Evaluation of a Training Manual to Teach Multiple-Stimulus Preference Assessment

by

Duong Ramon

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Winnipeg

Abstract

A self-instructional training manual for conducting a multiple-stimulus without replacement (MSWO) preference assessment procedure was evaluated. The manual describes the MSWO procedure and how to summarize and interpret its results. Review questions that highlighted the key target behaviours were included at the end of each section of the manual, followed by an answer key. The manual was compared to a method description, adapted from the method sections of research articles published by DeLeon and Iwata (1996) and Roscoe, Fisher, Glover, and Volkert (2006). Eighteen undergraduate university students were assigned to two groups using a matched-pairs random assignment for the first 10 participants and random assignment for the next eight participants. Group 1 received the manual training first, followed by a crossover to the method description training if the pre-determined mastery criterion (85% correct or higher) was not met during simulated assessments. Group 2 received the training procedures in reverse order. The self-instructional manual was statistically significantly more effective than the method description in improving performance accuracy for conducting the MSWO procedure with an actor (a graduate student) simulating a person with an intellectual disability. Four out of nine participants in Group 1 met the mastery criterion after studying the manual only and one participant achieved mastery after the crossover. In contrast, none of the nine participants in Group 2 met the mastery criterion after studying the method description only and seven participants achieved mastery after the crossover. The remaining six participants all met the mastery criterion after observing a live demonstration of the procedure. Within each group, interventions were implemented in a modified multiple-baseline design across participants and the results

showed that improvements occurred only after an intervention had been introduced. All participants who achieved mastery showed strong retention and generalization performances with novel actors and clients. On the social validity questionnaire, mean participant ratings showed that the manual was easier to follow and understand, and provided the necessary information for the assessment compared with the method description. The self-instructional manual has considerable potential as a low cost and effective tool to teach individuals to conduct the MSWO procedure.

Acknowledgements

I thank my academic advisor Dr. C.T. Yu for his guidance and continued support throughout my study; Drs. Garry Martin, Angela Cornick, and Dennis Hrycaiko, for their valuable input and recommendations; and to my research assistants for their help with data collection. I express my sincere gratitude to the participants and clients without whom this research would not be possible. This research was supported in part by fellowships from the Canadian Institutes of Health Research and the Social Sciences and Humanities Research Council of Canada.

Table of Contents

List of Tables

List of Figures

Evaluation of a Training Manual to Teach Multiple-Stimulus Preference Assessment

Introduction

Many studies have demonstrated that the opportunity for individuals with intellectual and developmental disabilities to exercise choice has a positive impact on their quality of life (Hughes, Pitkin, & Lorden, 1998). Often, these individuals do not have opportunities throughout their day to make choices even though there is ample evidence to demonstrate that they can (Cannella, O'Reilly, & Lancioni, 2005; Hagopian, Long, & Rush, 2004; Steege, Wacker, Berg, Cigrand, & Cooper, 1989; Tullis et al., 2011; Wacker, Berg, Wiggins, Muldoon, & Cavanaugh, 1985; Wacker, Wiggins, Fowler, & Berg, 1998). Furthermore, some studies have demonstrated that the opportunity to engage in choice-making behaviour can result in reduction in frequency of challenging behaviour for this population (Cooper, Wacker, Thursby, & Plagmann, 1992; Harding, Wacker, Berg, Cooper, & Asmus, 1999; Peck, et al., 1996).

Preference assessments are tools that caregivers can use to provide opportunities for these individuals with no speech to make choices to identified preferred activities and food items (Fisher, Piazza, Bowman, & Amari, 1996; Lohrmann-O'Rourke & Browder, 1998). Considerable research has shown that preference assessments can be used to identify reinforcers for persons with developmental disabilities, which is an important component of any skills acquisition program (Graff, Gibson, & Galiatsatos, 2006; Green, Reid, Canipe, & Gardner, 1991; Logan & Gast, 2001). The importance of preference assessment for supporting this population is underscored by recent research efforts on teaching individuals to carry out this procedure. Several studies have successfully trained

students and staff to conduct stimulus preference assessments through one-to-one or small-group instruction; however, such training is resource intensive. Developing a lower-cost, yet effective alternative that does not require face-to-face instruction would be highly desirable. The purpose of this study was to evaluate a self-instructional training manual to teach undergraduate university students to conduct the multiple-stimulus without replacement (MSWO) preference assessment procedure. In the ensuing sections, I review the concept of preference, research on direct preference assessment procedures, research on teaching individuals to conduct preference assessments, describe the method and results of this study, and discuss its findings.

Stimulus Preference Assessment

Preference is defined as a pattern of selections among available stimuli (Martin, Yu, Martin, & Fazzio, 2006). Stimulus preference assessments with individuals with intellectual and developmental disabilities can be direct or indirect. Indirect assessments rely on the opinions of informants such as staff or caregivers who are familiar with the individual. Indirect preference assessments are less time consuming than direct assessments; however, they are less reliable in identifying reinforcers (Fisher et al., 1996; Green et al., 1991; Green et al., 1988; Parsons & Reid, 1990; Windsor, Piché, & Locke, 1994).

Direct preference assessments may be conducted using free-operant or trial-based approaches. In free-operant assessments, a participant is given access to two or more stimuli concurrently for a period of time (e.g., Roane, Vollmer, Ringdahl, & Marcus, 1998). The frequency of interactions or duration of engagement with each stimulus is measured, and the stimulus with the highest rate of interaction or the longest

engagement duration is the most preferred. In trial-based assessments, one or more stimuli are presented on each trial for a brief period (e.g., 5-10 s) and a participant's approach or selection response is recorded. Across trials, the stimulus that has been approached most frequently is considered the most preferred. Compared to the free-operant procedure, trial-based approaches provide greater differentiation among the items assessed, not only identifying highly preferred items, but also less-preferred items (Kodak, Fisher, Kelly, & Kisamore, 2009; Ortiz & Carr, 2000; Roane et al., 1998), whereas free-operant procedures do not produce discrete rankings for each item assessed. Furthermore, the free-operant procedure may produce false negatives if the individual does not interact with all stimuli presented. In other words, items that are not selected during the free-operant procedure may still be highly preferred. As a result, for individuals with intellectual and developmental disabilities, trial-based assessments have received more research attention than the free-operant approach. Therefore, research on trial-based, direct preference assessments are reviewed in more detail below.

Single-, Paired-, and Multiple-Stimulus Presentations

In trial-based preference assessments, the number of stimuli presented on each trial has received considerable attention. In single-stimulus (SS) presentations, one stimulus is presented on each trial and stimuli are presented in a random order (Pace, Ivancic, Edwards, Iwata, & Page, 1985). In paired-stimulus (PS) presentations, two stimuli are presented concurrently on each trial and each stimulus is paired with every other stimulus (Fisher et al., 1992). In multiple-stimulus (MS) presentations, more than two stimuli are presented concurrently (DeLeon & Iwata, 1996; Windsor et al., 1994).

Pace et al. (1985) evaluated the SS procedure with six individuals with profound developmental disabilities. In Experiment 1, the participants' preferences among 16 stimuli were determined by their approach response to each stimulus (i.e., either pointing to or reaching for a stimulus). On each trial, if the participant approached the stimulus within 5 s, he or she was given immediate access to the stimulus for 5 s. If the participant did not approach the stimulus within 5 s, the experimenter prompted him/her to sample the stimulus, and then repeated the trial. If the participant still did not approach the stimulus, the next stimulus was presented. All participants approached at least one stimulus on 80% or more of the trials. In Experiment 2, the authors compared the reinforcing effects of the preferred and less-preferred stimuli (approached on less than 50% of the trials) on instruction-following behaviours in a reversal design. The preferred stimuli resulted in higher response rates relative to baseline than the less-preferred stimuli for all participants.

Several studies have found that the PS procedure was more effective than the SS procedure in identifying preferences (DeLeon & Iwata, 1996; Fisher et al., 1992; Fisher et al., 1996; Pace et al., 1985; Piazza, Fisher, Hagopian, Bowman, & Toole, 1996). For example, Fisher et al. (1992) compared the two procedures with four persons with severe and profound developmental disabilities. The PS procedure yielded a more differentiated preference hierarchy among 16 stimuli than did the SS procedure. The latter procedure was less sensitive because participants could and tended to approach the item on every trial, whereas they could select only one of two items on each trial in the PS procedure.

The preferred stimuli identified by the PS procedure have also been shown to be potent reinforcers (DeLeon, Frank, Gregory, & Allman, 2009; Lee, Yu, Martin, &

Martin, 2010; Piazza et al., 1996). For example, Lee et al. found that more preferred items tended to be more powerful reinforcers. Piazza et al. tested the reinforcing effects of top-ranked (most preferred), middle-ranked, and bottom-ranked (least preferred) stimuli identified by a PS procedure with four participants with severe and profound developmental disabilities. Results indicated that the top-ranked stimuli functioned as reinforcers for all participants, the middle-ranked stimuli functioned as reinforcers for two participants, and the bottom-ranked stimuli did not function as reinforcers for any participant.

Windsor et al. (1994) were the first to compare the PS procedure to a six-item MS procedure to identify food preferences for eight participants with severe to profound developmental disabilities. During the MS procedure, all six stimuli were presented on each trial. The results showed that both the PS and MS procedures identified the same stimuli as most preferred for all participants. The PS procedure yielded more consistent preferences across assessments, but the MS procedure required less time to administer than the PS procedure (means of 7 min vs. 16 min, respectively). A potential limitation of the MS procedure is that because all stimuli are presented on each trial, participants tend to select the same and presumably the most-preferred stimuli on every trial thereby yielding little information about the relative preference among the stimuli.

DeLeon and Iwata (1996) compared the PS, MS, and multiple-stimulus without replacement (MSWO) procedures. The MSWO procedure differed from the MS procedure in that the stimulus selected on a trial was not presented on subsequent trials of that session. Trial presentations would continue in the MSWO procedure until either the last two items had been presented or no selection had been made. Seven items were

used and the maximum number of trials was 30 (i.e., until the last two items were presented). Five consecutive sessions were conducted for each procedure with each participant and the order of the three procedures was varied across participants. Although all three procedures identified the same stimuli as the most preferred for four of the seven participants, the PS and the MSWO procedures produced more consistent stimulus preference rankings across sessions and identified the same three stimuli as the most preferred. Moreover, the preferred stimuli identified by the PS and the MSWO procedures, and which were not selected during the MS procedure, functioned as reinforcers during subsequent tests for three of the four participants. Lastly, the PS procedure required an average of 53.3 min, whereas the MS and MSWO procedures required an average of 16.5 min and 21.8 min, respectively.

Higbee, Carr, and Harrison (2000) evaluated the effectiveness of an abbreviated MSWO procedure and tested the reinforcing effects of the top-ranked stimuli with nine adults with severe or profound intellectual disabilities. The abbreviated MSWO procedure involved conducting the assessment for three instead of five sessions as described in the DeLeon and Iwata (1996) study. During reinforcer tests conducted after the preference assessments, the most- or second-most-preferred stimulus increased responding above baseline levels for seven of the nine participants. The three-session MSWO procedure required 15 min or less to complete.

Carr, Nicolson, and Higbee (2000) evaluated the effectiveness of a three-session MSWO assessment in an individual's everyday setting and whether an individual's preference varied over a one-month period. Participants were three children with autistic disorder. Each session began with eight items on the first trial and continued on

successive trials until the last two items had been presented, and the session was repeated three times. A reinforcer assessment was conducted following the preference assessment for each child using the most-preferred item (stimulus that was ranked first), a moderately preferred stimulus (ranked fourth or fifth), and the least-preferred stimulus (ranked eighth) to increase a low-frequency response selected from the child's curriculum. The least-preferred stimulus did not increase responding relative to baseline for any participant. The middle-ranked stimulus produced moderate increase for one child. The most-preferred stimulus produced the largest increase in responding relative to baseline and to the other stimuli for all children. In addition, eight additional MSWO assessments were conducted over a four-week period for each child. Preference rankings over time were stable for two children. The authors also found that the results obtained from the first of the three sessions were highly correlated with results obtained from all three sessions combined. Spearman rank correlation averaged .85 for Participant 1 (range, .63 to .98), .74 for Participant 2 (range, .46 to .92), and .89 for Participant 3 (range, .82 to .96). Based on these findings, the authors suggested that a one-session MSWO assessment might be adequate. The average duration of the brief MSWO assessment (three sessions) was 4 min 49 s for Participant 1, 4 min 37 s for Participant 2, and 5 min 23 s for Participant 3.

A limitation of the Carr et al. (2000) study above was that the authors did not compare the 3-session assessment to a 5-session assessment. Therefore, it was not clear whether the preference hierarchies observed would have been comparable to a 5-session assessment. Graff and Ciccone (2002) addressed this limitation by comparing results obtained from an extended MSWO preference assessments, which consisted of seven

sessions of seven trials each, to results obtained in fewer sessions and trials. The authors also repeated the preference assessment with 12 of the 15 participants six months after the initial assessment yielding a total of 27 data sets. When results from the seven 7-trial sessions were compared to those obtained after five sessions, the same top-ranked stimulus was identified in 22 of the 27 data sets. When results from the seven 7-trial sessions were compared to those obtained after three sessions, the same top-ranked stimulus was identified in 19 of the 27 data sets. When results from the seven 7-trial sessions were compared to those obtained after seven sessions of three trials each, the same top ranked stimulus was identified in 25 of the 27 data sets. When the results from the seven 7-trial sessions were compared to those obtained after five 3-trial sessions, the same top ranked stimulus was identified in 22 of the 27 data sets. In addition, reinforcer assessments with four participants showed that all of the most-preferred stimuli identified based on five 3-trial sessions functioned as reinforcers.

In summary, the above research showed that the SS and the MS procedures are not very effective in differentiating preferences because individuals tend to approach the stimulus on every trial during the SS presentation or approach the same stimulus on every trial during the MS presentation. In contrast, both the PS and MSWO procedures have been shown to be effective in differentiating relative preferences among stimuli. While the PS procedure yields more clear preference hierarchies, the MSWO procedure requires less time to administer. Furthermore, Graff and Ciccone (2002) demonstrated how most-preferred stimuli, identified using five 3-trial sessions, also functioned as reinforcers which further decreased the time required to conduct the MSWO assessment.

Teaching Individuals to Conduct Stimulus Preference Assessment

To date, a small number of studies have evaluated procedures to train staff to conduct stimulus preference assessments. Lavie and Sturmey (2002) successfully trained three assistant teachers to conduct the PS procedure in a multiple-baseline design across participants. These assistant teachers had reported prior to the study that they had difficulties with identifying potential reinforcers to use for their students. Participants were given a blank piece of paper, pen, and the stimuli to be assessed during baseline to conduct stimulus preference assessments with a child with autism. After baseline, training was provided on how to conduct the PS preference assessment procedure. The authors completed a task analysis of the PS preference assessment procedure and identified nine steps required to successfully conduct one trial (i.e., put two stimuli on the table in front of the child and wait 5 s, if the child chooses a stimulus, then remove the other stimuli, etc.) and each session consisted of 21 trials. These steps were taught using a training procedure that consisted of nine training components: (1) describing the skills required for the assessment to the staff; (2) providing staff with a 9-item behaviour checklist; (3) describing the skills listed on the checklist to the staff; (4) showing a videotape of someone modeling the nine skills; (5) observing the staff conduct an assessment with a child with autism spectrum disorders; (6) providing feedback on skills executed correctly and incorrectly; and (7) repeating steps 4 through 6, if necessary, until at least 85% of steps contained on the checklist were performed correctly for two consecutive sessions. The mean accuracy during baseline was 16%, 23%, and 20% for the three participants, respectively. Following training, the mean accuracy increased to 98%, 100%, and 100% across the participants, respectively. The participants required two

to six training sessions, averaging 80 min of training per participant. One limitation of the study was that it did not compare the training procedure to an alternate procedure. Another limitation was that the 9-item checklist only included skills on how to present an assessment trial and did not include recording, ranking the items, or interpreting the results.

Roscoe, Fisher, Glover, and Volkert (2006) evaluated the relative effectiveness of performance feedback and contingent money on the acquisition of PS and MSWO preference assessment procedures. Participants consisted of four individuals with undergraduate degrees in Psychology or a related discipline and who had no prior experience in conducting preference assessments. During baseline, the experimenter asked each participant to conduct a PS or MSWO preference assessment with a simulated client (adult playing the role of a client) and with real clients. No reading materials and no instructions on the procedures were provided, and no feedback on performance was provided during or after the baseline assessments. Performance accuracy was evaluated on each trial using a behaviour checklist, which included six to nine behaviours depending on the responses of the person being assessed. The second phase of the study consisted of providing all participants with written instructions that consisted of brief summaries of the essential points to conduct the PS and MSWO procedures, adapted from the method sections of papers published by Fisher et al. (1992) and DeLeon and Iwata (1996). Participants were asked to assess simulated clients and real clients after reading the procedures. Data sheets specifically designed for the PS and MSWO procedures were provided. Except for the data sheets, the written instructions were not available during the assessments. The mean percent of correct responses

increased slightly from baseline for all four participants during the PS assessments and for three participants during the MSWO assessments. One participant achieved 85% correct after the written instruction during the MSWO assessment. During the third phase, performance feedback and contingent money were compared. In the feedback condition, the trainer reviewed each participant's last videotaped assessment, recorded the participant's target behaviours as correct or incorrect, and conducted a 5-10 min feedback session with the participant. During the feedback session, the trainer provided the participant with the completed data sheet from the preceding session. While viewing the taped session with the participant, the trainer identified each target behaviour as correct or incorrect, and explained why a target response was incorrect. All feedback statements were given in a neutral tone. During the contingent money condition, participants received a maximum of $10 contingent on the proportion of correct responses performed during the last session. No other feedback was provided. Participants 1 and 2 received feedback for the PS procedure and contingent money for the MSWO procedure. Participant 3 received feedback on the MSWO procedure and contingent money for the PS procedure. Participant 4 received contingent money for both procedures. Feedback resulted in performances of 90% correct or higher for at least three consecutive sessions for all three participants and for both preference assessment procedures with confederates. Contingent money did not produce large increases in correct responses for any of the participants in either procedure. However, when feedback was added to the condition, it resulted in criterion performance of 90% correct or higher across three consecutive sessions for all participants. Assessments with real clients matched the results obtained with confederates. Overall, performance feedback

was more effective than contingent money.

Roscoe and Fisher (2008) evaluated performance feedback and role-play for training staff to conduct the PS and MSWO procedures in a multi-element design. Participants were eight behavioural technicians with Bachelor degrees and who had some experience working with individuals with developmental disabilities before the study. During a baseline phase, participants were given 30 min to review a written summary of the procedure, extracted from the Fisher et al. (1992) paper, before they were asked to conduct an assessment with a confederate. This was repeated for the MSWO procedure using a procedural summary extracted from the DeLeon and Iwata (1996) paper. Participants were randomly assigned to two groups (four per group) following baseline. During the first training phase, Group 1 received feedback and role-play on the MSWO assessment while Group 2 received feedback and role-play for the PS preference assessment. During the second training phase, Groups 1 and 2 received feedback and role-play training for each of the PS and MSWO assessments in a counterbalanced order, prior to conducting the assessments with the confederates. Feedback and role-play sessions were 15-20 min long. Feedback consisted of reviewing the last videotaped session with the participant and identifying correct and incorrect responses, similar to Roscoe et al. (2006). In addition, feedback was also provided during a role-playing session, in which the experimenter played the role of a client and demonstrated all possible client responses that the participant may encounter. The participants were instructed to respond accordingly and the experimenter provided feedback on whether the behaviour was performed correctly. During baseline, Groups 1 and 2 showed mean correct responses of 44% and 47%, respectively, for the MSWO assessment, and mean

correct responses of 47% and 34%, respectively, for the PS assessment. Following the first feedback and role-play training phase on the MSWO procedure, Group 1 had a mean performance of 96% while Group 2, who had not received training on the MSWO procedure, had a mean performance of 54% and the difference was statistically significant ($p < .001$) with baseline performance as the covariate. Following feedback and role-play on the procedure, Group 2 had a mean performance of 94% while Group 1, who had not received training on the procedure, had a mean performance of 46% and the difference was statistically significant ($p < .001$) with baseline performance as the covariate. Following the second training phase, where both groups had received training on both procedures, the mean performance across all participants was 95% for the MSWO procedure and 96% for the PS preference assessment procedure. Feedback and role-play training produced correct responding for all 16 assessments (two each for eight participants), and over 90% in 14 assessments. The authors concluded that feedback and role-play were effective for teaching staff to conduct common stimulus preference assessments.

Machalicek et al. (2009) evaluated the use of videoconferencing to teach three graduate student teachers to conduct PS assessments with children with autism. Prior to the video conferencing training sessions, teachers were provided with a 5-step task analysis of the PS procedure and were asked to practice the protocol with their students. The teachers were also provided with a list of eight items to be assessed using the PS procedure and the sequence of pairings of the items across trials.

The videoconferencing training sessions consisted of setting up cameras in the classroom to ensure that trials were visible and audible to the supervisors. Supervisors

consisted of three doctoral students, who were board certified behaviour analysts and all had extensive experience conducting research on the assessment and treatment of children with developmental disabilities. Teachers wore headsets so that they could hear the supervisors. At the start of the videoconference training session, the supervisor indicated to the teacher where to begin on the list of pairings and asked the teacher to begin the trials. During each trial, the supervisor would collect data on the number of steps performed correctly, the item chosen by the child, and delivered positive feedback to the teacher following the successful completion of the trial. The supervisors interrupted immediately if an error occurred and the teacher was provided with specific corrective feedback. If an error occurred during a trial, then the trial was repeated.

All three teachers implemented the PS procedure with 100% accuracy during the videoconferencing training sessions. The authors suggested that the results might have been due to the teachers receiving the task analysis ahead of time and were asked to practice the protocol with the children prior to commencement of the videoconference training sessions. These results do support previous research that demonstrates the use of instruction coupled with positive feedback to produce rapid learning of assessment skills. However, the study did not demonstrate that the video conference sessions made a significant contribution to the teacher's skill acquisition.

Bishop and Kenzer (2012) evaluated group classroom instruction and in-vivo feedback to teach 11 behavioural therapists to conduct a brief PS preference assessment (i.e., a 1-minute procedure that consisted of one presentation of two items identified as highly preferred, when to conduct preference assessments, and how to interpret the data during therapy sessions with children with autism). During baseline, the experimenters

observed the therapists in the child's home and asked them to complete five discrete-trials teaching sessions. During training, group classroom instruction consisted of lecture, video modeling, role-play, and feedback with simulated clients. The lecture portion included the following topics: (a) a review of reinforcement and reinforcers, (b) types of reinforcers, (c) guidelines for delivering reinforcers, (d) the importance of frequent systematic preference assessments, (e) how to conduct a brief PS preference assessment, and (f) how to interpret the results of a preference assessment. Video-modeling was used to demonstrate the PS procedure. Role-play consisted of the therapists practicing the procedure with simulated clients while the experimenters provided feedback on their performances. Post-intervention observations were conducted immediately following training and the procedures were identical to baseline.

Participants' performance was scored using the following checklist: (a) identify an opportunity to conduct a preference assessment, (b) present two stimuli simultaneously and have the child select one, (c) re-present the selected stimulus with a third stimulus, (d) continue presenting stimulus pairs until the same stimulus is selected when presented with different stimuli on two consecutive trials, and (e) deliver the preferred stimulus during the subsequent trial block. An opportunity was defined as (a) the beginning of a trial block, (b) returning from a break longer than 30 s, (c) child refusing an item delivered, (d) child not consuming an edible, interacting with a tangible, or participating in an activity, and (e) child requesting a different item. Follow-up observations were conducted four weeks post-mastery. Follow-up observation and scoring procedures were identical to baseline.

If participants did not complete the steps at ≥ 75% accuracy during post-training observations, then in-vivo feedback was provided. In-vivo feedback consisted of the experimenter observing a therapy session and providing trial-by-trial feedback as needed to the therapist. Less intrusive prompts (i.e., vocal instruction and gestures) were used first, followed by more intrusive prompts (i.e., modeling a preference assessment with the child), if necessary. In-vivo feedback continued until the therapist was able to perform all steps of the preference assessments independently after which a second post-training observation was conducted. Data were also collected on the number of different stimuli assessed and identified as preferred during baseline and following training, the amount of time required to conduct the brief assessment, the number of trials needed to identify a stimulus as preferred, and the number of brief PS preference assessments conducted per session.

Results indicated that none of the participants met the mastery criterion (≥ 75% correct) during baseline (range, 0 to 43%). Seven of the 11 participants met the mastery criterion following group classroom instruction while the remaining four participants met criterion after in-vivo modeling. Eight participants were able to maintain their performance of ≥ 75% correct at 4-week follow-up. In addition, there was an increase in the variety of stimuli assessed during the therapy sessions with their clients. The authors concluded that group classroom instruction was usually sufficient to teach therapists the skills necessary to conduct preference assessments. Limitations that the authors cited included not knowing which component of the classroom instruction was the most efficient instructional method (i.e., lecture, video-modeling, role-play, or feedback). Furthermore, the interaction between group classroom instruction and in-vivo feedback

was unknown. In other words, perhaps in-vivo feedback in the absence of group classroom instruction may have been effective.

Pence, St. Peter and Tetreault (2012) evaluated the effectiveness of pyramidal training on teachers' implementation of three types of preference assessments: PS, MSWO, and free operant. Nine teachers participated in Experiment 1. All teachers reported prior to the start of the study that they had: (1) provided services to children with learning developmental disabilities (ranging from 3 to 25 years of experience), (2) were enrolled in a course sequence designed to prepare them to become board-certified behaviour analysts, (3) completed two graduate-level courses in behaviour analysis, (4) accrued practicum hours, and (5) received didactic instructions on preference assessment procedures. Prior to the start of Experiment 1, three of the nine teachers were trained on how to perform the three different types of preference assessments during their practicum supervision with a board-certified behaviour-analyst consultant and had met the mastery criterion defined as 90% accuracy for one session. These three first-tier teachers served as trainers in Experiment 1. Each of the three trainers were randomly assigned one type of preference assessment to train and the experimenter provided each of them with copies of the data sheets, preference assessment protocol that outlined the step-by-step breakdown of each preference assessment, and a training protocol. The training protocol included instructions that the trainer had to read out loud to the trainees, the frequency and types of errors that the trainer had to make during each session, and outlined how to provide positive and constructive feedback to trainees. The remaining six teachers served as trainees. Baseline and training sessions were conducted during regularly scheduled class periods for a course on behavioural assessment and intervention. Baseline consisted of

providing the six trainees with readings that described the three preference assessment types. The trainees were then required to conduct all three preference assessments with a trainer and were not provided with feedback. The six trainees' performances ranged from 0% to 80% during baseline across the three types of preference assessments.

During individual training sessions, the trainer stated to the trainee the type of preference assessment that he/she was going to learn, reviewed how to conduct the preference assessment, and modeled each step of the procedure. Trainees were not required to record data, rank the items, or interpret the results. The trainer was allowed to answer any questions from the trainee. After each training session, trainees were required to conduct the preference assessment with the trainer. If the trainee completed at least 90% of the steps correctly, then the trainees received training on the next preference assessment procedure. If the trainee did not meet the mastery criterion, the trainer would provide feedback on the errors observed during the session and modeled the correct response. Training continued until the trainee met the mastery criterion. Training resulted in rapid mastery for all three assessments across three of the six trainees. The three remaining trainees met the mastery criterion for the free-operant and PS procedures and maintained performance at high levels; however, two of the three participants required one additional training session for the MSWO procedure to meet the mastery criteria. The remaining participant's performance on the MSWO procedure increased to 80%, but due to trainer error, did not receive additional training. During generalization sessions, trainees implemented all three preference assessment types with children in their classrooms and clinics with 90% or greater accuracy 1 to 11 weeks after training. It took approximately 60 to 90 min to complete all baseline and training sessions.

Eight of the nine teachers in Experiment 1 participated in Experiment 2. The three first-tier teachers were now assigned to take procedural fidelity data and provide feedback to the five teachers (i.e., second-tiered trainers) on how to train 18 preschool teachers. Prior to the start of Experiment 2, all preschool teachers reported that they had received didactic instructions on preference assessments. The baseline and training sessions were identical to the first phase of the study, with the exception of the five second-tier teachers serving as trainers. During baseline, 15 of the 18 trainees demonstrated 0% accuracy across all three preference assessment types. The remaining three trainees' baseline performances ranged from 0% to 70%. Training resulted in immediate increases in performance accuracy for all 18 trainees in one to two sessions. Only three of the trainees required additional training session to meet the mastery criteria. Generalization sessions were omitted from the second phase of the study. It took approximately 60 to 120 min to complete all baseline and training sessions. The authors suggested that by training community members on how to conduct behaviour analytic procedures, the social validity and adoptability of these procedures could be improved in various community settings.

Most recently, Graff and Karsten (2012) evaluated the effectiveness of a self-instructional package to train staff to implement the MSWO and PS preference assessment procedures. They recruited 11 teachers who worked with individuals with autism and related developmental disabilities. All teachers held a bachelor's or master's degree and did not have any previous exposure to preference assessments. All participants completed a pretest designed to assess their knowledge of both procedures.

All participants met the inclusion criterion of 50% correct or lower (*M* = 28%; range, 0% to 50%).

Three different training tools were evaluated: Written instructions (baseline), enhanced-written instructions, and written instructions plus data sheet. The baseline procedure consisted of providing teachers with written instructions on how to conduct both preference assessments that were extracted from the method sections of previously published articles (DeLeon & Iwata, 1996; Fisher, et al., 1992). Teachers were given up to 30 minutes to read the instructions before they were asked to conduct their first preference assessment session with a simulated consumer. The enhanced-written instructions consisted of providing teachers with step-by-step instructions that included minimal technical jargon, diagrams on how to conduct the procedure, and a detailed data sheet. The written instructions plus data sheet phase included the baseline written instructions and a data sheet from the enhanced written instructions. Participants were allowed to refer to the written text while conducting assessments for all phases.

A multiple-baseline design across assessment procedures (MSWO and PS) was used to evaluate the different training tools for each teacher. The teachers were divided into two groups and each group was exposed to one of two training sequences. Group 1 consisted of six teachers who were exposed to written instructions (baseline), enhanced written instructions, and generalization probes. Group 2 consisted of five teachers who were exposed to written instructions (baseline), written instructions plus data sheet, enhanced written instructions, and generalization probes. Training sessions were conducted with one teacher at a time. Immediately following intervention, the teachers were asked to conduct preference assessments with simulated consumers. Data were

collected on how accurately the teachers conducted the procedure, recorded the data, summarized the results, and interpreted the assessment outcomes. The mastery criterion for each preference assessment was 90% or greater accuracy. Generalization probes with actual consumers were conducted after each teacher achieved the mastery criterion.

Results indicated that none of the six teachers in Group 1 met the mastery criterion for either preference assessment procedure using written instruction alone (*M* = 34%, range, 9% to 56% for PS; *M* = 46%, range, 13% to 61% for MSWO). Following enhanced instructions training, five of the six teachers met or exceeded the mastery criterion. None of the teachers in Group 2 met the mastery criterion for either preference assessment procedure during baseline (PS procedure *M* = 38%, range, 20% to 64%; MSWO procedure *M* = 38%, range, 8% to 58%). Following written instruction plus data sheet, mean performance accuracy of PS assessments increased to 73% (range, 65% to 89%) and the mean performance accuracy of MSWO assessments increased to 59% (range, 12% to 79%). When enhanced written instructions were introduced, the total accuracy further increased performance accuracy to a mean of 98% for the PS procedure and 99% for the MSWO procedure. Participants implemented both assessment procedures between 80-100% accuracy during the generalization probes with actual consumers. Lastly, participants completed a social-validity questionnaire following the enhanced-instructions and written-instructions training phases. Results indicated that the teachers preferred the enhanced-instructions over the written instructions because it was easier to use and that they would use the enhanced-instructions in the event that they had to train new staff. The results of this study suggests that written instructions that include diagrams and minimize technical jargons may be sufficient to train staff without having

to provide individualized feedback thereby reducing face-to-face time with a trainer. The authors noted that one limitation to their study was that all participants were exposed to the method description prior to receiving the enhanced-written instructions and therefore, the results obtained may have been influenced by a sequence effect.

Statement of the Problem

Previous studies have demonstrated that direct stimulus preference assessments can be used to identify reinforcers for individuals with intellectual or developmental disabilities. Although it would be desirable to train practitioners to use stimulus preference assessments, only a small number of training studies have been conducted. Further investigations are needed to address some limitations of those training studies, and to extend this area of research. First, limited data are available on maintenance of the learned skills and on generalization to assessing an individual with disabilities following training. Second, individual or small-group training requiring face-to-face instruction can be resource intensive if a large number of individuals require training. It would be highly beneficial to develop a training tool that is effective and requires minimal resources. To date, only one study has trained individuals to conduct preference assessments using written instructions (Graff & Karsten, 2012). However, one of the limitations to the Graff and Karsten study was that the written instructions were preceded by the method description and a sequence effect might have influenced the results. It should be noted that the paper by Graff and Karsten was published after the present study had been initiated.

Self-instructional training manuals have been used successfully to teach students and staff to conduct assessments with individuals with intellectual and developmental

disabilities including discrimination skills (DeWiele, Martin, & Garinger, 2000), adaptive behaviours (Yu, Martin, Hardy, Leader, & Quinn, 1985), and discrete-trials teaching (Arnal et al., 2007; Fazzio, Martin, Arnal, & Yu, 2009). The above self-instructional manuals involve presenting complex procedures in small units, describing the procedures in nontechnical language, providing study questions or exercises after each unit on skills that need to be learned, and a final review exam at the end of the manual for the reader to evaluate their mastery of the information. Therefore, I have prepared a self-instructional training manual for the MSWO preference assessment procedure (Ramon & Yu, 2010) and incorporated the above features into the manual.

The purpose of this study was to evaluate the manual's effectiveness in teaching undergraduate university students to conduct the MSWO preference assessment. Retention and generalization of the learned skills to clients were examined. Ethical approval was obtained from the University of Manitoba Psychology/Sociology Research Ethics Board and written informed consent was obtained from participants or substitute decision makers before the study began.

Method

Participants and Setting

Eighteen undergraduate students from the University of Manitoba were recruited in two waves. The first wave occurred in the Fall of 2010 and included three males and seven females from a second year psychology course. None of the students reported that they had any prior training on conducting preference assessments. Each participant received an honourarium of $45 for participating in the study regardless of his or her performance. Participants also received 2% credit for their course from the instructor for

taking part in the study, regardless of their performance. The Experimenter notified the instructor of a student's participation in the study, but did not indicate how he/she had performed.

The second wave of student participants was recruited in the Fall of 2011 and included two males and six females from a second year psychology course. None of the students reported that they had any prior training on conducting preference assessments. Each participant also received an honourarium of $45 for participating in the study regardless of his or her performance. However, the course instructor was different from the one in the first wave and participants received 10% credit for their course from their course instructor for taking part in the study, regardless of their performance. As with the first wave, the Experimenter notified the instructor of a student's participation, but did not indicate how he/she had performed.

The age range, sex, and years of university for the participants are shown in Table 1. The characteristics of the two waves of participants were fairly comparable except for years of university. The mean ages of each wave was computed using 20 for ≤ 20, the mid-point for each range, and 50 for ≥ 50. The mean ages were 23.8 years and 26.3 years for the first and second waves, respectively. Seventy percent of the first wave were females and 75% of the second wave were females. On average, the first wave had 2.9 years of university whereas the second wave had 2.0 years.

Table 1

Characteristics of Student Participants

Participant	Age Range (Years)	Sex	Years of University
First Wave			
1	21-25	F	2
2	≤ 20	F	2
3	21-25	M	4
4	21-25	F	2
5	41-45	F	6
6	≤ 20	M	4
7	21-25	F	3
8	≤ 20	M	2
9	21-25	F	2
10	≤ 20	F	2
Second Wave			
11	21-25	M	2
12	≥ 50	F	2
13	21-25	F	2
14	21-25	F	2
15	≤ 20	F	2
16	≤ 20	F	2
17	21-25	M	2
18	26-30	F	2

Two clients with developmental disabilities were recruited from a community residential facility to participate in the study for generalization assessment. Client 1 was a 47-year-old man who was non-verbal and non-ambulatory, and his functioning level was severe according to his health record. Client 2 was a 46-year-old woman who was non-verbal and non-ambulatory, and her functioning level was moderate.

All sessions were conducted at St.Amant Research Centre in a room furnished with a large table, a small table, two chairs, a tripod, and a video camera. The Experimenter was present during all sessions. During simulated or client assessments (described below), the student participant sat at the large table and items required to conduct the preference assessment were placed on the small table next to the participant. During all assessments, the person being assessed sat across the table facing the student participant.

Materials

All sessions were videotaped, with consent, for scoring. Training materials included the MSWO self-instructional manual and a method description of the MSWO procedure (described below). Data sheets, stopwatch, calculator, and a variety of items used during preference assessments were provided to participants at the beginning of all assessment sessions. The items used during preference assessments are shown in Table 2.

Research Design

For the first wave of participants, an unbalanced crossover design (Toutenburg, 2002) with matched-pairs random assignment was used to compare the interventions (self-instructional *manual* vs. *method* description). In this design, pairs of participants were matched, as closely as possible, based on their performance scores obtained during

Table 2

Items Used for Preference Assessment

First Wave (Participants 1-10)

Simulated Baseline and Post-Training Assessments	Simulated Retention/Generalization Assessment with New Actor and Items
Playdoh®	maraca
squishy toy	beading
toy cars	toy dump truck
wooden blocks	stickers
Lego®	puzzle
Book	xylophone

Second Wave (Participants 11-18)

Simulated Baseline and Post-Training Assessments	Post-Training Assessment with Client 1	Post-Training Assessment with Client 2
toy train	wrestling calendar	hand lotion
Lego®	Playdoh®	glow stick
Playdoh®	Batman figurine	Lego®
glow stick	hand lotion	Playdoh®
wooden blocks	glow stick	comb
toy car	squishy toy	nail polish

the first simulated baseline preference assessment (described below). Members of each pair were then randomly assigned to two groups. Group 1 was trained using the manual first, followed by the method; whereas the order was reversed with Group 2. The crossover design was "unbalanced" in that only participants who did not meet the mastery criterion (defined as 85% correct or higher) after the first intervention received the second intervention. The mastery criterion was based on previous training studies (range, 75% to 90%). Participants who met the mastery criterion after either the first or second intervention were asked to return to complete a simulated retention/generalization assessment (described below).

Participants who did not meet the mastery criterion after both interventions were offered a third intervention that involved observing a model (live demonstration of the procedure). Participants in the second wave who met the mastery criterion were asked to return to complete a retention/generalization assessment after this intervention. The study phases are shown in Table 3.

In addition to the crossover design, a modified multiple-baseline design (Johnston & Pennypacker, 2009; Martin & Pear, 2011) was implemented across four participants within each group of the first wave. Scheduling difficulties prevented the fifth participant in each group from being included in the concurrent time series. For the modified multiple-baseline design, an initial simulated baseline assessment was completed for all four participants. The intervention was then introduced for the first participant only. After the first participant had received the post-intervention simulated assessment, the simulated baseline assessment was repeated for the next participant before the

Table 3

Study Phases

Baseline	Group	Training 1	Post-Training 1	Training 2[a]	Post-Training 2[a]	Training 3[b]	Post-Training 3[b]	Retention/ Generalization[c]	Client Assessment[d]
Simulated Assessment	1	Manual	Simulated Assessment	Method	Simulated Assessment	Modeling	Simulated Assessment	Simulated Assessment with New Actor	Assessment with Client
	2	Method		Manual					

[a] Training 2 and Post-Training 2 were provided for participants who did not meet criterion at Post-Training 1.

[b] Training 3 and Post-Training 3 were provided for participants who did not meet criterion at Post-Training 2.

[c] Assessments were conducted with a new actor and new set of activities. In the first wave (Participants 1-10), retention/generalization was evaluated for participants who met the mastery criterion at Post-Training 1 or at Post-Training 2. In the second wave (Participants 11-18), retention was assessed regardless of whether the mastery criterion was met at Post-Training 1, 2 or 3.

[d] Client assessment (i.e., assessments conducted with a person with developmental disabilities) was evaluated for participants who met the mastery criterion at Post-Training 1, Post-Training 2, or Post-Training 3. Client assessments were conducted only by the participants in the second wave (Participants 11-18).

intervention was introduced for that participant. This was repeated for each participant successively.

For participants in the second wave, due to time constraints (course coming to an end) and scheduling difficulties, no attempt was made to implement the modified multiple-baseline design or to match the participants. However, participants were assigned randomly to the two groups (in alternation in the order the participants were recruited).

Simulated Baseline Preference Assessment Procedure

Simulated baseline preference assessments were conducted before training. During a simulated baseline assessment, a participant was asked to conduct a preference assessment with an actor (graduate student). The participant was informed that the actor was a graduate student who was playing the role of a person with developmental disabilities with no speech. At the beginning of the assessment, the participant received a bin containing six items (see Table 2 for items used), an MSWO Preference Assessment Baseline data sheet, pencil, calculator, stopwatch, and the following written instructions:

> Thank you for helping me with this study. Today, you will conduct a multiple stimulus without replacement preference assessment using six items with <name of actor>. I cannot provide you with any additional information about the assessment procedure. Please do your best to find out what <name of actor> likes and dislikes. You can begin whenever you are ready. You have up to 20 minutes to complete the assessment or you can let me know when you are

finished and we can stop. Let me know when you are ready to begin.

The baseline data sheet (see Appendix A) provided a space for: (a) the client's name, (b) date, (c) tester's name, (d) space for three 5-trial sessions to record which item was selected on each trial, (e) a space for participants to list the items used from most to least preferred, and (f) an interpretation question (i.e., which item is likely to be the most powerful reward). Instructions were not provided on the data sheet on how to determine the ranking of the items.

During all simulated assessments, the actor's responses on each trial were scripted (see Appendix B for a sample script). The same script was used for all participants during each phase, but different scripts were used across phases. The script ensured that all skill items on the behaviour checklist was sampled equally often across participants. The actor had access to the script during each session and the script was placed out of the participant's view.

Target Behaviours

Each simulated and client assessment consisted of three sessions and each session was scored using a behaviour checklist (Appendix C). The first section of the checklist, *Preparation Before an Assessment*, has three target responses: (1) enter the client's (person being assessed) name, the tester's name, date of the assessment, and the items to be assessed on the data sheet; (2) prepare the edible items to be assessed if edibles were being used; and (3) provide the person being assessed an opportunity to sample each item by allowing them to consume each edible or interact with each leisure item for up to 30 s. These responses were expected to occur only once per assessment, at the beginning of the

first session. The second section, *Antecedent Behaviours,* has three target responses: (1) ensure the person being assessed had attended to each item by providing the verbal instruction “look” when introducing each item singly in front of the person; (2) present the correct items on the table and place the items in an order different from the last trial; and (3) deliver the verbal instruction “pick one” after placing the items on the table. The third section, *Consequences for Responses,* consists of four to eight target responses depending on the response of the person being assessed. There are four target behaviours following an appropriate selection, six target behaviours following attempts to select two items simultaneously, and up to eight target behaviours following no selection after 15 s. The responses in the second and third sections of the checklist were applicable to every assessment trial and were scored for each of the three sessions. The last section of the checklist, *Summarizing and Interpreting the Results*, consists of two target responses: (1) identify the ranking of each item correctly based on the assessment results; and (2) identify which item would serve as the most effective reward. These responses were expected to occur only once per assessment, at the end of the third session.

The percentage of correct responses per session was computed for each participant. The percentage correct for Session 1 was based on the participant's performance on the first section of the checklist and the first five trials of the preference assessment. The percentage correct for Session 2 was based on the participant's performance on the second set of five trials of the preference assessment. The percentage correct for Session 3 was based on the participant's performance on the last set of five trials and the last section of the checklist.

Method Description Training

The method description, shown in Appendix D, was adapted from the method sections of published papers by DeLeon and Iwata (1996) and Roscoe et al. (2006). Adaptations were made so that the procedures would match the one described in the self-instruction manual. For example, modifications included describing the procedure in the present tense, modifying the number of items from seven to six, reducing the number of sessions from five to three, replacing the term "participant" with "client", and replacing the term "experimenter" with "staff".

During the training session, participants were presented with the method description and were given as much time as they needed to read the written procedures. The amount of time taken by each participant to study the materials was recorded by the Experimenter.

Simulated Post-Method Assessment

The simulated post-training simulated assessment commenced immediately after the participant had finished studying. The procedures were the same as the simulated baseline assessment described above except that the method data sheet was provided (see Appendix E). The method data sheet was similar to the baseline data sheet with the following exceptions: There was a space for participants to determine the percentage selection for each item used and instructions on how to determine the rank of each item, extracted from the method description, were included in the data sheet. The participant did not have access to the method description during the simulated assessment.

Self-Instructional Manual Training

The self-instructional manual was written in nontechnical language. The manual starts with an overview on preference assessment and its importance for people with developmental disabilities (1 page). It then outlines the steps required to conduct an MSWO preference assessment and the consequences provided for each potential client response (6 pages). Next, the manual describes how to summarize and interpret the results of the preference assessment (3 pages). The target behaviours are highlighted in each section of the manual. Review exercises, with answer keys, were included at the end of each section and at the end of the manual. A final review exercise, with answer key, was also included at the end of the manual and reviewed the key target behaviours from all sections of the manual (see Appendix F for the manual's Table of Contents). In the self-instructional manual, the reader is asked to complete the exercise at the end of the unit without referring to the preceding section and then check their answers against the answer key. The manual also instructs the reader to review the relevant section(s) of the unit for each error before proceeding to the next unit.

A 25-item behaviour checklist for the MSWO preference assessment was provided at the end of the manual (Appendix G). The face validity of the behaviour checklist was evaluated prior to the commencement of the study by three graduate students who had each conducted at least 30 MSWO preference assessments with clients. The graduate students were asked to complete a questionnaire (see Appendix H) to rate the importance of each item on the checklist using a 7-point scale (1 = not important; 4 = neutral; 7 = essential). The 25 items received a mean rating of 6.8 (range, 5.7 to 7.0).

Simulated Post-Manual Assessment

The simulated post-manual assessment commenced immediately after the participant had finished studying the self-instructional manual. The procedures were the same as the simulated baseline assessment except for the data sheet. The data sheet from the self-instructional manual (see Appendix I) included space and instructions for participants to determine the preference ranking for each item. The participant did not have access to the manual during the simulated assessment.

Modeling Procedure

Modeling was provided for participants who did not meet the mastery criterion after receiving both the method description and manual training, regardless of order. This intervention consisted of the Experimenter demonstrating the assessment with an actor covering all the target behaviours. After the Experimenter had modeled the assessment, the participant could request to see specific portions of the demonstration again as many times as needed, but he/she could not ask the Experimenter any questions related to the procedure.

Simulated Post-Modeling Assessment

The simulated post-modeling assessment procedures were the same as in the simulated post-manual assessment.

Simulated Retention/Generalization Assessment with Novel Actor and Items

Participants were asked to return approximately one week after an intervention to conduct a simulated assessment with a new actor and new items. However, the retention interval was usually longer due to participants' availability. Participants in the first wave who met the mastery criterion following modeling were not asked to conduct the

retention/generalization assessment. The procedures were the same as that used during the simulated baseline assessment except for the new actor and items.

Preference Assessment with a Client

Generalization assessment with a real client following training was not conducted with participants in the first wave due to difficulties with scheduling the client. Only one child with autism was recruited and he was hospitalized due to an illness two days before the assessment. As a result of his illness, the family withdrew from the study.

Participants in the second wave who met the mastery criterion, regardless of whether it followed manual training, method training, or modeling, were asked to return approximately one-week after training to conduct a preference assessment with a client with developmental disabilities. The procedures were the same as that used during the simulated baseline assessment except that the assessment was conducted with a real client and new items.

Social Validity

After each intervention, all participants were asked to complete a brief 6-item survey (see Appendix J). Statements in the survey were designed to evaluate the importance of the intervention (one item), the acceptability of the training procedure (two items), the perceived effectiveness of the training (two items), and how likely participants would use preference assessments if they were working with a person with developmental disabilities (one item). Participants indicated the extent to which they agreed with each statement (1 = strongly disagree, 2 = disagree, 3 = neither agree/disagree, 4 = agree, and 5 = strongly agree).

Inter-observer Reliability

Observer training. Before conducting reliability checks, all observers were trained by reviewing the MSWO assessment procedure (i.e., method description and self-instructional manual), rehearsing the procedure with the Experimenter, and practicing scoring using a videotaped simulated assessment that depicted how to sample each item, how to present items prior to the start of each trial, and the consequences for all potential client responses. Training continued until the observer scored 10 consecutive trials (two 5-trial blocks) correctly and this was determined by comparing their scores to the Experimenter's scores.

Reliability checks on dependent measure. All simulated and client assessments conducted by participants were reviewed for reliability checks. The Experimenter and an observer scored the assessments independently (live or from videotape) using the data sheet shown in Appendix C, and then compared their recordings. The Experimenter scored each participant's performance live during 95% of the assessments and the remaining 5% were scored from videotapes. The observer scored most of the participants' performances from videotape (95%). On each trial, a checklist item was scored as an agreement if both the Experimenter and the observer had the same recording (i.e., correct, incorrect, or not applicable for each item). A disagreement was scored if the recordings differed. The percentage of agreement per session was calculated by dividing the number of agreements by the sum of agreements and disagreements, and multiplying by 100%. The mean agreement score for all sessions was 96% (range, 73% to 100%). Note that the lower IOR scores occurred during simulated baseline assessments.

Actor's script adherence checks. The actor's adherence to the scripted responses on each trial was also evaluated for all simulated assessments. Percentage of script adherence was computed for each session by dividing the number of scripted responses that occurred correctly by the total number of scripted responses and multiplied by 100%. The mean adherence score across all sessions was 98% (range, 82% to 100%).

Results

Individual Results for Participants from the First Wave

Figure 1 shows the percentage of correct responses per session during simulated assessments before training (baseline), after intervention, and during retention/ generalization for each participant in the first wave (Participants 1 to 10). The horizontal dash line in each graph indicates the 85% mastery criterion. Participants 1, 3, 5, 7, and 9 (left column, Group 1) received the self-instructional manual training first, followed by the method description training and then modeling, if necessary. Participants 2, 4, 6, 8, and 10 (right column, Group 2) received the method description training first, followed by the self-instructional manual training and then modeling, if necessary.

In the modified multiple-baseline design across four participants in Group 1 (top four graphs in the left column of Figure 1), all four participants showed clear and immediate improvement in performance after manual training, strongly suggesting that the observed improvement was due to the manual intervention. The fifth participant (Participant 9) in Group 1 showed improvement during the first two sessions of the post-manual simulated assessment, but performance accuracy declined to near zero in the third session because she missed conducting the third session (last five trials) and computed the ranking based on the first two sessions. Despite the improvement in performance

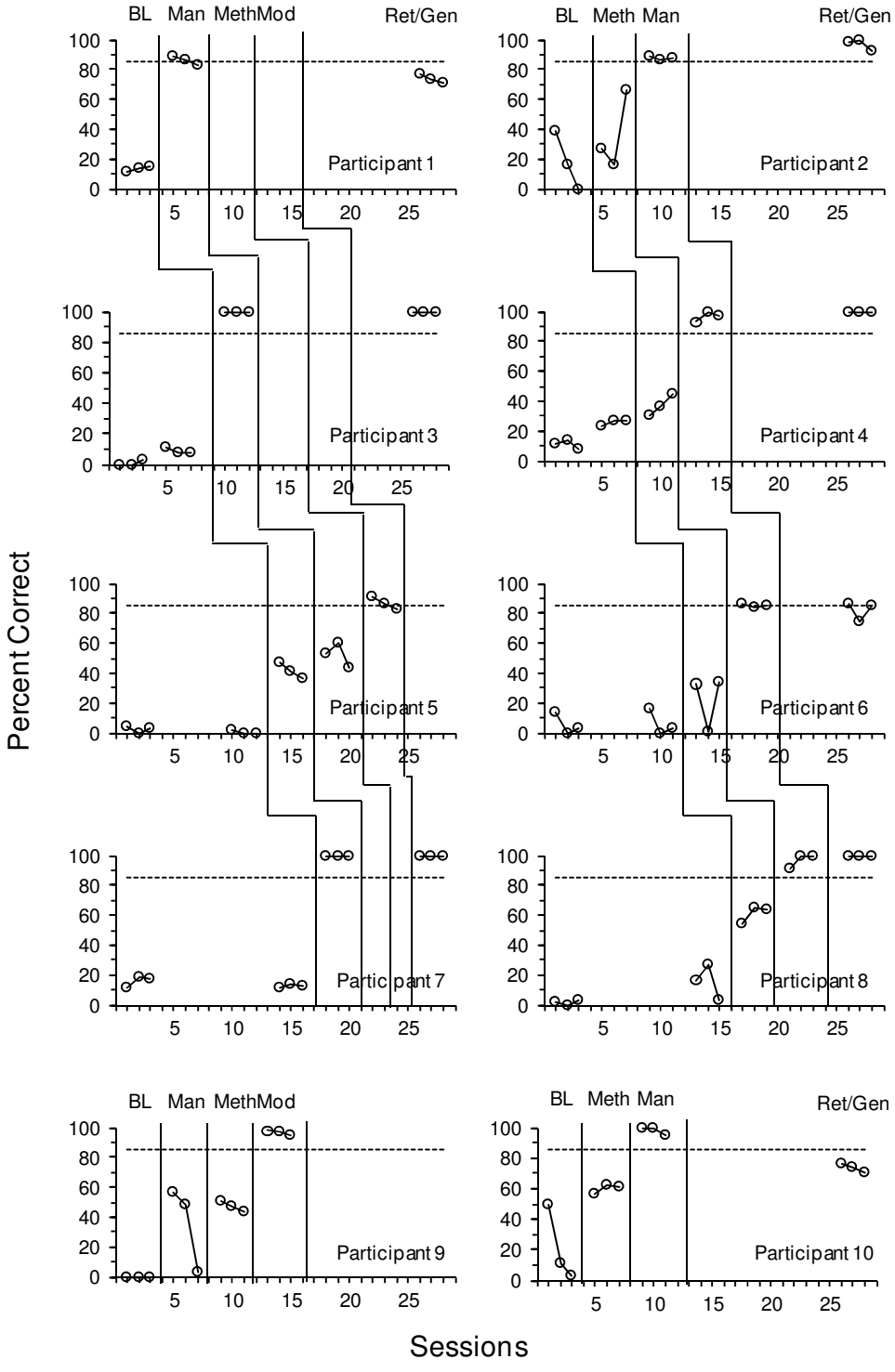

Figure 1. Percent correct responses during simulated preference assessments before and after each intervention, and during retention/generalization assessments for Participants 1 to 10. Group 1 (left column) received self-instructional manual first and Group 2 (right column) received method description first. BL = baseline; Man = manual; Meth = method; and Ret/Gen = retention/generalization.

accuracy for all five participants, only three met the mastery criterion (Participants 1, 3, and 7). Participants 5 and 9, who did not meet the mastery criterion, were given the method description training and neither showed much improvement after studying the method description. Modeling was then provided after which both participants met the mastery criterion.

Participants 1, 3, and 7 in Group 1 were available for the simulated retention/generalization assessment (with a novel actor and different items) ranging from 17 days to 22 days (M = 19.3) after they achieved the mastery criterion. Participant 1's performance declined to a 3-session mean of 73.7% (below the mastery criterion) during the simulated retention/generalization assessment, but both Participants 3 and 7 maintained their performance at 100% for retention/generalization assessments.

The method description training was evaluated in a modified multiple-baseline design across four participants in Group 2 (top four graphs in the right column of Figure 1). Improvements were small at best for Participants 2, 4, and 6. Participants 2 and 6's improvement was variable and overlapped with baseline during two and one sessions, respectively. Participant 4's improvement appeared to be consistent with a gradual upward trend from baseline. Improvement for Participant 8 was clear (no overlap with baseline) and moderate. The fifth participant (Participant 10), who was not part of the modified multiple-baseline design, showed similar performance as Participant 8. Thus, none of the five participants in Group 2 met mastery criterion after studying the method description. All five participants were then exposed to the self-instructional manual and all met the mastery criterion after studying the manual.

All participants in Group 2 were available for the simulated retention/ generalization assessment ranging from 7 to 22 days with one outlier at 64 days (*Mdn = 16* days) after they had achieved the mastery criterion. Participants 2, 4, and 8 maintained their accuracy above the mastery criterion during the simulated retention/generalization assessment with 3-session means of 96.8%, 100%, and 100%, respectively. Participants 6 and 10's performance accuracy declined slightly, averaging 81.9% and 73.7%, respectively.

Individual Results for Participants from the Second Wave

Figure 2 shows the percentage correct responses per session during simulated assessments before training (baseline), after intervention, and during the simulated retention and generalization assessment for each participant in the second wave (Participants 11 to 18). The horizontal dash line in each graph indicates the 85% mastery criterion. Participants 11, 13, 15, and 17 (left column, Group 1) received the self-instructional manual training first, followed by the method description training and modeling, if necessary. Participants 12, 14, 16, and 18 (right column, Group 2) received the method description training first, followed by the self-instructional manual training and modeling, if necessary.

Similar to the results of the participants in the first wave, performance accuracy improved for all four participants in Group 1 (left column, Figure 2) after studying the manual. However, only Participant 13 met the mastery criterion. Therefore, Participants 11, 15, and 17 received the method description training for their second intervention and only Participant 17 met the mastery criterion after the method description training.

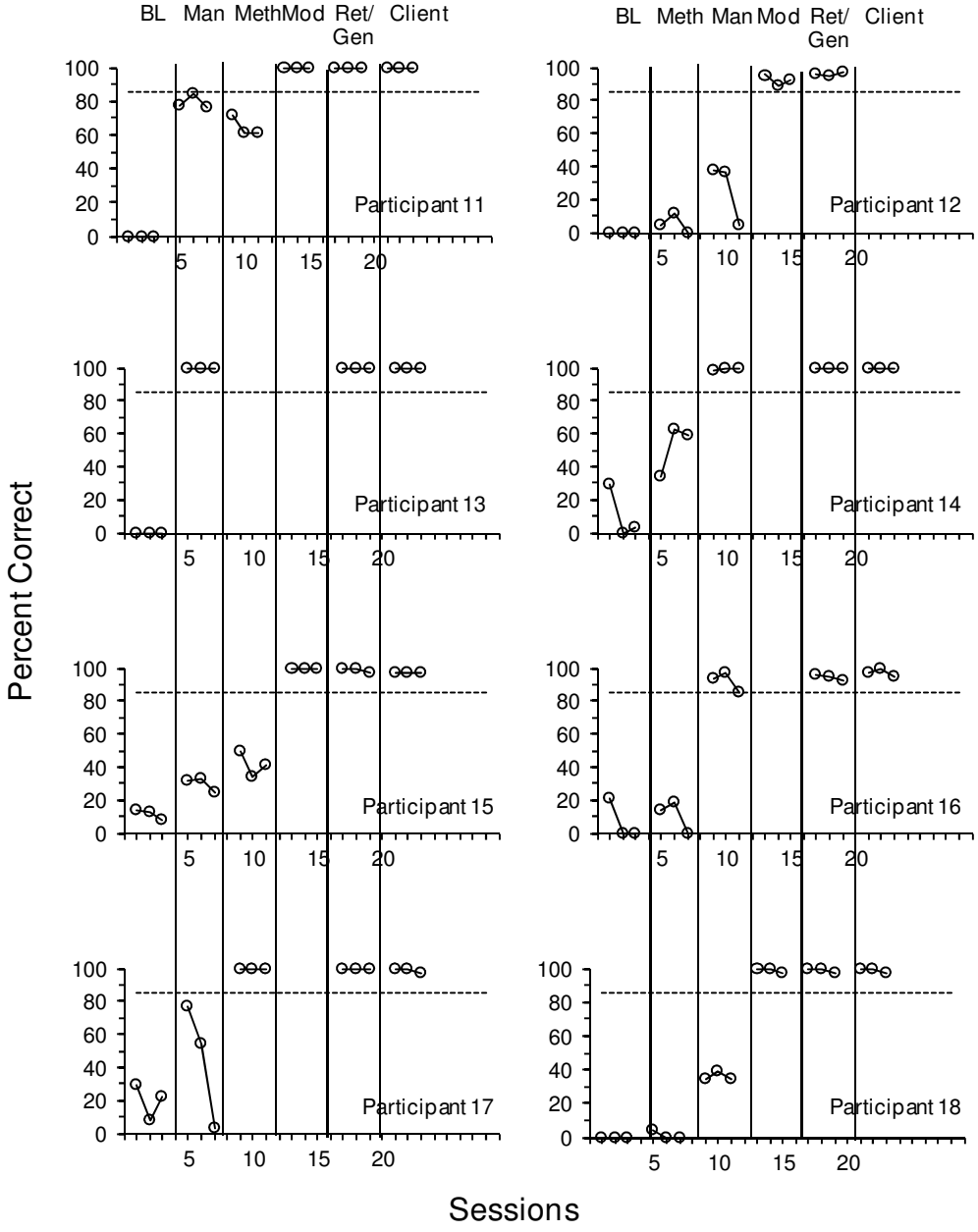

Figure 2. Percent correct responses during simulated preference assessments before and after each intervention, during retention/generalization assessments, and client assessments for Participant 11 to 18. Group 1 (left column) received self-instructional manual first and Group 2 (right column) received method description first. BL = baseline; Man = manual; Meth = method; Ret/Gen = retention/generalization assessment; and client = assessment with a real client.

Participants 11 and 15 were then exposed to modeling and both met the mastery criterion after observing the live demonstration. The simulated retention/generalization assessments and client assessments were completed for all four participants in Group 1 (left column, Figure 2), ranging from 5 to 7 days (M = 6.5 days) after they had achieved the mastery criterion. All participants performed near 100% accuracy in all sessions (range, 97.4% to 100%).

For the four participants in Group 2 (right column), method description resulted in small (Participants 12, 16, and 18) to moderate (Participant 14) improvements in performance accuracy. However, none of the participants met the mastery criterion. All participants went on to receive the self-instructional manual and two (Participants 14 and 16) met the mastery criterion after studying the manual. Participants 12 and 18, whose performance accuracy had increased slightly after manual training, but did not reach mastery, met the mastery criterion after modeling.All four participants in Group 2 (right column, Figure 2) completed the simulated retention/generalization assessments with a novel actor and different items, and all participants performed above the mastery criterion in all sessions (M = 97.3%, range, 92.7% to 100%). Three of the four participants in Group 2 completed the retention/generalization assessments with a client ranging from 5 to 9 days (M = 7.0) after they had met the mastery criterion and all participants performed above the mastery criterion in all sessions (M = 98.9%, range, 94.7% to 100%).

Inter-tester Agreement on Results of Assessment with Clients

Since each client was assessed by more than one participant, the inter-tester agreement was examined by comparing the results obtained by different participants. Table 4 shows the preference ranking obtained for each item by each participant for the

Table 4

Mean Item Ranking Obtained by Student Participants during Generalization Assessments with Clients

Client 1

Items Assessed	*Participant 11*	*Participant 15*
Wrestling calendar	1	1
Glow stick	2	3
Batman figurine	2	4
Playdoh®	2	5
Squishy toy	5	2
Hand lotion	6	6

Client 2

Items Assessed	*Participant 13*	*Participant 14*	*Participant 18*
Playdoh®	1	1	3
Silver nail polish	2	2	1
Hand lotion	3	3	4
Comb	4	5	5
Lego®	5	6	6
Glow stick	6 [a]	4 [a]	2 [a]

Client 2

Item Assessed	*Participant 16*	*Participant 17*
Glow stick	1 [a]	1 [a]
Comb	1	3
Gold nail polish	3	4
Silver nail polish	4	5
Hand lotion	5	6
Lego®	6	2

[a] Variability of rating could be due to client's change in preference during assessments.

same client. For Client 1, the wrestling calendar was ranked highest during assessments by Participants 11 and 15. For Client 2, Playdoh® was ranked highest during assessments by Participants 13 and 14 and ranked third for Participant 18. Lastly, when Playdoh® was replaced by gold nail polish for Client 2, glow stick was ranked highest during assessments by Participants 16 and 17 (although comb tied for first for Participant 16). The variability for the glow stick could have reflected a change in Client 2's preference over the course of the study. Overall, there was moderately-high agreement on items that were identified as the most-preferred for the same client.

Group Comparisons

Baseline to Post-Training 1. In order to determine whether Group1 (manual first) performed differently than Group 2 (method first) at post-training 1, a 2 x 2 repeated measures analysis of variance (RM ANOVA) was conducted, with group (manual vs. method) as the between-subjects factor and time (baseline vs. post-training 1 assessments) as the within-subject factor. All 18 participants from both waves were included for this analysis, which yielded nine participants for each intervention group. All tests were evaluated with $p = .05$, two-tailed.

Results from the RM ANOVA showed a significant main effect for time and a significant group x time interaction (see Table 5 and Figure 3). While both groups showed statistically significant increases in performance accuracy at post-training 1, the manual group showed a much larger improvement (mean increase of 61.26%) relative to the method group (mean increase of 21.86%). Although improvement in performance accuracy at post-training 1 was statistically significant, only four of the nine participants

Table 5

Performance Accuracy (Percent Correct) During Simulated Assessments at Baseline and Post-Training 1

	Baseline *Mean* (*SD*)	Post-Training 1 *Mean* (*SD*)	Δ score, *p*
Group 1 Manual (*n* = 9)	7.21 (7.54)	68.47 (30.18)	61.26, $p < .001$
Group 2 Method (*n* = 9)	9.86 (7.78)	31.72 (23.00)	21.86, $p = .007$

Repeated Measures ANOVA (group x time) – Significant main effects of time: Wilks' $\Lambda = .255$; $F(1, 16) = 46.66$, $p < .001$; partial $\eta^2 = .745$; and significant interaction: Wilks' $\Lambda = .604$; $F(1, 16) = 10.48$, $p = .005$; partial $\eta^2 = .396$.

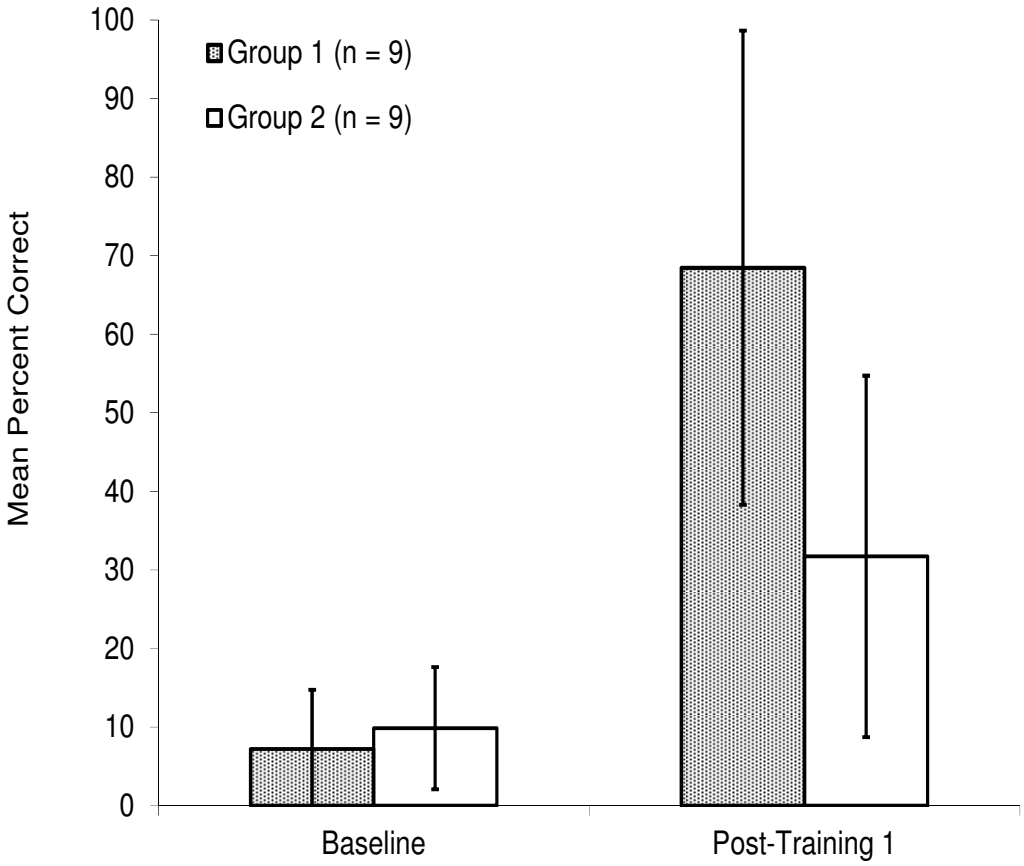

Figure 3. Mean percent correct during simulated preference assessments at baseline and post-training 1 for Groups 1 (self-instructional manual) and 2 (method description). Error bars represent +/– one standard deviation.

in Group 1 and none of the Group 2 participants met the mastery criterion after studying the method description, as reported earlier in the individual results.

Baseline to Post-Training 2. To determine whether Group 1 (manual, then method) performed differently than Group 2 (method, then manual) at post-training 2, a 2 x 3 RM ANOVA was conducted, with group (manual vs. method) as the between-subjects factor and time (baseline vs. post-training 1 vs. post-training 2) as the within-subject factor. Only participants who received both interventions were included for this analysis (recall that only participants who did not meet the mastery criterion went on to receive the second intervention). Therefore, this analysis included five participants for Group 1 (Participants 5, 9, 11, 15, and 17) and all nine participants for Group 2.

Results from the 2 x 3 RM ANOVA showed a significant main effect for time and a significant group x time interaction (see Table 6 and Figure 4). Both groups showed significant mean performance improvements from baseline to post-training 2, with Group 2 showing a larger improvement (mean increase of 70%) relative to Group 1 (mean increase of 54.66%).

Group 1 received training using the self-instructional manual followed by training using the method description. Mean performance accuracy for Group 1 improved significantly after the manual intervention at post-training 1. There was a further mean increase of 15.18% after exposure to the method description at post-training 2, but that was not statistically significant (see Table 6). Although the overall mean increase for

Table 6

Performance Accuracy (Percent Correct) During Simulated Assessments at Baseline, Post-Training 1, and Post-Training 2

	Baseline *Mean* (*SD*)	Post-Training 1 *Mean* (*SD*)	Post-Training 2 *Mean* (*SD*)	Baseline to Post-Training 1 Δ score, *p*	Post-Training 1 to Post-Training 2 Δ score, *p*	Baseline to Post-Training 2 Δ score, *p*
Group 1 Manual-Method ($n = 5$)	6.56 (8.84)	46.04 (19.23)	61.22 (23.27)	39.48, $p = .020$	15.18, $p = .248$	54.66, $p = .003$
Group 2 Method-Manual ($n = 9$)	9.86 (7.78)	31.72 (23.00)	79.86 (28.11)	21.86, $p = .007$	48.14, $p = .001$	70.00, $p = .001$

Repeated Measures ANOVA (group x time) – Significant main effects of time: Wilks' $\Lambda = .101$; $F(2, 24) = 57.51$, $p < .001$; partial $\eta^2 = .827$; and significant interaction: Wilks' $\Lambda = .592$; $F(2, 24) = 4.03$, $p = .031$; partial $\eta^2 = .251$.

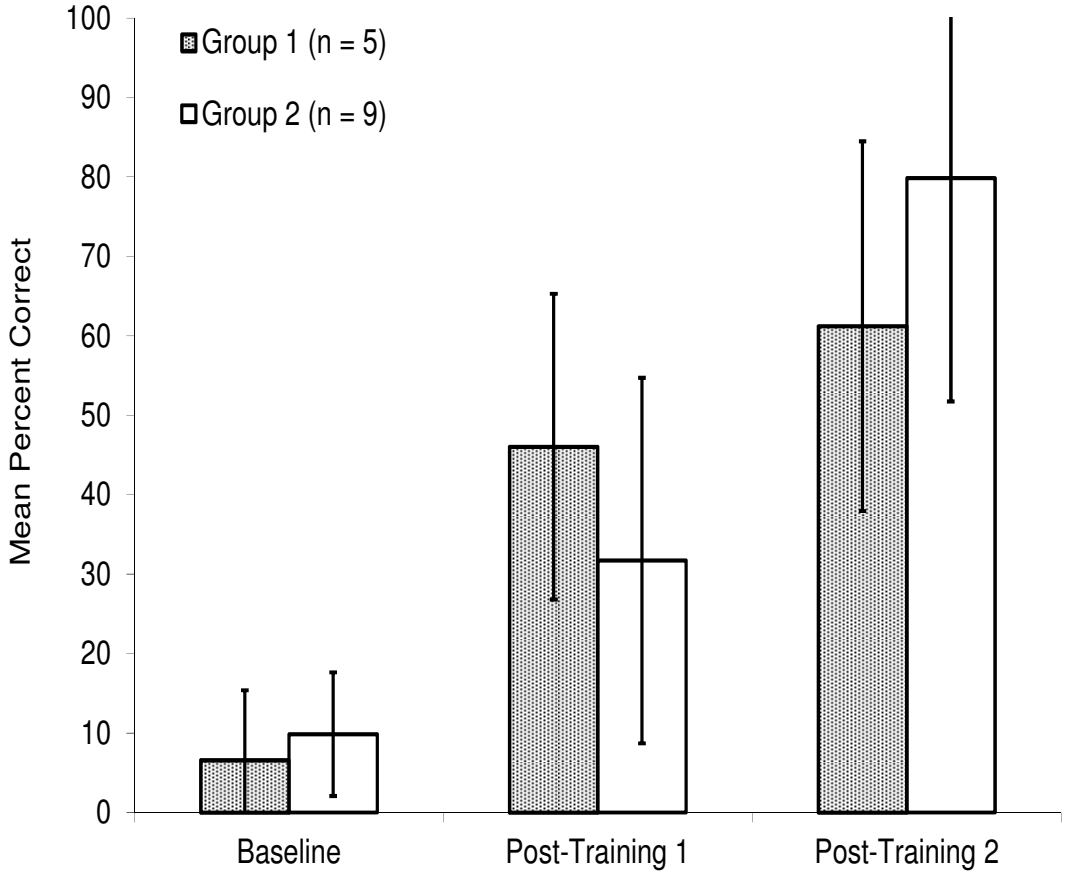

Figure 4. Mean percent correct during simulated preference assessments at baseline, post-training 1, and post-training 2 for Groups 1 (self-instructional manual) and 2 (method description). Error bars represent +/– one standard deviation.

Group 1 from baseline to post-training 2 was 54.66% and statistically significant, only one (Participant 17) of the five participants met mastery criterion. Group 2 received training using the method description followed by training using the self-instruction manual. Mean performance accuracy for Group 2 improved significantly after the method description at post-training 1. There was a further significant mean increase of 48.14% after exposure to the self-instructional manual at post-training 2. The overall mean increase for Group 2 from baseline to post-training 2 was 70.00% and statistically significant (Table 6). Moreover, seven of the nine participants met the mastery criterion at post-training 2.

In summary, for participants who did not meet mastery criterion after studying the self-instructional manual, exposure to the method description resulted in only a small improvement in performance accuracy and in the number of participants who met the mastery criterion. However, for participants who did not meet the mastery criterion after studying the method description, exposure to the self-instructional manual resulted in substantial improvement in both accuracy and in the number of participants who met the mastery criterion.

Social Validity

The results of the social validity questionnaire are summarized in Table 7 for 14 participants who had received both training procedures regardless of order. First, the mean rating for Item 1 - the importance of the goal of the intervention, was similar after each intervention (manual $M = 4.6$; method $M = 4.4$). That is, participants rated the goal of the intervention as quite important regardless of the intervention procedure. Second, on average, participants rated the manual easier to understand (Item 2, $M = 4.8$ vs. 3.6) and

Table 7

Social Validity Questionnaire Results for 14 Participants Who Received Both Training Procedures Regardless of Training Order

Questionnaire Items [a]	**Post-Manual**	**Post-Method**
1. It is important for staff working with individuals with developmental disabilities to learn to conduct preference assessments.	*M* = 4.6 (range, 3 to 5)	*M* = 4.4 (range, 3 to 5)
2. The written material was easy to follow and understand	*M* = 4.8 (range, 4 to 5)	*M* = 3.6 (range, 3 to 5)
3. The written material provided all the necessary information for me to do the assessment.	*M* = 4.6 (range, 4 to 5)	*M* = 3.6 (range, 2 to 5)
4. I believe I have successfully learned how to conduct the MSWO preference assessment from studying the materials.	*M* = 4.4 (range, 4 to 5)	*M* = 3.4 (range, 2 to 5)
5. I feel confident and ready to conduct MSWO preference assessment with clients after studying the materials.	*M* = 4.0 (range, 3 to 5)	*M* = 3.1 (range, 1 to 5)
6. I will likely use this assessment with clients in my job.	*M* = 3.3[b] (range, 1 to 5)	*M* = 3.3[b] (range, 1 to 5)

[a] Each item was rated from 1 (strongly disagree) to 5 (strongly agree).

[b] All 14 participants responded to items 1 through 5 and 12 participants responded to item 6.

provided the necessary information for conducting the assessment (Item 3, *M* = 4.6 vs. 3.6). Third, on average, participants perceived the manual to be more effective than the method description in teaching them to conduct the procedure (Item 4, *M* = 4.4 vs. 3.4) and they felt more confident about their ability to conduct the assessment with clients after studying the manual (Item 5, *M* = 4.0 vs. 3.1). Lastly, the average ratings for how likely participants would incorporate preference assessments into their work with clients were similar for both interventions (Item 6, manual *M* = 3.3; method *M* = 3.3).

Intervention Time

The amount of time each participant spent to study the manual and method description and to observe the model is reported in Table 8. For Participants 1 through 10 (first wave), the mean study time was 53.4 min. for the self-instructional manual and 9.7 min. for the method description. The mean amount of time participants spent to observe the model was 8.5 min. For Participants 11 through 18 (second wave), the mean study time was 66.1 min. for the self-instructional manual and 12.1 min. for the method description. Thus, on average, the second wave took slightly longer than the first wave to study for both interventions but the mean difference was not statistically significant ($p > .05$). The mean amount of time participants spent to observe the model was 8.4 min.

Discussion

This study yielded several findings. First, the MSWO self-instructional manual was more effective than the method description in improving performance accuracy and in producing mastery performance. This finding was supported by the effects observed in

Table 8

Time (minutes) Spent by Student Participants to complete the Method, Manual, and Modeling Phases

Participant	Manual	Method	Modeling
First Wave			
1	45	-	-
2	55	14	-
3	49	-	-
4	70	5	-
5	75	13	10
6	19	5	-
7	44	-	-
8	38	6	-
9	95	4	7
10	44	21	
Mean	53.4	9.71	8.5
Second Wave			
11	53	5	7
12	100	25	10
13	69	-	-
14	60	7	-
15	57	6	7
16	75	6	-
17	35	25	
18	80	11	10
Mean	66.1	12.1	8.4

the modified multiple-baseline design across participants where three of four participants achieved mastery after studying the manual alone and none of the four participants met mastery after studying the method description alone (Figure 1). This finding was also supported by the group comparisons, which showed statistically significant main effects of time and interaction between groups at post-training 1 and post-training 2 (Tables 5 and 6). Seven of the nine participants (77%) achieved mastery after receiving the method description and then the manual training, but only one of the five participants (20%) achieved mastery after receiving the manual and then the method description training. In other words, for those who did not achieve mastery after studying the method description, exposure to the self-instructional manual resulted in very substantial gains; whereas for those who did not achieve mastery after studying the manual, exposure to the method description resulted in little improvement.

Another finding of this study was that modeling appeared effective for participants who did not achieve mastery after studying both the manual and method description. All six participants who were exposed to modeling achieved mastery afterwards. However, since modeling was preceded by both method and manual interventions, it was not clear whether modeling alone would have been sufficient to produce mastery performance. Future research is needed to examine the contribution of modeling.

A third finding of this study was that participants who achieved mastery showed very strong retention and generalization performance. For example, Participant 2 maintained her performance above the mastery criterion 64 days after training. In all, 14 out of the 16 participants who completed a retention/generalization assessment with an

actor performed at or above the mastery criterion, and all seven participants who completed an assessment with a real client exceeded the mastery criterion.

Lastly, the social validity results showed that the mean ratings for the importance of the goal of training were similar after each intervention. However, the mean ratings were higher for the manual than for the method description on being easier to follow and understand, and that it provided the necessary information for the assessment (Items 2 through 5, Table 7). Considering that both groups rated Item 1 comparably and Items 2 through 5 differently, it lends support to the sensitivity of the questionnaire (i.e., participants should not be expected to show a difference on Item 1 regardless of training procedures). The mean ratings for item 6 were lowest among the items and similar for both groups. This was not surprising given that participants were undergraduate university students who had not and at the time of the study were not working with people with developmental disabilities. Future research should consider modifying the questionnaire by removing item 6 for participants not working with this population.

The present research extends previous research in several ways. First, this study joins Graff and Karsten's (2012) study by demonstrating the potential for using a self-instructional approach to teach preference assessment procedures. Second, the present study examined the effects of the self-instructional manual without it being preceded by other interventions. Lastly, the present study showed that the self-instructional approach was feasible with undergraduate university students whereas participants in Graff and Karsten's study were teachers with bachelor’s or master’s degrees.

The current study also contributes to the limited data on skill maintenance and generalization in previous research. Of all the training studies reviewed, only one (Bishop

& Kenzer, 2012) directly evaluated maintenance of skills learned over time. Overall, participants in the present study demonstrated a high level of retention and generalization after achieving the mastery criterion during both simulated assessments and assessments with real clients.

Several limitations of the study should be noted. First, the face validity of the behaviours for MSWO preference assessment described in the manual is limited because it was evaluated by only three graduate students. However, the assessment results obtained by different participants for the same client showed moderately high agreement on the most preferred items. This might be seen as a functional measure of the validity of the target behaviours and the mastery criterion used in this study. That is, participants who mastered the skills identified the same item as the most preferred for the same client. Future research on teaching behavioural assessment procedures should consider adding functional evaluations to establish the validity of the target behaviours.

Another limitation of the present study is that the participants in the study consisted of only university students and who were mostly in their early to mid 20s in age. This may limit the generality of the findings to individuals with different educational backgrounds. Future replications with different groups of participants such as front-line staff and parents should be conducted.

Another limitation to the present study involved the difference in consequences for student participants in the two waves. Although all students received an honourarium of $45, those in the first wave received 2% of their course credit for participating in the study whereas students in second wave received 10%. The amount of course credit was a decision of the course instructor and was beyond the Experimenter's control. While we

might speculate that the second wave should have performed better because of the larger reinforcer, the results were the opposite – 80% of the participants in the first wave achieved mastery after being exposed to one or both interventions (without modeling), but only 50% of the participants in the second wave accomplished the same. Thus, the difference in course credit did not appear to have influenced the results in the anticipated direction.

The self-instructional manual evaluated in the present study taught only one preference assessment procedure (MSWO). Future research is needed to develop and evaluate self-instructional manuals for the paired-stimulus procedure. Although the MSWO procedure has been shown to be more efficient than the PS procedure, the latter can be useful for clients who may have difficulty responding to multiple-stimulus procedures because of a limited scanning repertoire.

Future research on the MSWO self-instructional manual should also consider some revisions based on the findings of the present study. A list of common errors that participants made during the study was compiled. Table 9 shows the mean percent correct across participants for each target response on the MSWO behaviour checklist after studying the manual. The three most common errors made were: determining the rank for each item, providing the instruction “look” when presenting each item singly to the client at the start of each trial, and answering the interpretation question correctly (i.e., which item based on rank would be considered the most powerful reward), respectively. Revising the relevant sections of the self-instructional manual to reduce these errors should hopefully further improve the manual's effectiveness.

Table 9

Mean Percent Correct across Participants (n = 18) for each Target Response after Studying the Manual

Target Response	Mean % Correct
Determine the rank for each item	44.4%
Provide instruction "look" when presenting each item singly to ensure client attends	54.8%
Answered interpretation question correctly	55.6%
If client selects item within 15 s allow access to item	64.0%
After 30 s Participant remove item from client	64.8%
Present correct items one at a time & in the correct sequence	68.2%
Remove all unselected items from table	68.5%
After 30 s Participant removed item from client	69.4%
Remove all unselected items from table	80.6%
Participant will remove all items if client doesn't respond	83.3%
Record client response on data sheet	83.3%
A new session will be initiated (if applicable)	83.3%
Allow access to item	84.0%
Participant blocks access to both items (if possible)	85.2%
Record client response on data sheet	87.0%
Represent instruction "Pick one"(following no response)	88.0%
All necessary information on data sheet (i.e., client name, date, etc.)	89.0%
Provide instruction to "Pick one"	90.7%

Future research should also examine the addition of video modeling to accompany the self-instructional manual. In the current study, six participants who did not meet the mastery criterion after studying both the method description and the manual did so after they observed the live demonstration. The live modeling provided in the current study attempted to simulate a video presentation in that participants could request to see the demonstration or a portion of it as many times as needed, but they were not allowed to ask the Experimenter any questions. Using video modeling would reduce the resources needed for live modeling.

Third, not surprisingly, manual training took longer to complete compared with method description and modeling. The mean study time was 59.1 min (n = 18) for the self-instructional manual, 10.9 min (n = 14) for the method description, and 8.5 min (n = 6) for observing the modeling. Although the longer study time may be justifiable given its effectiveness, future research should examine the possibility of streamlining the manual to reduce study time to further increase its practicality. If video demonstration is included as part of the self-instructional package, it may be possible to reduce the length of the manual while maintaining its effectiveness. In summary, the current study extended previous research on staff training by demonstrating that a self-instruction training manual could be an effective training tool for teaching undergraduate university students to conduct the MSWO procedure. The study also added to the limited data on generalization of learned skills to a real client, as well as on the maintenance of skills learned over time. Considering the importance of stimulus preference assessment in behavioural and educational programming, the need to facilitate uptake of this technology for a large number of staff and instructors, and the need to train new staff regularly in this

field, it is important that a training tool be effective and low cost. The self-instructional manual shows great promise in meeting both of those criteria.

References

Arnal, L., Fazzio, D., Martin, G. L., Yu, C. T., Keilback., L., & Starke, A. (2007). Instructing university students to conduct discrete-trials teaching with confederates simulating children with autism. *Developmental Disabilities Bulletin, 35,* 131-147.

Bishop, M. R., & Kenzer, A. L. (2012). Teaching behavioural therapists to conduct brief preference assessments during therapy sessions. *Research in Autism Spectrum Disorders, 6,* 450-457.

Cannella, H. I., O'Reilly, M. F. O., & Lancioni, G. E. (2005). Choice and preference assessment research with people with severe to profound developmental disabilities: a review of the literature. *Research in Developmental Disabilities, 26*, 1-5.

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33,* 353-357.

Cooper, L. J., Wacker, D. P, Thursby, D., & Plagmann, L. A. (1992). Analysis of the effects of task preferences, task demands, and adult attention on child behavior in outpatient and classroom settings. *Journal of Applied Behavior Analysis, 25,* 823-840.

DeLeon I. G., Frank M. A., Gregory M. K., & Allman M. J. (2009). On the correspondence between preference assessment outcomes and progressive-ration schedule assessments of stimulus value. *Journal of Applied Behavior Analysis. 42(2),* 729–733.

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.

DeWiele, L., Martin, G. L., & Garinger, J. (2000). Field testing a self-instructional manual for the ABLA test. *Journal on Developmental Disabilities, 7,* 93-108.

Fazzio, D., Martin, G. L., Arnal, L., & Yu, D. C. T. (2009). Instructing university students to conduct discrete-trials teaching with children with autism. *Research in Autism Spectrum Disorders, 3,* 57-66.

Fisher, W. W., Piazza, C. C., Bowman, L. G., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment to enhance reinforcer identification. *American Journal on Mental Retardation, 101,* 15-25.

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.

Graff, R. B., & Ciccone, F. J. (2002). A *post hoc* analysis of multiple-stimulus preference assessment results. *Behavioral Interventions, 17,* 85-92.

Graff, R. B., Gibson, L., & Galiatsatos, G. T. (2006). The impact of high- and low-preference stimuli on vocational and academic performances of youths with severe disabilities. *Journal of Applied Behavior Analysis, 39, 131-135.*

Graff, R. B., & Karsten, A. M. (2012). Evaluation of a self-instruction package for conducting stimulus preference assessments. *Journal of Applied Behavior Analysis, 45,* 69-82.

Green, C. W., Reid, H. D., Canipe, V. S., & Gardner, S. M. (1991). A comprehensive evaluation of reinforcer identification processes for persons with profound multiple handicaps. *Journal of Applied Behavior Analysis, 24,* 537-552.

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21,* 31-43.

Harding, J. W., Wacker, D. P., Berg, W. K., Cooper, L. J., & Asmus, J. (1999). An analysis of choice making in the assessment of young children with severe behavior problems. *Journal of Applied Behavior Analysis, 32,* 63-82.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities, 21,* 61-73.

Hagopian, L. P., Long, E. S., & Rush, K. S. (2004). Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28*, 668-677.

Hughes, C., Pitkin, S. E., & Lorden, S. W. (1998). Assessing preferences and choices of persons with severe and profound mental retardation. *Education and Training in Mental Retardation and Developmental Disabilities, 33(4),* 299-316.

Kodak, T., Fisher, W. W., Kelley, M. E., & Kisamore, A. (2009). Comparing preference assessments: Selection- versus duration-based preference assessment procedures. *Research in Developmental disabilities, 30,* 1068-1077.

Johnston, J. M., & Pennypacker, H. S. (2009). Strategies and tactics of behavioral research (3rd ed.). New York: Routledge.

Lavie, T., & Sturmey, P. (2002). Training staff to conduct a paired-stimulus preference assessment. *Journal of Applied Behavior Analysis, 35,* 209-211.

Lee, M. S. H., Yu, C. T., Martin, T. L., & Martin, G. L. (2010). On the relation between reinforcer efficacy and preference. *Journal of Applied Behavior Analysis, 43*, 95-100.

Logan K. R. & Gast, D. L. (2001). Conducting preference assessments and reinforce testing for individuals with profound multiple disabilities: Issues and procedures. *Exceptionality, 9(3), 123-134.*

Lohrmann-O'Rourke, S. & Browder, D. M. (1998). Empirically based methods to assess the preferences of individuals with severe disabilities. American Journal on Mental *Retardation, 193 (2), 146-161.*

Martin, G. L., & Pear, J. J. (2011). *Behavior modification: what it is and how to do it (9th edition)*. Upper Saddle River, NJ: Prentice-Hall.

Martin, T. L., Yu, C. T., Martin, G. L., & Fazzio, D. (2006). On choice, preference, and preference for choice. *The Behavior Analyst Today, 7,* 234-241.

Machalicek, W., O'Reilly, M., Chan, J. M., Rispoli, M., Lang, R., Davis, T., Shogren, K., Sorrells, A., Lancioni, G., Sigafoos, J., Green, V., & Langthorne, P. (2009). Using videoconferencing to support teachers to conduct preference assessments with students with autism and developmental disabilities. *Research in Autism Spectrum Disorders, 3,* 32-41.

Ortiz, K. R. & Carr, J. E. (2000). Multiple-stimulus preference assessments: A comparison of free-operant and restricted-operant formats. *Behavioral Interventions, 15,* 345-353.

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18*, 249-255.

Parsons, M. B., & Reid, D. H. (1990). Assessing food preferences among persons with profound mental retardation: Providing opportunities to make choices. *Journal of Applied Behavior Analysis, 23,* 183-195.

Peck, S. M., Wacker, D. P., Berg, W. K., Cooper, L. J., Brown, R. A., Richman, D., McComas, J. J., Frischmeyer, P., & Millard, T. (1996). Choice-making treatment of young children's severe behavior problems. *Journal of Applied Behavior Analysis, 29*, 263-290.

Pence, S. T., St. Peter, C. C., & Tetreault, A. S. (2012). Increasing accurate preference assessment implementation through pyramidal training. *Journal of Applied Behavior Analysis, 45,* 349-359.

Piazza, C. C., Fisher, W. W., Hagopian, L. P., Bowman, L. G., & Toole, L. (1996). Using a choice assessment to predict reinforcer effectiveness. *Journal of Applied Behavior Analysis, 29*, 1-9.

Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of a brief stimulus reference assessment. *Journal of Applied Behavior Analysis, 31,* 605-620.

Ramon D. & Yu C. T. (2010). *How to conduct direct preference assessments for persons with developmental disabilities using a multiple-stimulus without replacement procedure: A self-instruction manual.* Unpublished manuscript, University of Manitoba and St.Amant Research Centre, Canada.

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.

Steege, M.W., Wacker, D. P., Berg, W. K., Cigrand, K. K., & Cooper, L. J. (1989). The use of behavioral assessment to prescribe and evaluate treatments for severely handicapped children. *Journal of Applied Behavior Analysis, 22*, 23-33.

Toutenburg, H. (2002). *Statistical analysis of designed experiments* (2nd ed.). Springer Verlag.

Tullis, C. A., Cannella-Malone, H. I., Basbigill, A. R., Yeager, A., Fleming, C. V., Payne, D., & Wu, P. F. (2011). Review of the choice and preference assessment literature for individuals with severe to profound disabilities. *Education and Training in Autism and Developmental Disabilities, 46(4)*, 576-595.

Wacker, D. P., Berg, W. K., Wiggins, B., Muldoon, M., & Cavanaugh, J. (1985). Evaluation of reinforcer preferences for profoundly handicapped students. *Journal of Applied Behavior Analysis,* 18, 173-179.

Wacker, D. P., Wiggins, B., Fowler, M., & Berg, W. K. (1998). Training students with profound or multiple handicaps to make requests via microswitches. *Journal of Applied Behavior Analysis, 21,* 331-343.

Windsor, J., Piché, L. M., & Lock, P. A. (1994). Preference tests: A comparison of two presentation methods. *Research in Developmental Disabilities, 15*, 439-455.

Yu, D., Martin, G. L., Hardy, L., Leader, C., & Quinn, G. (1985). Developing a behavioral assessment system for the mentally handicapped: A behavioral approach. *Canadian Journal for Exceptional Children, 1,* 117-123.

Appendix A

Multiple-Stimulus without Replacement Preference Assessment Baseline Data Sheet

Client Name/Code:_____ Tester:_____
Date: _____

Item	List items below	Item	List items below
A		D	
B		E	
C		F	

For each trial, write the letter (A through F) of the item chosen.

Session 1		Session 2		Session 3	
Trial	Item Chosen (e.g., B or Ø if none)	Trial	Item Chosen (e.g., B or Ø if none)	Trial	Item Chosen (e.g., B or Ø if none)
1		1		1	
2		2		2	
3		3		3	
4		4		4	
5		5		5	

List the six items from the most to least preferred in the space provided below:

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____

Question: Based on the results of the rank for each item, which of the six items is likely to be the most powerful reward? _____

Appendix B

A Sample Script for the Actor's Responses During a Simulated Assessment

Items: Playdoh®, squishy toy, cars, wooden blocks, Lego®, and book

Sampling and/or When Given an Item: Engage in the activity until the participant takes it away from you.

SESSION 1	First Response	Second Response
Trial 1	Touch only the book	
Trial 2	No response	Touch the Lego®
Trial 3	Simultaneously touch more than one toy	Touch only the Playdoh®
Trial 4	Touch only the squishy toy	
Trial 5	Touch the wooden blocks	

SESSION 2	First Response	Second Response
Trial 1	Touch the Playdoh® with one hand, wait 1 sec, then touch the book with your other hand	
Trial 2	Touch only the book	
Trial 3	Simultaneously touch more than one toy	Touch only the Lego®
Trial 4	Do not respond	Touch the purple squishy toy
Trial 5	Touch the car first, wait 1 sec and then touch the wooden blocks	

SESSION 3	First Response	Second Response
Trial 1	Touch the Lego® first, wait 1 sec and then touch the squishy toy	
Trial 2	Touch the squishy toy	
Trial 3	Touch the toy cars	
Trial 4	Simultaneously touch more than one toy	Touch only the Playdoh®
Trial 5	No response	No response

Appendix C

Checklist and Data Sheet for Scoring Participant's Performance

Multiple-Stimulus without Replacement Data Sheet for Tracking Participant's Target Behaviours

Date: **Participant Name:** **Observer:**
Training/Phase: **Confederate:**

Please put a "✓ " for correct responses, an X" for incorrect responses, and a n/a for not applicable for each response.

Preparation before an Assessment	**1**	**2**	**3**	**4**	**5**	**6**
1. Readying choices to be assessed (food only)						
2. Allow client to interact with item for 30 sec or consume item						

Presenting MSWO Trials *Antecedent Behaviours*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Provide instruction "look" when presenting each item singly to ensure client attends															
Present correct items one at a time & in the correct sequence															
Provide instruction to "pick one"															

Consequence 1. Appropriate Selection Responses Selects 1 item OR Selects 2 Items Sequentially OR Rejects 1 & Selects another	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
If client selects item within 15 s allow access to item															
Remove all unselected items from table															
After 30 s Participant removed item from client															
Record client response on datasheet															

Consequence 2. Selects two items simultaneously	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Participant blocks access to both items (if possible)															
Represent instruction "pick one"															
If client selects item within 15 s allow access to item															
Remove all unselected items from table															
After 30 s Participant remove item from client															
Record client response on datasheet															

Consequence 3. No response	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Participant repeats instruction "pick one"															
Scenario 3.1. Client does not response to second "pick one"															
Participant will remove all items															
Record client response on datasheet															
A new session will be initiated (if applicable)															
Scenario 3.2 Client selects item within 15 s															
Allow access to item															
Remove all unselected items from table															
After 30 s Participant removed item from client															
Record client response on datasheet															

Summarizing and Interpreting the Results	**Yes / No**
Calculate the rank for each item correctly	
Answered interpretation question correctly (identify most effective reward)	
All necessary information on data sheet (i.e., client and Participant name, date, etc.	

Appendix D

Method Description of the MSWO Procedure[1]

Prior to the beginning of the assessment, clients are given a sample of each of the edible items and are given 30 s access to each of the leisure items.

For this assessment procedure, each session begins with all items sequenced randomly in a straight line on the table, about 5 cm apart. While a client is seated at the table approximately .3 m from the stimulus array, the staff will instruct the client to select one item. After a selection is made, the item is either removed from the immediate area (leisure item) or it is not replaced (food item). Prior to the next trial, the sequencing of the remaining items is rotated by taking the item at the left end of the line and moving it to the right end, then shifting the other items so that they are again equally spaced on the table. The second trial then follows immediately. This procedure continues until all items are selected or until a client makes no selection within 15 s from the beginning of a trial. In the latter case, the session ends and all remaining items are recorded as not selected. The staff will complete two additional sessions, as described above, using the same six items.

---

[1] This description was adapted from the Method sections from the published papers by DeLeon and Iwata (1996) and Roscoe et al. (2006). Adaptations included describing the procedure in present tense, modifying the number of items from 7 to 6, reducing the number of sessions from 5 to 3, replacing the term “participant” with “client”, reducing the duration that clients could access reinforcing items from 30 seconds to 15 seconds, and replacing the term “experimenter” with “staff”.

A selection response is recorded when a client makes physical contact with one of the presented items. The client has 15 s to select an item. If the client makes contact with more than one item, the first item contacted will be recorded as the selection. If no item is selected within the 15 s period, the staff will repeat the instruction “pick one.” If no item is selected within the 15 s period following the repeated instruction, the trial and session will end. When a selection is made, the trial ends after the client receives 30 s access to the item (leisure stimuli) or after the client has complete consumed the item (edible stimuli).

If the client selects two items in a sequential fashion, the client will be permitted access to the item selected first for 30 s. After 30 s, the item will be removed from the array. If the client selected two items simultaneously, the client will be prevented from obtaining either of the items, and the staff will reinitiate the trial by rotating the items and saying “pick one.” If the client simultaneously selects two items a second consecutive time, all stimuli will be removed, and a new session (i.e., all six items will be presented) will be initiated. Staff will record client responses for each trial by indicating the item chosen using its corresponding item letter on the data sheet. The staff will summarize client data by adding the total number of selections for each item and dividing that number by the total number of presentations for that item and multiplying by 100% to yield a percentage selection measure for each item in the array.

Appendix E

Multiple-Stimulus without Replacement Preference Assessment Method Data Sheet

Client Name/Code:_____ Tester:_____
Date: _____

Item	List items below	Item	List items below
A		D	
B		E	
C		F	

For each trial, write the letter (A through F) of the item chosen.

Session 1		Session 2		Session 3	
Trial	Item Chosen (e.g., B or Ø if none)	Trial	Item Chosen (e.g., B or Ø if none)	Trial	Item Chosen (e.g., B or Ø if none)
1		1		1	
2		2		2	
3		3		3	
4		4		4	
5		5		5	

Determine the percentage selection for each item by adding the total number of selections for each item and divide that number by the total number of presentations for that item and multiply by 100%.

Item	% selection for each item
A	
B	
C	
D	
E	
F	

Question: Based on results of the rank for each item, which of the six items is likely to be the most powerful reward? _____

Appendix F

[2] Table of contents from Ramon and Yu (2010).

Appendix G

25-Item Behaviour Checklist for MSWO Assessments (from Ramon & Yu, 2010)

Setting
- Do your assessment in a quiet place.
- Have a table and chairs, and a place for your data sheet and materials out of reach of the client.

Materials
- Gather the six items you want to test.
- If you're testing food items, have at least four bite-size pieces of each item.
- Fill in the data sheet: client's name, your name, date, and items.
- Bring the client in and sit at the table facing each other.

Sampling
- Present one item at a time in front of the client, ask him/her to "look" at the item, and ask him/her to take it.
- Let the client eat the food item or interact with the non-food item for up to 30 seconds, remove the item, and present the next item.
- Repeat until each item has been sampled once before you start presenting trials.

Trial Presentation
- Present all items in a row in front of client on the ***first*** trial of each session.
- Do not include the chosen item on subsequent trials of the session.
- On each subsequent trial, shift all the items one position to the left (or right).
- On each trial, ask the client to "look" at each item before putting it on the table, and then say "Pick one".
- Give the client up to 15 seconds to choose.

After a Selection Response
- Say "Good"
- Give the selected item to the client to eat or interact with for 30 seconds
- Remove other items from the table.
- Retrieve the item after 30 seconds.
- Record the selected item on the data sheet and present the next trial.

After No Response
- If the client did not choose an item after 15 seconds, repeat the request "Pick one" once.
- If the client still does not select an item after 15 seconds, the trial ends.
- Say nothing, remove all items, record a zero on the data sheet and end the session.

After Client Tries to Take More than 1 Item
- Block gently and retrieve items if necessary (saying "no" is optional).
- Remove all items from table.
- Repeat the same trial.

Ranking Items
- Transfer the ranks from each session to table at the bottom of the data sheet.
- Add the rankings for each item across the three sessions and divide by 3 to obtain an average rank for each item.

Appendix H

Rating Form for Face Validity Evaluation

Date: _____

Graduate Student: _____

Instructions: Please use the 7-point scale to rate the following items regarding IMPORTANCE when conducting a multiple-stimulus without replacement preference assessment.

Setting

1. Do your assessment in a quiet place.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

2. Have a table and chairs, and a place for your data sheet and materials out of reach of the client.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

Materials

3. Gather the 6 items you want to test.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

4. If you're testing food items, have at least 4 bite size pieces of each item.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

5. Fill in the dat asheet: client's name, your name, date, and items.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

6. Bring the client in and sit at the table facing each other.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

Sampling

7. Present one item at a time in front of the client, ask him/her to look at the item, and ask him/her to take it.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

8. Let the client eat the food item or interact with the non-food item for 15-30 seconds.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

9. Repeat until each item has been sampled once.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

Trial Presentation

10. Present all items in a row in front of client on the ***first*** trial of each session.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

11. Do not include the chosen item on subsequent trials of the session.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

12. On each subsequent trial, shift all the items one position to the left (or right).

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

13. On each trial, ask the client to look at each item before putting it on the table, and then say “pick one”.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

14. Give the client up to 15 seconds to choose.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

After a Selection Response

15. Say “Good”

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

16. Give the selected item to the client to eat or play with for 15-30 seconds and remove other items from the table.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

17. Record the selected item on the data sheet and present the next trial.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

After No Response

18. If the client did not choose an item after 15 seconds, repeat the request “pick one” once.

1= not important, 4= important, 7= very important / essential

1	2	3	4	5	6	7

19. If the client still does not select an item after 15 seconds, the trial ends.

1= not important, 4= important, 7= very important / essential

1	2	3	4	5	6	7

20. Say nothing, remove all items, record a zero on the data sheet and end the session.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

After Client Tries to Take More than 1 Item

21. Block gently and retrieve items if necessary.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

22. Remove all items from table.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

23. Repeat the same trial.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

Ranking Items

24. Transfer the ranks from each session to table at the bottom of the data sheet.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

25. Add the three rankings in each row and divide by three to obtain an average rank for each item.

1 = not important, 4 = important, 7 = very important / essential

1	2	3	4	5	6	7

Are there any important items that are missing from the Behaviour Checklist for the MSWO Assessment?

Additional Comments & Suggestions

Appendix I

Multiple-Stimulus without Replacement Preference Assessment Manual Data Sheet

Client Name/Code:_____ Tester:_____
Date: _____

Item	List items below	Item	List items below
A		D	
B		E	
C		F	

For each trial, write the letter (A through F) of the item chosen.

Session 1			Session 2			Session 3		
Trial	Item Chosen **(e.g., B or Ø if none)**	Rank	Trial	Item Chosen **(e.g., B or Ø if none)**	Rank	Trial	Item Chosen **(e.g., B or Ø if none)**	Rank
1		1	1		1	1		1
2		2	2		2	2		2
3		3	3		3	3		3
4		4	4		4	4		4
5		5	5		5	5		5

Transfer the "rank" of items from each session above to the table below. Average the rank across the 3 sessions and enter the average in the last column. The item with the highest average rank (smallest number) is the most preferred item.

Item	Rank from Each Session			Average Rank
	Session 1	Session 2	Session 3	
A				
B				
C				
D				
E				
F				

Question: Based on the results of the rank for each item, which of the six items is likely to be the most powerful reward? _____

Appendix J

Social Validity Questionnaire

Name: _____ Date:_____

Please indicate how strongly you agree or disagree with each statement by circling the number after each statement.

1. It is important for staff working with individuals with developmental disabilities to learn to conduct preference assessments.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree/Disagree	Agree	Strongly Agree

2. The written material was easy to follow and understand.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree/Disagree	Agree	Strongly Agree

3. The written material provided all the necessary information for me to do the assessment.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree/Disagree	Agree	Strongly Agree

4. I believe I have successfully learned how to conduct the MSWO preference assessment from studying the materials.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree/Disagree	Agree	Strongly Agree

5. I feel confident and ready to conduct MSWO preference assessment with clients after studying the materials.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree/Disagree	Agree	Strongly Agree

6. I will likely use this assessment with clients in my job.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree/Disagree	Agree	Strongly Agree

Other Comments:_____