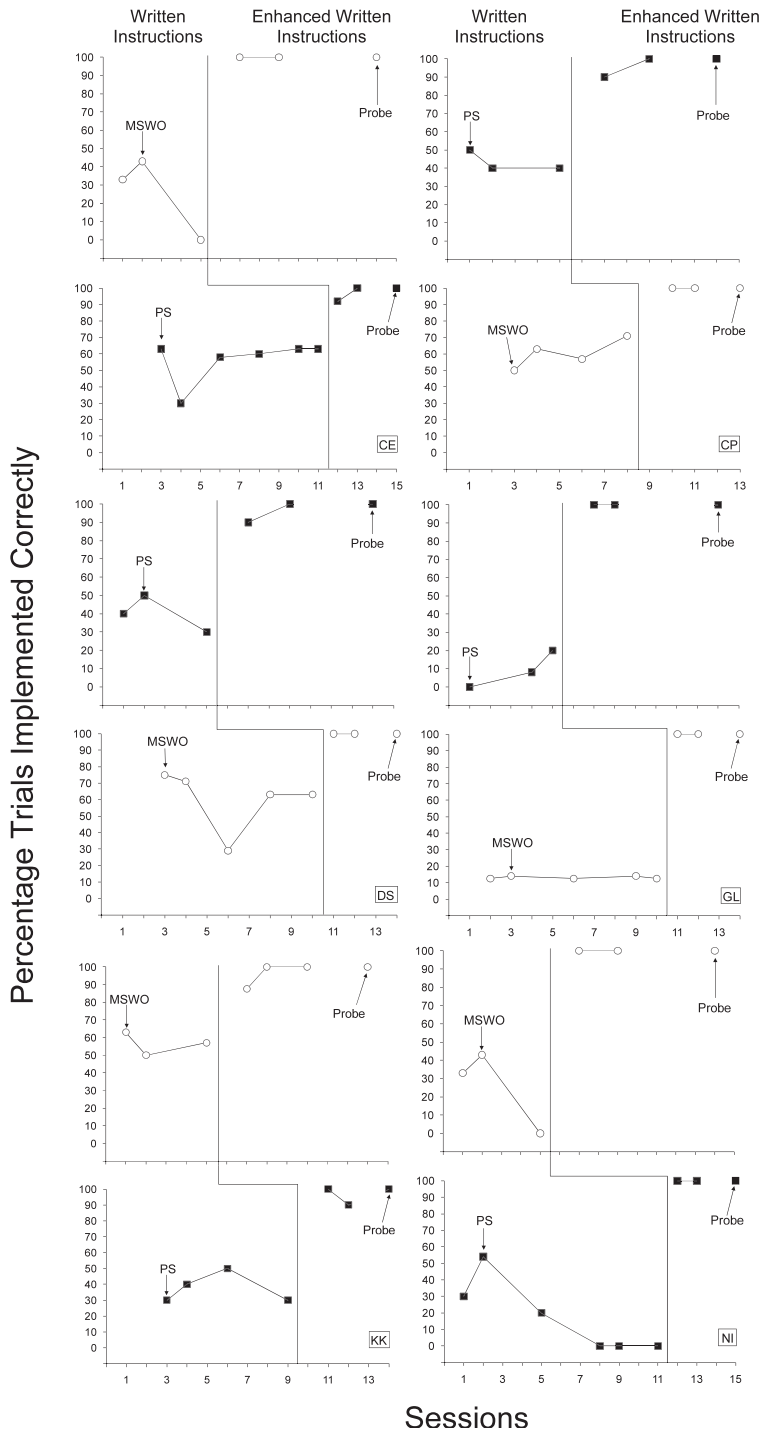


Figure 1. Percentage of trials implemented correctly across conditions for teachers who experienced written instructions alone (baseline), enhanced written instructions, and generalization probes.

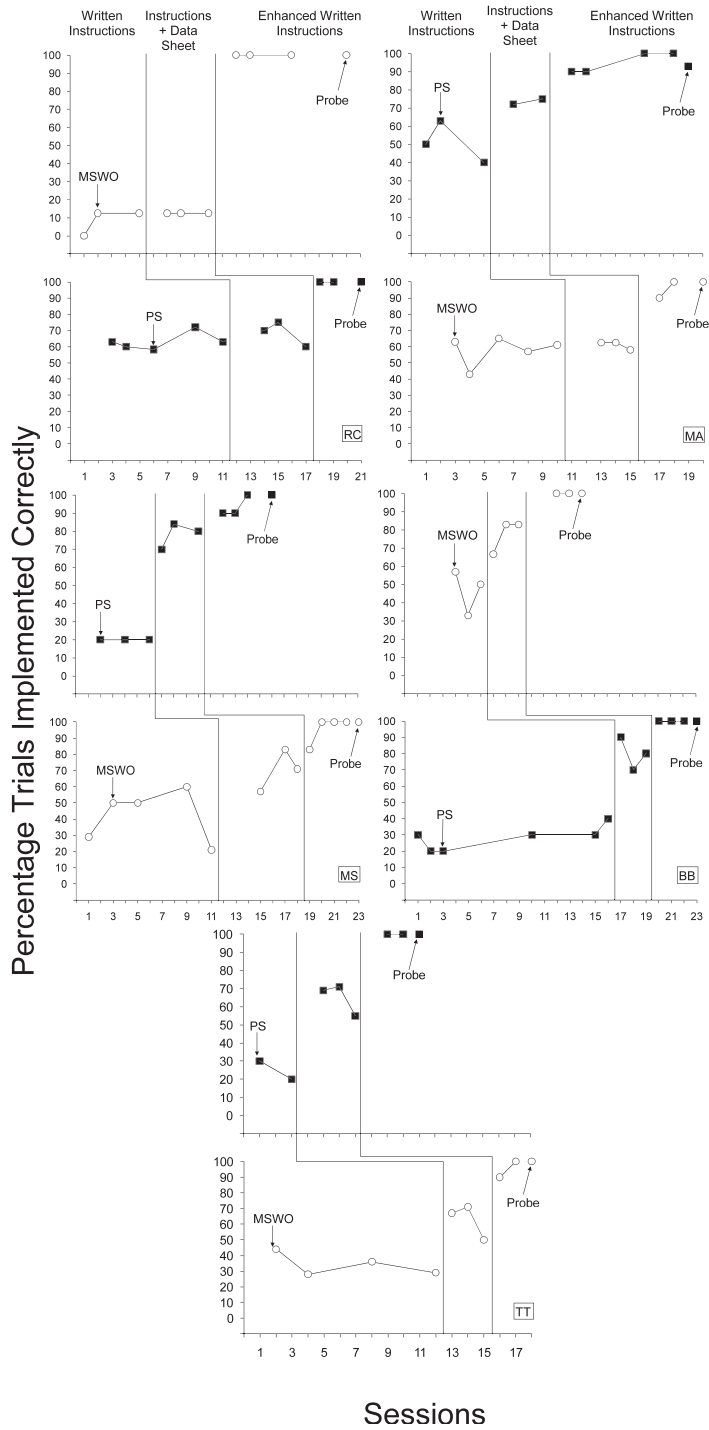


Figure 2. Percentage of trials implemented correctly across conditions for teachers who experienced written instructions alone (baseline), written instructions plus data sheet, enhanced written instructions, and generalization probes.

Table 1
Mean Accuracy (%) per Target Response for PS and MSWO Assessments

Assessment type	Instruction type	Number of cases	Stimulus presentation	Stimulus placement	Postselection response	Response blocking	Trial termination
PS	Written instructions	11	68	89	80	48	62
	Written + data sheet	5	100	94	99	38	60
	Enhanced instructions	11	100	99	100	98	98
MSWO	Written instructions	11	85	71	78	22	62
	Written + data sheet	5	100	100	97	25	67
	Enhanced instructions	11	100	100	98	97	99

None of the teachers correctly named the item associated with the highest point total on the MSWO. After the final session of the enhanced written instructions condition, all teachers indicated that they would use the item selected most frequently on the PS assessment and the item associated with the highest point total on the MSWO.

The social validity assessment indicated that participants preferred the enhanced instructions and found them easier to use than the traditional written instructions (see Table 2). When asked whether they would use traditional written instructions or enhanced instructions to

train new staff, all participants indicated that they would use the enhanced instructions.

DISCUSSION

Results of the current study demonstrated that individuals without previous experience in conducting stimulus preference assessments did not accurately implement PS and MSWO assessment procedures using written instructions alone (i.e., procedural description from the method section of published assessment studies). This finding is consistent with the results of Lavie and Sturmey (2002), Roscoe

Table 2
Results of the Social Validity Survey

Questions	Mean rating (1 = <i>strongly disagree</i> , 7 = <i>strongly agree</i>)	
It is important that individuals with special needs are motivated to learn	6.9 (range, 6 to 7)	
For some students with special needs, it is difficult to find things that will motivate them	5.3 (range, 4 to 7)	
Conducting systematic preference assessments is an acceptable way to identify things that will motivate children with special needs	6.4 (range, 5 to 7)	
	Enhanced instructions (step-by-step with pictures)	Written instructions (method from published research article)
Instructions were easy to follow	6.8 (range, 6 to 7)	3.5 (range, 1 to 5)
Instructions made the procedure easy to understand	6.7 (range, 6 to 7)	3.5 (range, 1 to 5)
I was confident I was implementing procedures correctly	6.2 (range, 5 to 7)	3.2 (range, 1 to 5)
I enjoyed conducting the assessment	6.3 (range, 5 to 7)	4.2 (range, 2 to 6)
If I was going to provide training materials to someone who had no previous exposure to this topic, I would recommend using these instructions	6.5 (range, 5 to 7)	2.5 (range, 1 to 4)
If I wanted to refresh my memory, I would use these instructions	6.6 (range, 5 to 7)	2.7 (range, 1 to 4)

et al. (2006), and Roscoe and Fisher (2008). One objective of the current study was to increase the efficacy of written instructions for those circumstances in which expert training and feedback are not readily available. Results for 11 of 11 teachers suggest that enhanced written instructions that included pictures and diagrams, step-by-step examples, and minimal technical jargon were sufficient to prepare inexperienced staff to implement assessment procedures accurately.

Roscoe et al. (2006) showed that trainer feedback was a crucial component of effective staff training on preference assessments because feedback provided trainees with clear discriminative information regarding errors and correct performance. In the present study, enhanced written instructions appeared to provide enough information for teachers to conduct PS and MSWO assessments accurately without individualized trainer feedback.

In addition to the fact that written instructions were enhanced with diagrams and non-technical language, participants had access to a copy of instructions while they conducted trials. Continuous access to written (baseline) and enhanced instructional materials was not available in previous studies. Anecdotally, during the current investigation, few participants referred to the enhanced instructions after a session began. It is possible that ongoing access to enhanced written instructions would be necessary to maintain procedural integrity over time, especially if teachers conduct preference assessments infrequently.

Participants in this study worked in a behavior-analytic school in which teachers were required to conduct preference assessments on a regular basis. Therefore, it is possible that these participants were more motivated to master assessment techniques than individuals who work in other settings. However, the preschool teacher participants (DS, GL, and NI) were not required to conduct preference assessments as part of their jobs, yet they performed as well or

better than their peers during the training. Thus, motivational variables may play a secondary role to adequate discriminative information in effective staff training. This possibility is supported by Roscoe et al. (2006), who found that the addition of money contingent on accurate responding was insufficient to train four staff members to implement PS and MSWO assessments.

Results from this investigation raise the question of whether enhanced written instructions could be used to train inexperienced staff on behavior-analytic technologies other than preference assessments. Although some procedures cannot be used safely or ethically by individuals who lack extensive training in behavior analysis (e.g., functional analysis of severe self-injury), low-risk and generally applicable procedures potentially could be disseminated using enhanced written instructions. Future research should examine whether practices such as selecting and implementing systems of data collection or effective reinforcer delivery, prompting, and prompt fading could be established and maintained using enhanced written instructions.

The current study has several limitations that should be noted. Each multiple baseline design included only two baselines, and the enhanced instructions differed slightly across assessment types (i.e., the instructions for the MSWO contained more diagrams, and the procedures for scoring the assessments differed). However, the introduction of the independent variable was associated with immediate increases in accurate performance of both assessment procedures, and this result was replicated across 10 additional participants. This increased our confidence that the intervention was responsible for increased accuracy.

Another limitation was that participants were exposed to the methods section from the relevant journal articles prior to experiencing the enhanced written instructions. Thus, it is possible that sequence effects may have influenced our results. Future research could evaluate whether

enhanced written instructions are as effective when participants do not have exposure to the relevant research articles.

Although the introduction of enhanced written instructions was associated with increased accuracy, the critical components of the enhanced written instructions are unknown. The introduction of the data sheet alone was associated with increased accuracy in 7 of 10 instances, suggesting that it may be an important component. Future research should more thoroughly evaluate this possibility. In addition, future research should isolate and systematically evaluate the role of nontechnical language, diagrams, and examples in training effects. If critical variables are identified, it is possible that they could be integrated with the data sheet in a more concise and accessible training document.

All teachers rated the enhanced written instructions as easier to use than traditional written instructions on the social validity survey. It may have been more informative, however, to ask teachers which specific features of the instructions they found most helpful. Given that some teachers' performance increased when the data sheet was added to the written instructions, it may have been interesting to see how teachers rated the helpfulness and acceptability of that condition compared to the enhanced written instructions. Future research might assess the social validity of each treatment component separately.

REFERENCES

- Behavior Analyst Certification Board. (2011). Retrieved from <http://www.bacb.com>
- Cannella, H. I., O'Reilly, M. F., & Lancioni, G. E. (2005). Choice and preference assessment research with people with severe to profound intellectual disabilities: A review of the literature. *Research in Developmental Disabilities, 26*, 1–15.
- Ciccione, F. J., Graff, R. B., & Ahearn, B. H. (2005). An alternate scoring method for the multiple stimulus without replacement preference assessment. *Behavioral Interventions, 20*, 121–127.
- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519–533.
- Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*, 491–498.
- Fisher, W. W., Piazza, C. C., Bowman, L. G., Kurtz, P. F., Sherer, M. R., & Lachman, S. R. (1994). A preliminary evaluation of empirically derived consequences for the treatment of pica. *Journal of Applied Behavior Analysis, 27*, 447–457.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197–209. (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2*, 3–20, 1982)
- Karsten, A. M., & Carr, J. E. (2009). The effects of differential reinforcement of unprompted responding on the skill acquisition of children with autism. *Journal of Applied Behavior Analysis, 42*, 327–334.
- Lavie, T., & Sturmey, P. (2002). Training staff to conduct a paired-stimulus preference assessment. *Journal of Applied Behavior Analysis, 35*, 209–211.
- Lerman, D. C. (2003). From the laboratory to community application: Translational research in behavior analysis. *Journal of Applied Behavior Analysis, 36*, 415–419.
- Lerman, D. C., Swiezy, N., Perkins-Parks, S., & Roane, H. S. (2000). Skill acquisition in parents of children with developmental disabilities: Interaction between skill type and instructional format. *Research in Developmental Disabilities, 21*, 183–196.
- Moore, J. W., & Fisher, W. W. (2007). The effects of videotape modeling on staff acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis, 40*, 197–202.
- National Center for Education Statistics. (2006). Retrieved from <http://nces.ed.gov/fastfacts/display.asp?id=59>
- Roscoe, E. M., & Fisher, W. W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavior Analysis, 41*, 249–254.
- Roscoe, E. M., Fisher, W. W., Glover, A. C., & Volkert, V. M. (2006). Evaluating the relative effects of feedback and contingent money for staff training of stimulus preference assessments. *Journal of Applied Behavior Analysis, 39*, 63–77.
- Wallace, M. D., Iwata, B. A., & Hanley, G. P. (2006). Establishment of mands following tact training as a function of reinforcer strength. *Journal of Applied Behavior Analysis, 39*, 17–24.

Received November 21, 2010
 Final acceptance August 3, 2011
 Action Editor, David Wilder