Teaching Individuals to Conduct a Preference Assessment Procedure using Computer-Aided Personalized System of Instruction

by

Lindsay Arnal

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

Acknowledgements

I extend my gratitude to my supervisor, Dr. C.T. Yu, my committee members, Drs. Garry Martin, Dennis Hrycaiko, and Joseph Pear and my research assistants, Carly Chand, Lilian Saltel, and Katrina Lavoie for their support during this study. I thank my family for their encouragement and support throughout my graduate program. Lastly, this research was supported in part by a doctoral fellowship from Social Sciences and Humanities Research Council of Canada.

Dedication

To my Parents, who have always instilled in me the value of education,

To Jason, for his unwavering support throughout my years of study,

To Dr. Yu, for his mentorship and for inspiring me to be passionate about behaviour analysis,

and

In memory of Dr. E.A. Kaprowy, who encouraged me to pursue graduate studies in the field of Psychology.

Abstract

Preference assessments are an evidence-based procedures used to identify potential reinforcers for persons with developmental disabilities. There is a need to develop effective and efficient procedures to teach students and staff to conduct preference assessments, but only a small number of studies have been conducted and only two have used self-instructional materials. A recent study by Ramon et al. (2012) found that a self-instructional manual was more effective than a method description extracted from published articles for teaching university students to conduct multiple-stimulus without replacement preference assessments for persons with developmental disabilities. The present study extended this research by (a) adapting the self-instructional manual from Ramon et al. for online delivery, (b) adding video modeling as a teaching component, and (c) delivering the training package using a modified computer-aided personalized system of instruction (CAPSI, Pear and Kinsner, 1988). The training package was evaluated using a multiple-baseline design across three university students, replicated across three more students; and a multiple-baseline design across a pair of staff members, replicated a across a second pair. During the baseline phase, participants studied a two-page written description of the assessment procedure adapted from published studies. During the self-instructional manual phase, participants completed all of the following online: studied the self-instructional manual presented in eight units, viewed video demonstrations of the procedure, and completed review exercises scored by the computer program to demonstrate mastery of each study unit. Performance accuracy of each participant was scored using a standard behaviour checklist during a simulated preference assessment conducted following each phase. Clear and immediate improvement in performance

accuracy was observed in all participants immediately following the self-instructional training package. Overall, students improved from a mean of 35% correct in baseline to a mean of 94% correct following CAPSI and staff improved from a mean of 23% correct in baseline to a mean of 87% correct following CAPSI. During retention and generalization assessments conducted from 7 to 17 days following self-instructional training, five of the six students and one of the four staff members performed at or above 85% correct (the mastery criterion). The findings showed that online delivery of the self-instructional manual plus video modeling has tremendous potential for providing an effective method for teaching a preference assessment procedure without face-to-face instruction.

Table of Contents

List of Tables

List of Figures

Teaching Individuals to Conduct a Preference Assessment Procedure using Computer-Aided Personalized System of Instruction

Introduction

Using positive reinforcement to strengthen behaviours is an important component in virtually all teaching programs for individuals with developmental disabilities, and direct preference assessment is an effective and efficient method for identifying reinforcers for individuals with severe and profound developmental disabilities (Tullis et al., 2011). Some recent research has focused on developing and evaluating effective and efficient methods for teaching individuals to conduct a preference assessment in order to promote its application by practitioners. Ramon and Yu (2010) developed a self-instructional manual for conducting a preference assessment procedure, called multiple-stimulus without replacement, and their evaluation showed that the manual was moderately effective in teaching university students to carry out the procedure and highly effective when it was combined with modeling (Ramon et al., 2012). Given the importance of preference assessments in behavioural programming and interventions, the need to develop effective training strategies is essential. The current study sought to improve the effectiveness of the above self-instructional manual by adding video modeling and delivering the training package online using a modified computer-aided personalized system of instruction (CAPSI, Pear & Kinsner, 1988).

Concept and Characteristics of Preference

Preference may be described as “the relative strength of discriminated operants” (Catania, 1998). We can measure a person’s preference based on their frequency of selection among choices through direct preference assessment (Martin, Yu, Martin, &

Fazzio, 2006). The most preferred stimulus identified through direct preference assessment has been shown to be a reinforcer, which can be used to strengthen a behaviour in a teaching situation (Hagopian, Long, & Rush, 2004; Tullis et al., 2011). Moreover, reinforcer efficacy appears to be positively correlated with preference (Lee, Yu, Martin, & Martin, 2010).

Research has shown that preferences may vary over time and across individuals (Ciccone, Graff, & Ahearn, 2007). Hanley, Iwata, and Roscoe (2006) demonstrated that preference instability is not necessarily related to the age or disability of the individual but may be the result of naturally occurring changes in establishing operations and conditioning histories. Preference for a particular stimulus may be impacted by the type of reinforcers (e.g., food versus leisure items) that are presented concurrently in an assessment (Bojak & Carr, 1999) and the reinforcing value of a stimulus may be altered by the array of high and low preference stimuli with which it is presented (Daly et al., 2009; Roscoe, Iwata, & Kahng, 1999). Overall, preferences have been shown to be fairly stable for individuals with developmental disabilities (Carr, Nicolson, & Higbee, 2000).

Both direct and indirect methods have been used to assess preferences. A common indirect method involves interviewing the individual or caregiver to determine a set of preferred stimuli. While reinforcer surveys and interviews are beneficial to identify an initial group of stimuli to be assessed, this method is inferior to direct methods in that it does not generate preference hierarchies (Fisher, Piazza, Bowman, & Amari, 1996; Cote, Thompson, Hanley, & McKerchar, 2007). In addition, research that compared direct and indirect methods of assessing preferences have shown that direct assessment is superior in indentifying stimuli that function as reinforcers (Green et al., 1988; Fisher et

al.). Further, direct measures that utilize tangible over pictorial stimuli have been shown to be superior in identifying potent reinforcers (Higbee, Carr, & Harrison, 1999). Therefore, the following sections will focus on direct preference assessment measures.

Direct Preference Assessment Methods

Several direct preference assessment procedures have been evaluated. The single-stimulus (SS) procedure involves presenting stimuli one at a time to an individual and measuring the percentage of trials the person approaches or interacts with each stimulus (Pace, Ivancic, Edwards, Iwata, & Page, 1985). Pace et al. were able to identify preferred stimuli for all six clients in their study using the SS procedure. Moreover, delivering the identified preferred stimulus contingent on a target response increased the frequency of that response.

However, some studies have found that a paired-stimulus (PS) procedure was more effective than the SS procedure. The PS procedure involves presenting two stimuli concurrently on each trial, and each stimulus is paired with every other stimulus (Fisher et al., 1992). Preference for a stimulus is defined as the percentage of trials it is selected based on the number of trials it is available. Fisher et al. found that the PS procedure generated a more distinct preference hierarchy among the stimuli than the SS procedure. Other studies have also reported findings indicating that the PS procedure is superior to the SS procedure (DeLeon & Iwata, 1996; Piazza, Fisher, Hagopian, Bowman, & Toole, 1996). Moreover, Piazza et al. tested the reinforcing effectiveness of top-ranked, middle-ranked and low-ranked stimuli with four individuals with developmental disabilities. The results revealed that the top-ranked stimuli functioned as reinforcers for all four participants whereas none of the low ranked stimuli functioned as reinforcers for any of the participants, thus highlighting the ability of the PS procedure to identify reinforcers.

A third procedure that has been examined is the multiple-stimulus with replacement (MSW) procedure (Windsor, Piché, & Locke, 1994). In MSW, several stimuli (usually six to eight) are presented concurrently on every trial (i.e., a selected item is replaced on subsequent trials). Preference for a stimulus is operationally defined as the percentage of trials it is selected as in the PS procedure. Windsor et al. found that the MSW procedure was superior to the PS procedure in efficiency, requiring less time to administer, but the MSW procedure yielded a less differentiated preference hierarchy because participants tended to select the same item (i.e., most preferred item) across trials.

DeLeon and Iwata (1996) modified the MSW procedure by not replacing the selected item on each trial. Thus, after a stimulus has been selected on a trial, it is removed from the array and only the remaining stimuli are presented on the next trial. This process is repeated until the last two stimuli have been presented. Preference for a stimulus in the multiple-stimulus without replacement (MSWO) procedure can be operationalized as the percentage of trials it is selected or the order in which it and the other stimuli are selected (i.e., the stimulus selected on the first trial is ranked one, the stimulus selected on the second trial is ranked two, and so on). DeLeon and Iwata compared the PS, MSW, and the MSWO procedures. Although the results indicated that all three procedures were equally effective in identifying the most preferred stimuli, MSWO was more efficient than PS in that it required half the time to administer. In addition, both the PS and MSWO procedures consistently identified the same top three most preferred items and produced more stable stimulus hierarchies across sessions.

Several studies have focused on improving the efficiency of MSWO procedures. Higbee, Carr, and Harrison (2000) evaluated the effectiveness of an abbreviated MSWO preference assessment with nine adults with mental retardation. The abbreviated MSWO procedure was identical to that described in the DeLeon and Iwata (1996) study, but it involved only three sessions instead of five. While DeLeon and Iwata reported that the five-session MSWO procedure required an average of 21.8 minutes to administer, the abbreviated MSWO procedure required 15 minutes or less to administer. Reinforcer tests conducted after the abbreviated preference assessments revealed that the top two most preferred stimuli produced an increase in responding compared to baseline levels for seven out of the nine participants.

Carr et al. (2000) evaluated the effectiveness of an abbreviated MSWO procedure with three children with autistic disorder. Following the assessment, a reinforcer test was conducted for the most-preferred stimulus (ranked 1), moderately- preferred stimulus (ranked 4 or 5), and the least-preferred stimulus (ranked 8). The authors found that the least-preferred stimulus did not increase responding compared to baseline for any participants whereas the most-preferred stimulus produced the greatest increase in responding relative to baseline for all three participants. In addition, a positive correlation was found between the results obtained from the first session with the results obtained from all three sessions combined, suggesting that a single-session MSWO assessment may be sufficient. The average time required to administer the three-session MSWO procedure was approximately 5 minutes.

A limitation of the studies by Higbee et al. (2000) and Carr et al. (2000) is that neither compared the results of the abbreviated three-session MSWO assessment to an

assessment involving more sessions. Graff and Ciccone (2002) addressed this limitation by comparing an extended preference assessment, containing seven sessions comprised of seven trials each, to results obtained with fewer sessions and trials involving four different combinations (five sessions of seven trials each, three sessions of seven trials each, seven sessions of three trials each, and five sessions of three trials each). Each combination was compared with the extended assessment, and the same top ranked stimulus was identified during 82%, 70%, 93%, and 82% of the comparisons, respectively.

The direct preference assessment procedures discussed above may be referred to as trial-based or restricted-operant procedures in that the stimuli are only made available for brief-trial intervals and stimulus selection can occur only during a trial interval. A small number of studies have examined stimulus availability (e.g., duration of stimulus presentation) during preference assessments. Kodak, Fisher, Kelley, and Kisamore (2009) compared the MSW procedure to a variation of a free-operant procedure in which duration of stimulus presentation was manipulated. While both restricted- and free-operant procedures have proven to be effective in identifying highly preferred items (Ortiz & Carr, 2000), restricted-operant techniques have been found to be superior in identifying multiple reinforcers and preference hierarchies that include high and low preference items (DeLeon & Iwata, 1996; Kodak et al., 2009; Roane, Vollmer, Ringdahl, & Marcus 1998).

In summary, restricted-operant preference assessment procedures are more commonly used than free-operant procedures in the literature. Among the restricted-operant direct preference assessment procedures, PS and MSWO procedures are both

superior to the SS procedure in producing a more differentiated preference hierarchy, and the MSWO procedure requires less time to administer than the PS procedure. Lastly, abbreviated MSWO assessments using fewer than three sessions may be less reliable in identifying the most-preferred stimulus.

Teaching Individuals to Conduct Preference Assessments

Face-to-Face Instruction. To date, a handful of studies have focused on teaching individuals to conduct preference assessments. In a multiple-baseline design across three staff members, Lavie and Sturmey (2002) successfully taught staff to use the PS procedure to assess children with autism spectrum disorders. The training strategy included: a description of the target skills, a behaviour checklist, and a description of the items on the behaviour checklist, video modeling, observation, feedback, and repetition of the last three components as necessary. Accuracy was assessed using the behaviour checklist and mastery was defined as 85% correct or higher. All three participants demonstrated 100% or near 100% accuracy after approximately 80 minutes of training.

Roscoe, Fisher, Glover, and Volkert (2006) used a multi-element design to train four undergraduate students to implement both the PS and MSWO procedures with simulated clients (i.e., confederates role playing children with a disability) and real clients (i.e., children referred to an outpatient clinic for assessment). The authors were primarily interested in examining the impact of two variables to improve performance: feedback, and a monetary contingency. In the feedback condition, the experimenter reviewed a videotaped session of the participant conducting a preference assessment and provided feedback to the participant on correct and incorrect target behaviours. In the monetary contingency condition, $10 was provided contingent on correct performance

during the previous session. Results indicated that feedback was more effective in developing the necessary skills to conduct preference assessments whereas the monetary contingency appeared to have little impact on performance.

Roscoe and Fisher (2008) extended the work of Roscoe et al. (2006) in an effort to further evaluate the feedback component of training. In a multi-element design across two conditions, eight behavioural technicians were taught to implement the PS and MSWO preference assessments using feedback and role-play. Feedback was administered in a manner similar to that described by Roscoe et al. Each participant received training in each condition, yielding a total of 16 assessments. Results revealed that accuracy of responding increased from mean baseline scores of 45% for the PS assessment and 40% for the MSWO assessment to 80% for all participants, and in 14 out of 16 assessments accuracy increased to 90%.

Recently, Pence, St. Peter, and Tetreault (2012) extended the application of pyramidal training strategies to train special-education teachers enrolled in a course sequence to become Board Certified Behavior Analysts (BCBA) to conduct preference assessments. During Experiment 1, three special-education teachers taught six trainees to implement the PS, MSWO, and free-operant preference assessments using one-to-one instruction that included a written description of the procedures, live modeling, role-play, and feedback. All three trainers had previously learned to conduct the preference assessment procedures with at least 90% accuracy using detailed written descriptions and one-to-one instruction delivered by their BCBA supervisor. Following training, all six trainees were able to conduct the assessment procedures with a minimum of 90% accuracy during training and generalization sessions and treatment integrity remained

high in the natural setting (i.e., classroom). In Experiment 2, five trainees from Experiment 1 taught 18 preschool teachers to conduct the same three preference assessments using the identical strategy. Results were consistent with Experiment 1 in that all 18 teachers achieved the 90% mastery criterion or greater. The results lend support to the use of face-to-face instruction to teach behaviour assessment procedures.

Self-Instructional Training. Self-instructional manuals have been shown to be an effective and efficient method for teaching staff and caregivers to conduct behavioural assessments (DeWiele, Martin, & Garinger, 2000; Hu, Pear, & Yu, 2012; Yu, Martin, Hardy, Leader, & Quinn, 1985) and discrete-trials teaching with individuals with developmental disabilities (Arnal et al., 2007; Fazzio, Martin, Arnal, & Yu, 2009; Salem et al., 2009; Thiessen, Fazzio, Arnal, Martin, & Keilback, 2009; Thomson et al., 2012). These manuals incorporated pedagogical features, found in a personalized system of instruction (PSI, Keller, 1968), such as presenting materials in small units and requesting learners to master each unit before proceeding to the next unit.

Recently, Graff and Karsten (2012) evaluated the effects of a self-instructional package to teach 11 teachers to implement, score, and interpret the results of preference assessments. A multiple-baseline design across two assessment types (MSWO and PS) was conducted with each teacher. The order in which each assessment was taught was varied across teachers to control for sequence effects. In the baseline phase, participants received a written description for how to conduct the PS and MSWO assessments. The descriptions were extracted from the method sections of previous studies (DeLeon & Iwata, 1996; Fisher et al. 1992). The training package consisted of two components: (a) enhanced written instructions which included step-by-step instructions written in

nontechnical language and a detailed datasheet, and (b) the same written instructions from the baseline phase plus the datasheet from the enhanced written instructions. In order to isolate the effects of specific training components, participants received one of two training sequences (i.e., a-b or b-a). The designated mastery criterion for conducting each assessment was 90% or greater accuracy for two consecutive sessions. The mastery criterion for scoring and interpreting the preference assessment results was 100% accuracy when calculating selection percentages and an accurate verbal report of the most preferred item for one complete set of preference assessment data. Generalization probes with a child with a developmental disability took place one week after meeting the mastery criterion. Results showed that none of the teachers met mastery for either assessment using the written instructions alone. Performance accuracy in baseline was 34% for PS and 46% for MSWO. When the written instructions were supplemented with a datasheet, accuracy increased for four out of five participants for PS and three out of five for MSWO. Those participants who received the enhanced written instructions showed an increase in performance to a mean of 98% for PS and a mean of 99% for MSWO, and performance was maintained at the one-week generalization probe. The results demonstrated that individuals without prior training on conducting preference assessments did not accurately conduct the PS and MSWO assessments using written instructions alone. However, enhanced written instructions were sufficient to train staff to implement preference assessment procedures accurately. One limitation of the study noted by the authors was that the enhanced instructions were always preceded by the procedural description in baseline.

Ramon et al. (2012) compared a self-instructional manual (developed by Ramon and Yu, 2010) and a description of the MSWO procedures extracted from the method sections of published studies (e.g., DeLeon & Iwata, 1996; Roscoe et al., 2006) using an unbalanced crossover design. Eighteen university students were assigned randomly to one of two groups (n = 9 per group). One group received the method description first and followed by the self-instructional manual if necessary. Another group received the interventions in the reverse order. The design was unbalanced in that participants received the second intervention only if they did not meet a mastery criterion after the first intervention. Performance accuracy during simulated assessments conducted before and after the first intervention (self-instructional manual or method description) showed that the self-instructional manual group was significantly more accurate than the method description group. Moreover, four out of the nine participants in the manual group met the mastery criterion of 85% correct or higher, whereas no one in the method description group met the mastery criterion. After the second intervention (crossover), only one participant (out of five) in the manual-then-method group met the mastery criterion, whereas all nine participants in the method-then-manual group met the mastery criterion. Participants who did not meet mastery criterion after receiving both interventions did so after observing a live demonstration of the procedure. High performance accuracy was also observed during retention and generalization assessments with actors and with real clients with developmental disabilities. Although the self-instructional materials differed from Graff and Karsten's study (2012), Ramon et al. showed that self-instruction alone (without it being preceded by the method description) was more effective than the

method description from published articles. In addition, the effectiveness of self-instruction could be enhanced by modeling.

Statement of the Problem

The above literature suggests that: (a) preference assessments are evidence-supported procedures that can be used to identify reinforcers for people with developmental disabilities; (b) there is a need to promote the acquisition of these procedures because the identification of effective reinforcers is important to the success of educational and treatment programs; (c) there is limited research on teaching individuals to conduct preference assessments in general; and (d) although there is considerable research on using self-instructional materials to teach other procedures such as discrete-trials teaching and adaptive behaviour assessment, evaluation of self-instructional materials for teaching preference assessment procedures is limited. Research by Graff and Karsten (2012) and Ramon et al. (2012) showed that it is feasible to develop self-instructional materials to teach individuals to conduct a preference assessment.

Although most participants in Ramon et al.'s study (2012) met the mastery criterion, five of the 19 participants did not meet criterion even after studying both the method description and self-instructional manual. These participants, however, met the mastery criterion after observing a live demonstration. It is possible that some individuals may learn more effectively through observation than from textual materials. Therefore, the effectiveness of the self-instructional manual might be enhanced by adding a modeling component. Since the involvement of a live model would detract from the self-instructional approach and increase cost, video modeling could be used as an alternative.

Video modeling has been frequently used to teach skills to individuals with developmental disabilities. However, only a small number of studies have examined video modeling as a strategy to produce high treatment integrity when training individuals to implement behavioural interventions (Catania, Almeida, Liu-Constant, & Digennaro-Reed, 2009; Collins, Higbee, & Salzberg, 2009) and assessments (Hu et al., 2012; Moore & Fisher, 2007). For example, Digennaro-Reed, Codding, Catania, and Maguire (2010) studied the effects of video modeling and performance feedback to improve treatment integrity when implementing behavioural interventions by teachers. Among the studies on teaching preference assessments reviewed earlier, Lavie and Sturmey (2002) was the only one that included video modeling as a teaching component. In addition to its practicality and low cost, video modeling has the advantage of always presenting a consistent model compared to live modeling.

Promoting the acquisition of new knowledge using the self-instructional approach poses additional challenges. In the self-instructional manual by Ramon and Yu (2010) for instance, the effectiveness of the manual rests on the assumption that the learner reads the materials and completes the exercises correctly at the end of each unit. How can we ensure that this has been accomplished? Second, if the self-instructional package proved effective, how can we disseminate the materials quickly and widely? These challenges may be addressed using a computerized instructional program. With the proliferation of the Internet and online learning, research has shown that a computer-aided personalized system of instruction (CAPSI) holds great potential as a training delivery system (Kinsner & Pear, 1988, 1990; Pear & Crone-Todd, 1999, 2002; Pear & Kinsner, 1988). CAPSI is a computerized version of PSI developed by Fred Keller (1968). PSI is a non-traditional

method of instruction that relies on learning principles and it has been used to teach university courses (Hiltz, 1986; Kinsner & Pear; Pear & Crone-Todd; Pear, 2003; Pear & Novak, 1996). The key characteristics of PSI include: (a) specifying the course material and learning objectives clearly, (b) breaking down the course material into small units for study, (c) requiring mastery of each unit before proceeding to the next unit, (d) using proctors to grade tests, (e) allowing students to pace themselves, and (f) minimizing the use of lectures and emphasizing practice and participation (Eyre, 2007; Grant & Spencer, 2003). In a meta-analysis of 20 studies comparing PSI and conventional lecture-based classes, final examination scores were higher for PSI, with an average effect size of .49 across studies (Kulik, Kulik, & Bangert-Drowns, 1990). Although CAPSI has typically been used to deliver university courses, recent research has shown that it could be used to teach participants to conduct a discrimination assessment procedure (Hu et al., 2012).

Therefore, the purpose of this study was to extend research on teaching a preference assessment procedure by evaluating the self-instructional manual by Ramon and Yu (2010), incorporating video modeling as a teaching component, and delivering the training package to participants online using a modified CAPSI system. For the purpose of this study, the CAPSI system was modified to: (a) deliver both the manual and videos online, (b) score the completed exercises automatically, and (c) manage the contingencies to allow participants to proceed to the next unit only after they have correctly completed the exercise in the previous unit. The review exercises in the manual were also modified to facilitate computer scoring.

Method

Participants and Setting

Six undergraduate university students and four staff participated. The university students, four males and two females, were recruited from the University of Manitoba via recruitment posters displayed in various locations throughout the campus. The student participants ranged from 18-33 years of age and had completed between two to five years of University. Areas of study included Science, Engineering, Fine Art, and Environmental Design (see Table 1). The staff participants, one male and three females, included three educational assistants and one respite worker. Staff participants were recruited from the Pembina Trails School Division in Winnipeg. They ranged from 25-48 years of age and had between 5-13 years of experience working with persons with special needs. Two of the staff participants were also attending university (see Table 2).

None of the participants reported prior training on preference assessment procedures or having taken any courses that used CAPSI. Written informed consent was provided by each participant before the study began and each participant received an honorarium of $40 regardless of how they performed (i.e., the honorarium was given to each participant at the beginning of their first session in the study). During all training sessions, participants completed their studying in a private office. The experimenter was seated at a desk outside the office and was available to the participant at all times. Participants had the option of closing the office door for privacy.

Table 1

Characteristics of Student Participants

Students	Age (Yrs)	Gender	Years of University	Area of Study
1	20	Male	2	Electrical Engineering
2	23	Male	5	Fine Arts
3	18	Female	2	Faculty of Science
4	28	Male	3	Environmental Design - City Planning
5	33	Male	2	Electrical Engineering
6	24	Female	4	Faculty of Science

Table 2

Characteristics of Staff Participants

Staff	Age (Yrs)	Gender	Position	Years of Experience	Currently in University?
1	28	Female	Educational Assistant	5	Environmental Design
2	36	Female	Educational Assistant	5	No
3	48	Female	Educational Assistant	13	No
4	25	Male	Respite Worker	5	Kinesiology

All simulated assessments sessions (described below) were conducted in a testing room at St.Amant Research Center. During simulated assessments, the simulated client was seated at a table across from the participant. All the necessary stimuli required for the session (see below) were presented at a smaller table beside the participant.

Materials

Participants were given access to a computer with an internet connection (to access the web-based CAPSI system). The computer was connected to a printer. Participants were given the option to use headphones or speakers to hear the videos during CAPSI delivery. Datasheets, a stop watch, calculator, pencil, and a variety of leisure items were provided to the participant to conduct preference assessments. A video camera was used to record all simulated assessment sessions.

Research Design and Dependent Variables

Design. A multiple-baseline design replicated across two clusters of three students each and two pairs of staff was used to evaluate the effectiveness of the self-instructional training package. The number of baseline samplings were reduced to no more than two for each participant. In this study, the baseline condition consisted of the method description and the intervention consisted of the self-instructional training package delivered by a modified CAPSI system (described below). After baseline, the intervention was introduced successively for each participant after an intervention effect had been observed with the preceding participant (Johnston & Pennypacker, 2009; Martin & Pear, 2011).

Dependent Variables. The main dependent variable was performance accuracy in applying the MSWO procedure during simulated assessments conducted after studying

the method description (post-Method), after studying the self-instructional package via CAPSI (post-CAPSI), and during retention and generalization tests that occurred from 7 to 17 days following training. The duration of the retention interval varied due to participants' availability. All simulated assessments were videotaped and scored by the experimenter and a second observer. Each assessment consisted of three sessions and accuracy in applying the procedure was scored using the 25-item behaviour checklist developed by Ramon and Yu (2010, Appendix A). The percentage of checklist items carried out correctly was computed and reported for each session. Target behaviours on the behaviour checklist were divided into four areas: pre-assessment preparation, antecedent behaviours, consequences for various responses, and summarizing and interpreting the results. A minimum of seven and a maximum of 11 target behaviours, depending on the client's response, were evaluated during each trial within an assessment. For each simulated assessment, two additional target behaviours were evaluated for the first session to assess the sampling procedure and an additional three behaviours were added for the third session to assess calculation and ranking of the items. In addition to performance accuracy during simulated assessments, the amount of time spent studying during each phase and the participants' perception of the social significance (Kazdin, 2011) of the study was evaluated using a questionnaire (described below).

Procedure

Method description. During the baseline phase, participants were asked to study the two-page method description of the MSWO procedure adapted from published papers by DeLeon and Iwata (1996) and by Roscoe et al. (2006). This adapted method

description was identical to the one implemented in Ramon et al.'s (2012) study. A written set of instructions that described what was expected of the participants was provided at the beginning of the method description (see Appendix B for the instructions to participants and the two-page method description). Participants were given as much time as they needed. After a participant had finished studying the procedure, he/she conducted a simulated assessment.

Self-instructional package. The self-instructional intervention package consisted of presenting the self-instructional manual and videos through a modified CASPI system. During the intervention, participants were provided with access to a computer with an internet connection, access to a printer so they could print materials if they wished to do so, a timer, pencil, calculator and a set of instructions (see Appendix C). The self-instructional manual (Ramon & Yu, 2010) was presented in eight study units (see Appendix D for a list of the study units for the MSWO procedure presented through the web-based CAPSI system). Unit 1 included a brief introduction to the concept of preference assessment and described how to prepare before a session (e.g., arranging the setting, and getting the materials ready). Unit 2 explained how to present assessment trials. Unit 3 described what to do after a client selects an item. Unit 4 described what to do if the client did not select an item. Unit 5 described what to do if the client selected more than one item simultaneously. Unit 6 described how to calculate the results and rank each item in the preference assessment. Unit 7 explained how to interpret the ranking of the items. Finally, Unit 8 provided a cumulative review of units one through seven (see Appendix E for the cumulative behaviour checklist for MSWO assessments).

Units 1 through 5 were accompanied by video clips of the experimenter demonstrating the target behaviours covered in each unit and Unit 8 was accompanied by a video clip demonstrating five consecutive trials of the MSWO procedure. Participants were able to view the videos online after accessing the unit content and could replay the videos as many times as needed. Participants also could review videos for units that they had completed.

In the web-based CAPSI system, participants accessed each unit online by clicking on links to open text files and video clips. After studying a unit, participants clicked on a link to complete the review exercise corresponding to that unit. Within the review exercise, participants were required to answer all questions and submit the answers by clicking a link. There was no time limit for completing a review exercise. During a review exercise, the CAPSI system prevented access to the study materials for that unit. Participants could also cancel a review exercise at any time and return to reading the written material if they decided that they had not sufficiently studied the material.

Participants were required to answer all of the questions correctly in a review exercise before proceeding to the next unit. In order to implement computer scoring, study questions that required short answers in the original manual (Ramon & Yu, 2010) were replaced with fill-in-the-blank questions that required one or two word answers. All answers were checked by the computer program in the CAPSI system and feedback was provided immediately. Participants were automatically shown their completed review exercise and all possible correct responses for each question after completing a review exercise. A “pass” was issued if the participant answered all the questions in the review

exercise correctly. After receiving a "pass", the CAPSI system presented a link to the next unit and prompted the participant to proceed. A "restudy" was issued if the participant had one or more incorrect responses in their review exercise. If a participant received a "restudy", the computer prompted the participant to review the material and complete the review exercise again. A link was then provided to return to the study material. All questions in the exercise were presented each time. If a participant was issued a "restudy", the computer also prompted the participant to "appeal" the decision if they believed that the computer had inaccurately scored their review exercise by seeking out the experimenter who was seated just outside room.

Simulated assessment. A simulated assessment involved conducting an MSWO assessment with an actor (graduate student) playing the role of a person with developmental disabilities with no speech. At the beginning of the assessment, participants were provided with written instructions asking them to find out what the actor liked or disliked (see Appendix F). Participants were also provided with an MSWO assessment datasheet (Appendix G), a container with six items, a timer, pencil, and a calculator. Participants did not have access to the study materials during the assessment. The actor's behaviours were scripted (see Appendix H for a sample script) to ensure consistency across participants and to ensure that all target behaviours in the 25-item behaviour checklist were sampled. Participants were not provided feedback on how they had performed during simulated assessments.

Retention and generalization assessments. To assess retention, participants were asked to conduct a simulated assessment at least one week after the post-CAPSI assessment with the same items. Instructions were identical to the baseline phase except

participants were provided with the cumulative behaviour checklist provided in Unit 8 of the manual. Participants were given as much time as they needed to review the behaviour checklist prior to completing the assessment. All participants took approximately 3 to 5 minutes to review the checklist.

Immediately following the retention test, participants were asked to conduct an assessment with a different actor and with food instead of leisure items. A different set of written instructions (see Appendix I) were provided at the beginning of the assessment.

Reliability

Interobserver reliability. Reliability checks were conducted for all simulated assessments. A trained observer blinded to the phases of the study viewed the videotaped sessions and scored the participant's behaviours independently using the behaviour checklist in Appendix A. An agreement was defined as the experimenter and observer recording a response the same; otherwise, it was considered a disagreement. Percent agreement was calculated for each session by dividing the number of agreements by the sum of agreements and disagreements, and multiplying by 100% (Martin & Pear, 2011). The second observer practiced scoring videotaped sessions (developed for training purposes) until 100% agreement between the experimenter and observer was achieved prior to scoring videotaped sessions for the current study. Mean agreement score across participants and sessions was 98% (range 84-100%) for students and 97% (range 89-100%) for staff.

Simulated assessment script adherence. The actor's script adherence was evaluated for all simulated assessments. A trained observer viewed the videotaped sessions and scored whether the actor's response on each trial was correct (followed the

script) or incorrect (did not follow the script). Percentage of correct script adherence was 100% for all sessions.

Social validity

Each student participant was asked to complete a brief 6-item questionnaire (see Appendix J) after the self-instructional intervention. Each staff participant was asked to complete the same questionnaire twice, after the method description and after the self-instructional intervention, respectively. Participants indicated the extent to which they agreed with each of the six statements in the questionnaire on a five-point scale from strongly disagree (score 1) to strongly agree (score 5).

Results

Student Participants

Figure 1 shows the performance accuracy across simulated assessment sessions during each phase for Students 1 through 3. After studying the method description in baseline, Student 1's baseline performance averaged 23% correct (range, 18-33%) and showed a small gradual increase across the three sessions. However, he showed an immediate and large improvement after studying the manual and watching the video via CAPSI, achieving mastery for all three sessions (M = 99%, range 98-100%). He maintained his performance near 100% correct across sessions during retention (M = 98%, range 93-100%) and generalization (M = 100%) assessments, conducted 11 days after training.

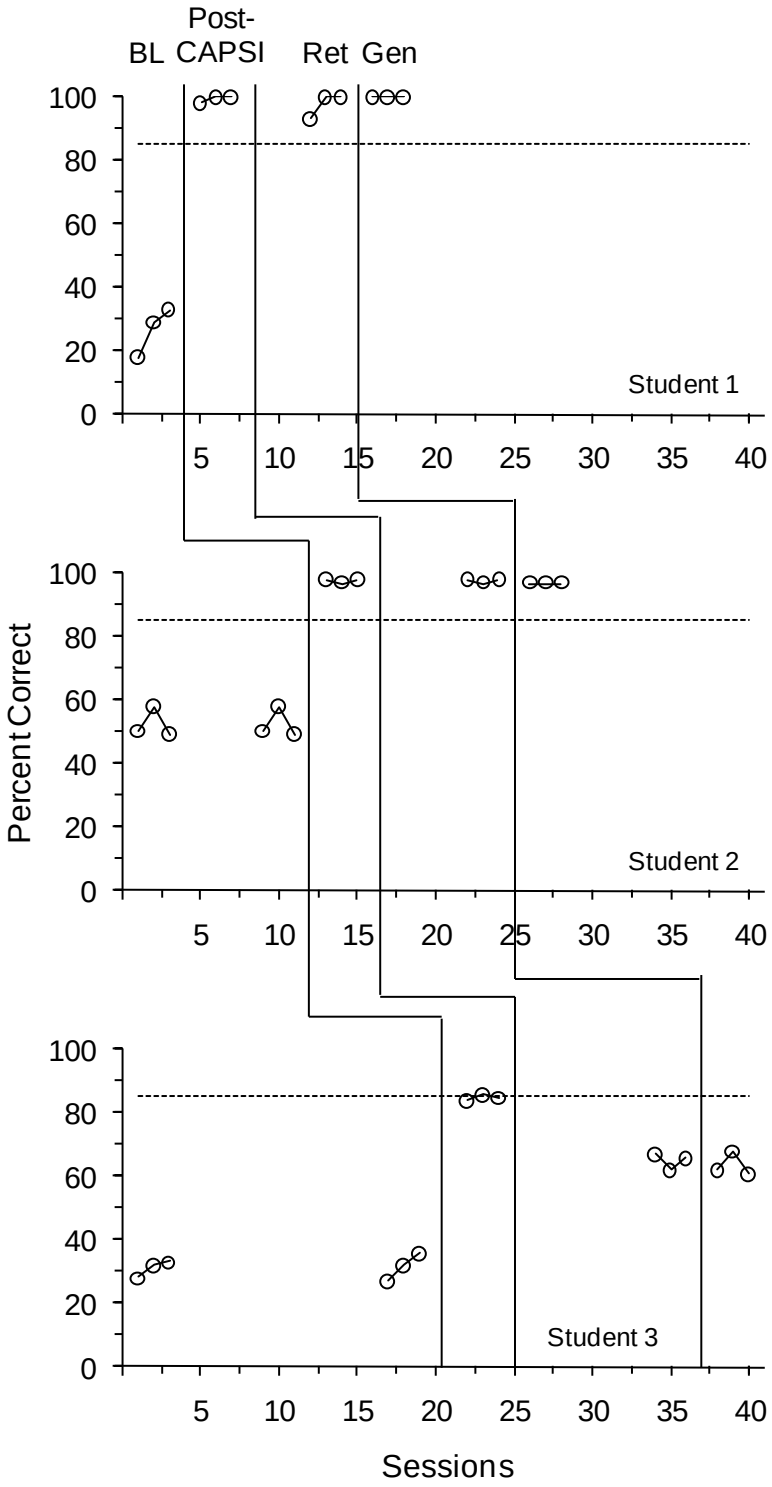

Figure 1. Percentage of correct responses per session during simulated preference assessments in baseline (BL), after manual and video delivered via CAPSI (Post-CAPSI), and during retention (Ret) and generalization (Gen) for Students 1 through 3. Dash lines indicate mastery criterion of 85% correct.

Student 2's baseline performance remained relatively stable across three sessions during the first simulated assessment. After Student 1 had achieved mastery, Student 2 was exposed to the baseline training (method description) again and then repeated the simulated assessment. His performance was quite stable with a mean of 52% (range 48-58%) across the six baseline sessions. He also achieved mastery after the CAPSI intervention (*M* = 98%, range 97-98%) and he maintained his performance near 100% correct during the retention (*M* = 98%, range 97-98%) and generalization (*M* = 97%) assessments, conducted 7 days after training.

Student 3's baseline performance was also relatively stable across the two simulated assessments after Students 1 and 2 had both achieved mastery. Her mean baseline performance accuracy was 31% across 6 sessions (range 26-36%). She achieved mastery after the CAPSI intervention during two of the three sessions although she missed the mastery criterion in the first session by only 1% (*M* = 85%, range 84-86%). Compared to Students 1 and 2, her performance at post-CAPSI was weaker. Moreover, she did not maintain her performance above the mastery criterion during the retention (*M* = 65%, range 62-67%) and generalization (*M* = 64%, range 61-68%) assessments, conducted 17 days after training.

Figure 2 shows the performance accuracy across simulated assessment sessions during each phase for Students 4 through 6. After studying the method description in baseline, Student 4's baseline performance remained relatively stable across the three sessions (*M* = 22%, range 18-27%). He showed an immediate and large improvement after studying the manual and watching the video in the CAPSI condition, achieving

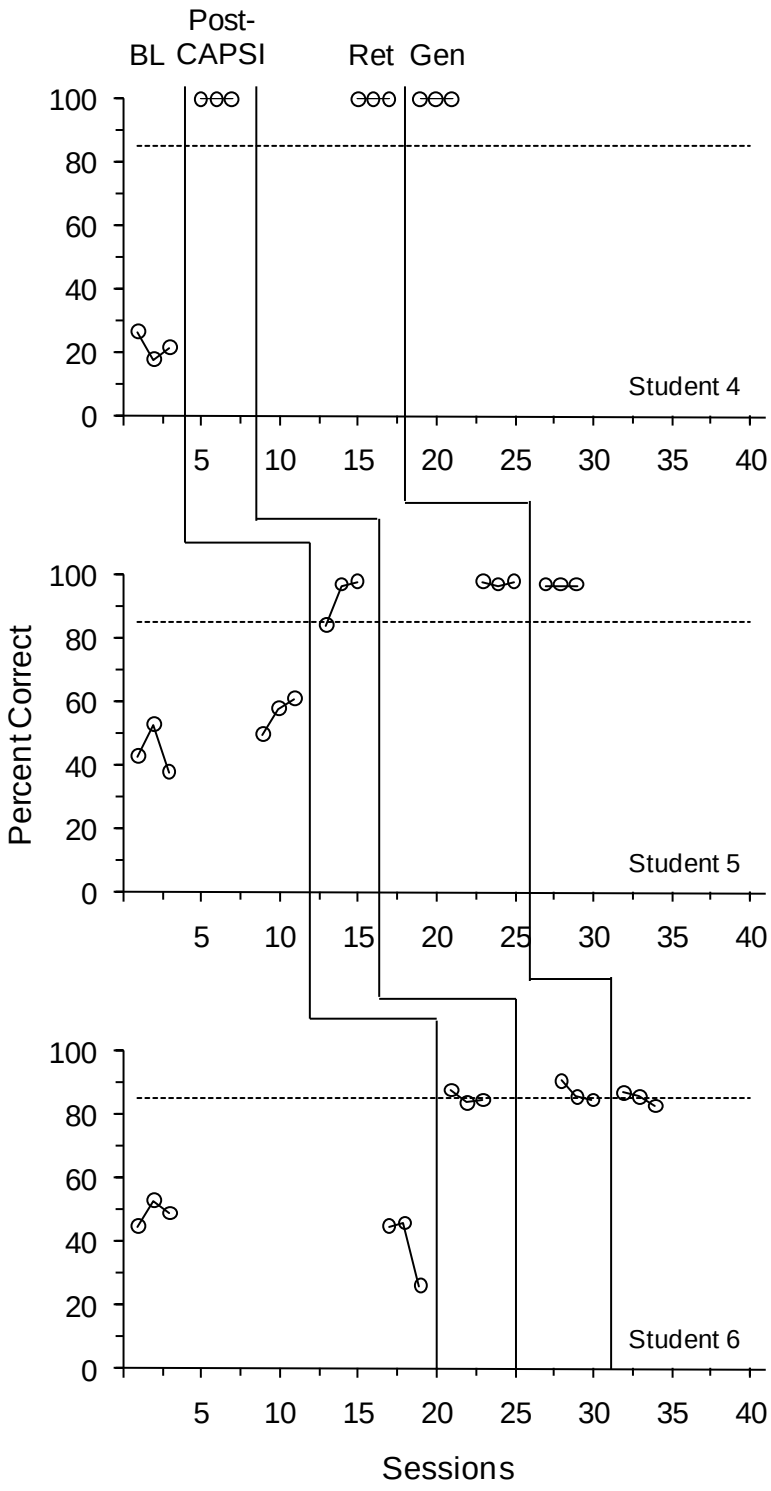

Figure 2. Percentage of correct responses per session during simulated preference assessments in baseline (BL), after manual and video delivered via CAPSI (Post-CAPSI), and during retention (Ret) and generalization (Gen) for Students 4 through 6. Dash lines indicate mastery criterion of 85% correct.

100% for all three sessions and maintaining his performance at 100% correct in every session during retention and generalization assessments, completed 7 days after training.

Student 5's baseline performance averaged 50% (range 38-61%) across six sessions. He also achieved mastery after the CAPSI intervention during two of the three sessions and missed the mastery criterion in the first session by only 1% (*M* = 93%, range 84-98%). He maintained his performance near 100% correct during retention (*M* = 97%, range 97-98%) and generalization (*M* = 97%) assessments, conducted 7 days after training.

Student 6's baseline performance accuracy decreased slightly across the simulated assessments with a six-session mean of 34% (range 26-49%). She achieved mastery after the CAPSI intervention during two of the three sessions although she missed the mastery criterion in the second session by only 1% (*M* = 86%, range 84-88%). She maintained her performance at approximately the same level during retention (*M* = 87%, range 85-91%) and generalization (*M* = 85%, range 83-87%) assessments completed 7 days after training. However, she missed the mastery criterion by 2% during the third session.

Staff Participants

Figure 3 shows the performance accuracy across simulated assessment sessions during each phase for Staff. The top two graphs show the multiple baseline design across Staff 1 and 2 and the bottom two graphs show the multiple baseline design across Staff 3 and 4. After studying the method description in baseline, Staff 1's baseline performance showed a small increase across the three sessions from 45% to 55% (*M* = 52%). Although she showed improvement after studying the manual and watching the video in the CAPSI condition, she did not demonstrate a large and immediate improvement in the first

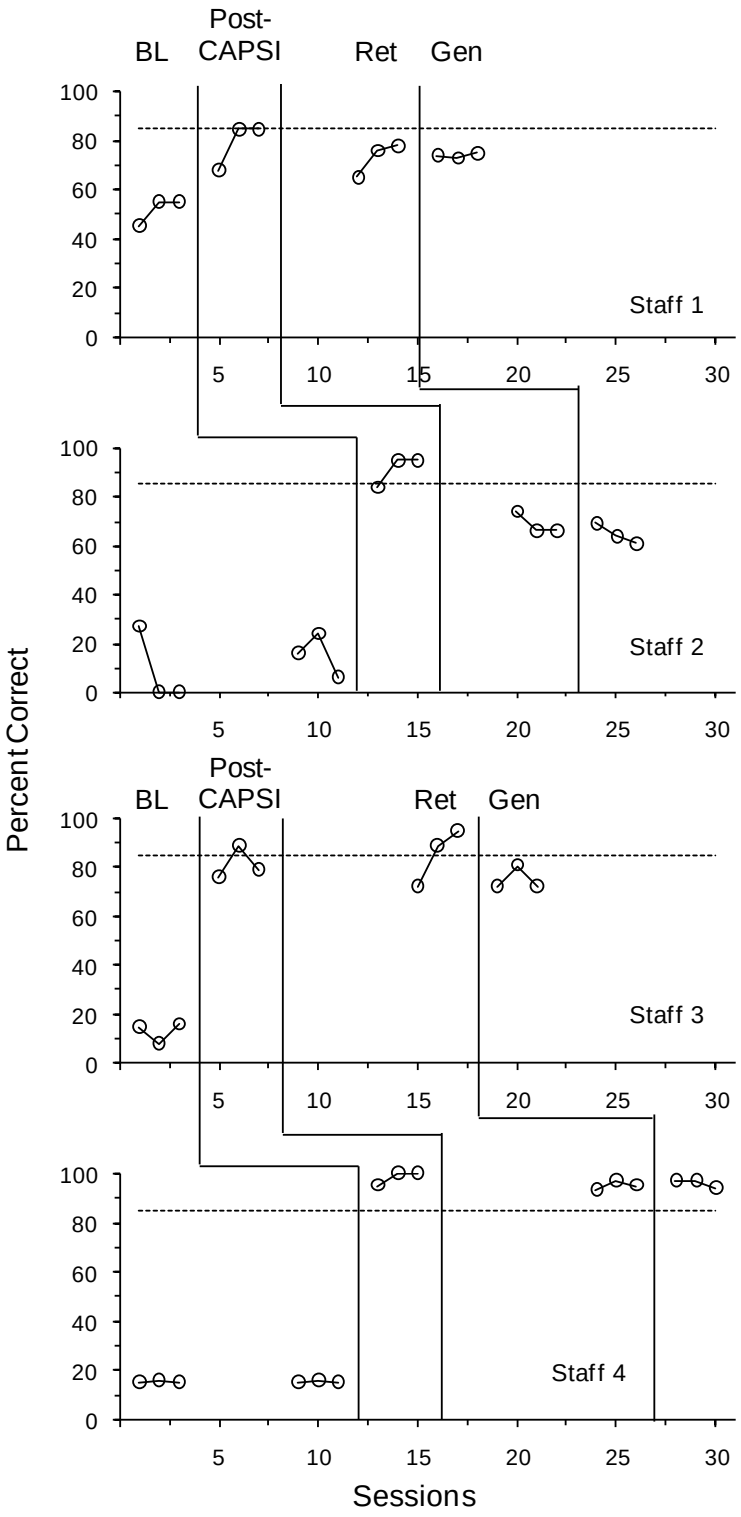

Figure 3. Percentage of correct responses per session during simulated preference assessments in baseline (BL), after manual and video delivered via CAPSI (Post-CAPSI), and during retention (Ret) and generalization (Gen) for Staff 1 through 4. Dash lines indicate mastery criterion of 85% correct.

session. Mastery was achieved in sessions two and three. Her performance averaged 79% across the three sessions, ranging from 68-85%. However, her performance accuracy decreased to below the mastery criterion during retention (*M* = 73%, range 65-78%) and generalization (*M* = 74%, range 73-75%) assessments completed 7 days after training. Staff 2's performance remained fairly low across all baseline sessions (*M* = 12%, range 6-27%). She showed a large and immediate improvement at post-CAPSI and achieved mastery during two of the three sessions, missing the mastery criterion by 1% in the first session (*M* = 91%, range 84-95%). However, performance accuracy decreased during the 11-day retention (*M* = 69%, range 66-74%) and generalization (*M* = 65%, range 61-69%) assessments.

Staff 3's baseline performance accuracy was low and relatively stable across sessions (*M* = 13%, range 8-16%). Although she showed an immediate and large improvement after studying the manual and watching the video, she met the mastery criterion during only one of the three sessions (*M* = 81%, range 76-89%). Her performance improved slightly during the 14-day retention assessment (two sessions met criterion, *M* = 85%, range 72-95%), but decreased to below the mastery criterion during generalization (*M* = 75%, range 72-81%). Staff 4's baseline performance was stable, with a mean of 15% (range 14-15%) across all 6 sessions in baseline. He achieved mastery at post-CAPSI (*M* = 98%, range 95-100%) and maintained his performance near 100% correct during the 14-day retention (*M* = 95%, range 93-97%) and generalization (*M* = 96%, range 94-97%) assessments.

Study Time

The average amount of time spent studying the method description during baseline was 9.5 minutes (range 4-15 min) across all 10 participants (both students and staff). The average amount of time spent studying the self-instructional manual and watching the videos via CAPSI was 85 minutes (range 49-125 min) across all participants.

Social Validity Ratings

Table 3 shows the mean and range of the ratings for each statement in the social validity questionnaire. At post-CAPSI, the mean ratings for the first four statements were at or above 4 for both student and staff participants. That is, most participants found the goal of the study to be important, the materials provided adequate information and were easy to use, and most felt that they had successfully learned to conduct the assessment. Mean ratings by staff participants at post-CAPSI were notably higher compared to the mean ratings obtained at post-Method across most statements. Although student participants did not receive the social validity questionnaire following the method description in baseline, the results obtained post-CAPSI appeared consistent across both students and staff.

Discussion

The results of the study showed that the self-instructional manual plus video modeling delivered using a modified CAPSI system was successful in teaching students and staff to conduct the MSWO preference assessment procedure. All participants demonstrated a large and immediate improvement over baseline in performance accuracy during simulated client assessments after studying the self-instructional manual and watching the videos via CAPSI. While a slight increase in baseline performance was

Table 3

Participants' Mean Ratings (Range) for Each Statement in the Social Validity Questionnaire (1= strongly disagree, 5 = strongly agree)

	Students (n = 6)	Staff (n = 4)	
Statements	Post-CAPSI	Post-Method	Post-CAPSI
1. It is important for individuals working with clients with developmental disabilities to learn to conduct preference assessments.	4.33 (range 1-5)	3.75 (range 1-5)	4.00 (range 1-5)
2. The material was easy to follow and understand.	4.17 (range 4-5)	3.00 (range 2-5)	4.50 (range 4-5)
3. The material provided all the necessary information for me to do the assessment.	4.50 (range 4-5)	2.75 (range 2-4)	4.75 (range 4-5)
4. I believe I have successfully learned how to conduct the MSWO preference assessment from studying the materials.	4.00 (range 3-5)	2.00 (no range)	4.25 (range 4-5)
5. I feel confident and ready to conduct MSWO preference assessment with clients after studying the materials.	3.83 (range 2-5)	1.75 (range 1-2)	4.00 (range 4-5)
6. I will likely use this assessment with individuals with developmental disabilities.	3.17 (range 2-5)	4.00 (range 3-5)	4.25 (range 4-5)

observed for students 1, 3, and 5, the immediate and sizeable increase in performance following the introduction of the intervention lends support to the conclusion that the observed effect was in fact due to the intervention. All of the students achieved mastery level performance following the CAPSI intervention. Only one of the four staff achieved mastery in all three sessions at post-CAPSI, although all staff demonstrated immediate improvement following the intervention and met mastery criterion on at least one or two of the three sessions. Five of the six students performed at or above 85% correct during retention and generalization assessments, but only one of the four staff members maintained mastery level performance during both retention and generalization assessments. The average amount of time spent studying the self-instructional manual with the addition of video modeling (i.e., 85 minutes) was similar to that reported by Ramon et al. (2012).

The results of this study contribute to the literature in the following ways. First, it extends Ramon et al.'s (2012) study by using a different delivery method. The present study is the first to evaluate the effectiveness of delivering the self-instructional training package online to teach individuals to conduct a preference assessment procedure. CAPSI has primarily been used to date for delivering university courses (Kinsner & Pear, 1988; Pear & Crone-Todd, 1999, 2002; Pear, 2003; Pear & Novak, 1996). The present study showed that a modified CAPSI system has tremendous potential for teaching practical assessment skills to individuals without face-to-face instruction.

In particular, the modifications made to the CAPSI system extend previous CAPSI research by delivering both textual and video materials online and using automated computer scoring. Typically, study materials in previous CAPSI research were

provided to learners in paper form (e.g., Hu et al., 2012). In addition, since the scoring was automated in the present study, CAPSI also automatically managed the contingency for participants to progress through each study unit.

The automated computer scoring could be viewed as a computerized version of programmed instruction (e.g., Holland & Skinner, 1961; Skinner, 1968). In the typical CAPSI system, exam questions usually require written short answers that need to be marked by the course instructor, proctors, or peer-reviewers (Pear, 2003; Pear & Crone-Todd, 2002). While this was necessary for essay-type answers, the present study implemented computer scoring by using fill-in-the-blank questions.

Although the intervention package proved effective, the computer scoring system used in the present study will require further improvement. To allow for automated scoring, all of the possible answers for each question were entered into a database in the CAPSI system; however, minor variations in spacing or spelling were detected as errors. For example, an answer with an extra space between two words or a misspelled word was detected as an error by the computer. To compensate for the imperfect automated scoring system, all participants were instructed to "appeal" a review exercise that they felt was inaccurately scored by the computer. Appeals were made directly to the experimenter who was sitting outside the office where the participants were studying. Improvement in the programming of the automated computer scoring system will be needed to reduce these errors.

The present study also extends previous research on the characteristics of participants. Ramon et al. (2012) included only undergraduate university students and Graff and Karsten (2012) included teachers with Bachelor or Master's degrees. The

present study included educational assistants and a respite worker as participants in addition to university students. However, it should be noted that two of the four staff participants were also attending university at the time of the study.

Lastly, the present study extends research on the use of video modeling as a component for teaching individuals to conduct preference assessment procedures. Although Hu et al. (2012) included video modeling in their study on teaching university students to conduct discrimination assessments using CAPSI, only one study (Lavie & Sturmey, 2002) has included video modeling as a component to teach preference assessment.

Further research is informed by a post-hoc analysis of procedural errors made by the participants. Table 4 shows the frequency of procedural errors made by participants after receiving the self-instructional training package. The analysis revealed five errors that were common to both students and staff participants. These included failing to: (a) instruct the client to "look" at the item on each trial when presenting each item, (b) block the response when the client tried to select two items simultaneously, (c) present the correct items one at a time in the correct sequence, (d) remove unselected items from the table following a response, and (e) ensure that all the necessary information is recorded on the datasheet. Moreover, three of the four staff commented in their social validity questionnaire that the process for calculating the ranking of each item following the assessment was not sufficiently described in the manual and requested more examples. Interestingly, while common errors across groups were evident, the most frequently made error by staff was not observed in the student participants. This involved inaccurate delivery of the sampling procedure or failing to conduct the sampling procedure prior to

Table 4

Frequency of Procedural Errors Made by Participants during Post-CAPSI Assessment

Behaviours	**Students**	**Staff**
• 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.	0	**2**
• Let the client sample the food item or interact with the non-food item before the assessment.	0	**12**
• Provide instruction "look" when presenting each item singly to ensure client attends.	**30**	**10**
• Present correct items one at a time and in the correct sequence.	**4**	**5**
• Provide instruction "pick one" and give the client up to 15 seconds to choose.	0	0
• If client selects item within 15 seconds, allow access to item.	0	**3**
• Remove all unselected items from the table.	**3**	**2**
• After 30 seconds, remove the item from the client.	0	0
• Record the selected item on the datasheet.	0	**1**
• If the client did not choose an item after 15 s, repeat the request "Pick one" once.	0	**3**
• If the client still does not select an item after 15 s, the trial ends.	0	0
• Say nothing and remove all items.	0	0
• Record a zero on the datasheet and end the session.	0	0
• If the client tries to select two items simultaneously, *block* gently (saying "no" is optional).	**13**	**6**
• Remove and represent the items and the instruction "pick one"	0	0
• At the end of the assessment, transfer the ranks from each session to table at the bottom of the datasheet.	0	0
• Add the rankings for each item across the three sessions and divide by 3 to obtain an average rank for each item.	0	**2**
• Correctly identify the most preferred item on the datasheet.	0	**2**
• Ensure all the necessary information is on the datasheet (i.e., client and participant name, date, items)	**2**	**1**

starting the first session. It is possible that some components of the MSWO procedure may not be necessary for all clients (e.g., not all clients will need to be asked to look at the items) and their omission may not influence the assessment outcome. If a component does not require a lot of time to perform, it might be more practical to teach a standardized method that includes all of these components rather than introducing the added complexity of determining the conditions under which a step is or is not needed. Therefore, future research with the self-instructional manual should consider revising the relevant sections to reduce these errors.

Several limitations of the study should be noted. First, generalization was assessed with a student playing the role of an individual with developmental disabilities instead of a real client. Considerable effort was made to recruit individuals with developmental disabilities to participate in generalization assessments. Unfortunately, recruitment was unsuccessful. Although the previous evaluation of the self-instructional manual and live modeling did show strong generalization to real clients following training (Ramon et al., 2012), this has not been established for the self-instructional training package delivered using CAPSI. Examining generalization to the assessment of real clients should be considered a high priority in future research.

Second, a comparison of social validity scores post-Method and post-CAPSI was available for staff but not for students. Students did not complete the social validity questionnaire at post-Method due to an oversight. Although the ratings obtained post-CAPSI were similar across students and staff, the addition of ratings from students at post-Method would have enhanced the comparison.

Third, due to scheduling difficulties, it was not feasible to conduct a multiple baseline across three participants for the staff as conducted with the students. A multiple baseline was implemented across two staff participants and replicated across two pairs. Considering the consistency of the results replicated across two sets of three students each and across two pairs of staff, it was quite likely that the observed results were due to the intervention.

Fourth, the self-instructional training package in the present study was always preceded by the method description, thus it was not clear whether the intervention package (i.e., self-instructional manual, video, and CAPSI delivery) would have produced the same results without the baseline method description. This was also the case in Graff and Karsten's study (2012) in which their enhanced instruction was preceded by the procedural description in baseline. Although it is possible that the baseline method description could have contributed to the performance at post-CAPSI, the results of Ramon et al.'s (2012) study did show that the self-instructional manual (paper version) was moderately effective without it being preceded by the method description. Future research should examine the effectiveness of the self-instructional manual, video, and CAPSI delivery without it being preceded by other interventions.

Lastly, the time between post-CAPSI and retention assessments was not the same for all participants due to their availability and this was beyond the experimenter's control. Four out of six students and one of the four staff completed the retention and generalization assessments seven days following the intervention. Student 1 completed the retention and generalization assessments 11 days after training and Student 3 completed the assessments 17 days following the intervention; however, a decrease in

performance was observed only for Student 3. Both Staff 3 and 4 completed the retention and generalization assessments 14 days following the intervention and Staff 2 completed the assessments after 11 days. A decrease was observed in both retention and generalization assessments for Staff 2 while a small decrease was observed for Staff 3 during only the generalization assessment. Thus, the retention and generalization results were relatively strong despite the variability in retention interval.

Given the findings of this study, several additional suggestions for future research can be made. While both Ramon and Yu's (2010) self-instructional manual and the self-instructional package delivered in this study demonstrated promising findings for teaching the MSWO assessment, the contribution of video modeling, as implemented in the current study, has yet to be determined. Therefore, future studies should examine the effectiveness of video modeling in isolation. Second, future studies should also consider examining the effects of video modeling versus video instruction (e.g., instructions embedded in the video demonstration). With video instruction, it may be possible to reduce the length of the self-instructional manual or eliminate it completely. Lastly, future research is needed to extend this approach to other preference assessment procedures.

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.

Bojak, S. L., & Carr, J. E. (1999). On the displacement of leisure items by food during multiple stimulus preference assessments. *Journal of Applied Behavior Analysis, 32,* 515-518.

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.

Catania, A. C. (1998). *Learning (4th Ed.).* Upper Saddle River, NJ: Prentice Hall, Inc.

Catania, C. N., Almeida, D., Liu-Constant, B., & Digennaro-Reed, F. D. (2009). Video modeling to train staff to implement discrete-trial instruction. *Journal of Applied Behavior Analysis, 42,* 387-392.

Ciccone, F. J., Graff, R. B., & Ahearn, W. H. (2007). Long-term stability of edible preferences in individuals with developmental disabilities. *Behavioral Interventions, 22,* 223–228.

Collins, S., Higbee, T. S., & Salzberg, C. L. (2009). The effects of video modeling on staff implementation of a problem-solving intervention with adults with developmental disabilities. *Journal of Applied Behavior Analysis, 42,* 849-854.

Cote, C. A., Thompson, R. H., Hanley, G. P., & McKerchar, P. M. (2007). Teacher report and direct assessment of preferences for identifying reinforcers for young children. *Journal of Applied Behavior Analysis, 40,* 157–166.

Daly, E. J., III, Wells, N. J., Swanger-Gagne, M. S., Carr, J. E., Kunz, G. M., & Taylor, A. M. (2009). Evaluation of the multiple-stimulus without replacement preference assessment method using activities as stimuli. *Journal of Applied Behavior Analysis, 42,* 563-574.

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.

Digennaro-Reed, F. D., Codding, R., Catania, C. N., & Maguire, H. (2010). Effects of video modeling on treatment integrity of behavioural interventions. *Journal of Applied Behavior Analysis, 43,* 291-295.

Eyre, H. L. (2007). Keller's personalized system of instruction: Was it a fleeting fancy or is there a revival on the horizon? *The Behavior Analyst Today, 8,* 317-324.

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. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). Evaluation of a Training Manual: 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., & Karsten, A. M. (2012). Evaluation of a self-instruction package for conducting stimulus preference assessments. *Journal of Applied Behavior Analysis, 45,* 69-82.

Grant, L. K., & Spencer, R. E. (2003). The personalized system of instruction: Review and applications to distance education. *Instructional Review of Research in Open and Distance Learning, 4,* 1-17.

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.

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

Hanley, G. P., Iwata, B. A., & Roscoe, E. M. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis, 39,* 189–202.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (1999). The effects of pictorial versus tangible stimuli in stimulus preference assessments. *Research in Developmental Disabilities, 20,* 63-72.

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.

Hiltz, S. R. (1986). "The virtual classroom": Using computer-mediated communication for university teaching. *Journal of Communication, 36,* 95-104.

Holland, J. G., & Skinner, B. F. (1961). *The analysis of behavior: A program for self-instruction.* New York: McGraw-Hill.

Hu, L., Pear, J. J., & Yu, C. T. (2012). Teaching university students to implement the Assessment of Basic Learning Abilities using computer-aided personalized system of instruction. *Journal on Developmental Disabilities, 18(1),* 12-19.

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

Kazdin, A. E. (2011). *Single-case research designs (2nd ed.)*. New York, NY: Oxford University Press, Inc.

Keller, F. S. (1968). "Good-bye, Teacher . . .". *Journal of Applied Behavior Analysis, 1,* 79-89.

Kinsner, W., & Pear, J. J. (1990). *A dynamic educational system for the virtual campus. In Studies in technological innovation and human resources.* In U.E. Gattiker (Ed.). End-user training (Vol. 2, 201-228). Berlin: Walter de Gruyter.

Kinsner, W., & Pear J. J. (1988). Computer-Aided Personalized System of Instruction for the Virtual Classroom. *Canadian Journal of Educational Communication, 17,* 21-36.

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.

Kulik, C. L. C., Kulik, J. A., & Bangert-Drowns, R. I. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research, 60,* 265–299.

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.

Martin, G. L., & Pear, J. J. (2011). *Behavior modification: What it is and how to do it (8th ed.)*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

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.

Moore, J. W., & Fisher, W. W. (2007). The effects of videotape modeling on staff acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis, 40,* 197-202.

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

Pear, J. J. (2003). *Enhanced feedback using computer-aided personalized system of instruction.* E-xcellence in Teaching e-column – PsychTeacher Electronic Discussion List, November, 2003.

Pear, J. J., & Crone-Todd, D. E. (1999). Personalized system of instruction in cyberspace. *Journal of Applied Behavior Analysis, 32*, 205-209.

Pear, J. J., & Crone-Todd, D. E. (2002). A social constructivist approach to computer-mediated instruction. *Computers and Education, 38,* 221–231.

Pear, J. J., & Kinsner, W. (1988). Computer-aided personalized system of instruction: An effective and economical method for short- and long-distance education. *Machine Mediated Learning, 2,* 213-237.

Pear, J. J., & Novak, M. (1996). Computer-aided personalized system of instruction: A program evaluation. *Teaching of Psychology, 23,* 119–123.

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,* 345-259.

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.

Ramon, D., Yu, C.T., Martin, T., Chand, C., Lee, M., Astacio, R., Lavoie, K., & Simon, K. (2012, October). Evaluation of a training manual to teach multiple-stimulus preference

assessment. Poster presented at the 7th Annual Conference of the Manitoba Association for Behaviour Analysis, Winnipeg.

Ramon, D., & Yu, C. T. (2010). *How to conduct preference assessment with persons with developmental disabilities using a multiple stimulus without replaces preference assessment procedure: A self-instructional manual.* Unpublished Manuscript, Winnipeg, MB.

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

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

Roscoe, E. M., Iwata, B. A., & Kahng, S. W. (1999). Relative versus absolute reinforcement effects: Implications for preference assessments. *Journal of Applied Behavior Analysis, 32,* 479–493.

Salem, S., Fazzio, D., Arnal, L., Fregeau, P., Thomson, K, Martin, G. L., & Yu, C. T. (2009). A self-instructional package for teaching university students to conduct discrete-trials teaching with children with autism. *Journal on Developmental Disabilities, 15,* 21-29.

Skinner, B. F. (1968). *The technology of teaching.* New Jersey: Prentice-Hall.

Thiessen, C., Fazzio, D., Arnal, L., Martin, G. L., & Keilback, L. (2009). Evaluating a self-instructional strategy plus monetary contingency to train instructors to

conduct discrete-trials teaching with children with autism. *Behaviour Modification, 33,* 360-373.

Thomson, K. M., Martin, G. L., Fazzio, D., Salem, S., Young, K., & Yu, C. T. (2012). Evaluation of a self-instructional package for teaching tutors to conduct discrete-trials teaching with children with autism. *Research in Autism Spectrum Disorders, 6,* 1073-1082.

Tullis, C. A., Cannella-Malone, H. I., Basbigill, A. R., Yeager, A., Fleming, C. V., Paine, D., & Wu, P. (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,* 576-595.

Windsor, J., Piché, L. M., & Locke, 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 behavioural assessment system for the mentally handicapped: A behavioral approach. *Canadian Journal for Exceptional Children, 1,* 117-123.

Appendix A

Behavior Checklist

Preparation before an Assessment	**1**	**2**	**3**	**4**	**5**	**6**
Readying choices to be assessed (food only)						
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.	

Note. From Ramon et al., 2012.

Appendix B

Baseline Instructions and Method Description[1]

Thank you for helping me with this study. Today you will attempt to conduct a preference assessment. The preference assessment procedure is described below. Take as much time as you need to read and familiarize yourself with the procedure. After you finish studying, I will ask you to conduct a preference assessment using six items with <*name*> using this procedure. You will not be able to refer to the written procedure during the assessment. Let me know when you have finished studying.

[1] The method description is identical to that used by Ramon et al. (2012) and it has been adapted from the Method sections of 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". Please see published studies by DeLeon and Iwata (1996) and Roscoe et al. (2006) for a copy of the Method Description.

Appendix C

Instructions for Participants for the Self-Instructional Training Package

You will now read the manual on how to conduct preference assessment. The entire manual is presented online. You will need the following login and password to access the materials:

Login: _____

Password: _____

Follow the instructions provided throughout the program to help you navigate through the training manual. I have provided you with all the materials you will need: a computer with an internet connection, access to a printer, a calculator, pencil, and a stop watch. Please do your best to complete the manual. I will be seated just outside the room. Please let me know if you have any difficulty with the computer. You can begin whenever you are ready by logging on to the training program. When you are done we will conduct the next assessment with the simulated client.

Appendix D

Study Units for the MSWO Procedure Presented Through the CAPSI System

Unit 1: What is Preference?
- What is a Direct Preference Assessment?
- Why is Assessing Preference Important?
- Multiple-Stimulus-Without-Replacement (MSWO) Preference Assessment
- Preparing Before an Assessment
- Sampling Procedure

Video 1
Review Exercise 1
Unit 2: Presenting Trials
Video 2
Review Exercise 2
Unit 3: Consequence provided when the client selects an item
- a) Consequence provided when a client selects an item within 15 seconds
- b) Consequence provided when a client selects one item and then another item

Videos 3a and 3b
Review Exercise 3
Unit 4: Consequence provided when a client does not select an item
Video 4
Review Exercise 4
Unit 5: Consequence provided when the client selects more than one item simultaneously
Video 5
Review Exercise 5
Unit 6: Ranking the Items
Review Exercise 6
Unit 7: What Does the Rank Mean? Interpreting the Results.
Review Exercise 7
Unit 8: Cumulative Behavior Checklist for MSWO Assessments
Video 6
Review Exercise 8

Note: The units correspond to the sections in the self-instructional manual developed by Ramon and Yu (2010).

Appendix E

Cumulative Behaviour Checklist for MSWO Assessments

Setting

- Do your assessment in a quiet place.
- Have a table and chairs, and a place for your datasheet 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 datasheet: 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 datasheet 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 datasheet 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 datasheet.
- Add the rankings for each item across the three sessions and divide by 3 to obtain an average rank for each item.

Note: This checklist was taken from the self-instructional manual developed by Ramon and Yu (2010).

Appendix F

Instructions to Participants for Simulated Assessments

You will now conduct a preference assessment with <*name*>. <*name*> is a graduate student playing the role of an individual with developmental disabilities with no speech. She will not be able to answer any of your questions during your session. I cannot provide you with any additional information about the assessment procedure. Please do your best to find out what <*name*> likes and dislikes using the materials provided (datasheet, pencil, calculator, and stopwatch) and the six activities: blocks, car, train, Lego, play dough, and glow stick. You can begin whenever you are ready. You have up to 15 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.

Appendix G

Multiple-Stimulus without Replacement Preference Assessment Datasheet

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 H

Sample Simulated Client Script

SESSION 1	First Response	Second Response
Trial 1	Touch only the train	
Trial 2	No response	Touch the Lego
Trial 3	Simultaneously touch more than one toy	Touch only the playdoh
Trial 4	Touch only the glow stick	
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 train with your other hand	
Trial 2	Touch only the train	
Trial 3	Simultaneously touch more than one toy	Touch only the Lego
Trial 4	Do not respond	Touch the glow stick
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 glow stick	
Trial 2	Touch the glow stick	
Trial 3	Touch the car	

Trial 4	Simultaneously touch more than one toy	Touch only the playdoh
Trial 5	No response	No response

Note. From Ramon et al., 2012.

Appendix I

Instructions to Participants for the Generalization Assessment

Thank you for helping me with this study. Today you will conduct a preference assessment using six items with <*name*>. <*name*> is a graduate student role playing a client with developmental disabilities with minimal to no speech. Using the knowledge you have acquired from your training please do your best to find out what <*name*> likes and dislikes using the materials I have provided (datasheet, pencil, calculator, stopwatch), and the six food items: crackers, smarties, chips, cereal, granola bar, and cheezies. I cannot provide you with any additional information about the assessment procedure. You can begin whenever you are ready. You can let me know when you are finished and we can stop. Let me know when you are ready to begin.

Appendix J

Social Validity Questionnaire

Name: _____ Date:_____

Thank you for helping me with this study. Please complete the following questionnaire before leaving today. Your feedback is greatly appreciated. You do not need to put your name on the form.

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

1. It is important for staff and/or parents 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 material was easy to follow and understand.

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

3. The 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 individuals with developmental disabilities 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 individuals with developmental disabilities (e.g., child, client).

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

Other Comments:_____

Note. Adapted from Ramon et al. (2012).