

Development and Evaluation of a Mobile Device Application and a Video Training Component
for Conducting Stimulus Preference Assessments

by

Flávia M. Julio

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Winnipeg

Copyright © 2019 by Flávia M. Julio

STIMULUS PREFERENCE ASSESSMENT

Acknowledgements

I am very grateful to my advisor Dr. C.T. Yu for his knowledgeable guidance and support. It has been an honour to learn from him. I would also like to thank my committee members, Dr. Kerri Walters, Dr. Shahin Shooshtari, and Dr. Joseph Pear, for their valuable inputs, and also Dr. Lorna Jakobson and Dr. Cameron for their recommendations during my proposal. I would like to acknowledge my colleagues from University of Manitoba for their collaboration, and single out Ryan Heckert for his great dedication assisting with data collection; May Lee and Lisa Hunter for helping with the video training by accepting being the actors; and Lindsay McCombe for assisting in many different ways throughout my research. I express my sincere gratitude to the participants and clients without whom this research would not have been possible. I am also grateful for the love, support and encouragement given by my immediate and extended family, including my parents, Sandra and Edmundo Julio, my sister, Marina Matielli, and my husband, Toby Martin.

STIMULUS PREFERENCE ASSESSMENT

To Toby and Isabella.

STIMULUS PREFERENCE ASSESSMENT

Abstract

Two studies were conducted to evaluate an iOS® application for conducting stimulus preference assessments. Study 1 examined the effectiveness of three video tutorials to teach university students and staff who work with individuals with intellectual and development disabilities to implement paired-stimulus, multiple-stimulus-without-replacement, and free-operant procedures. Fourteen participants met the mastery criterion to perform the preference assessment procedure from watching the videos and one participant needed additional instruction to learn the procedure. Following training, participants showed high procedural fidelity during generalization assessment with a client with autism or with a confederate, and all participants found that the videos were easy to understand and that the iOS® app was helpful. In Study 2, six staff participants were asked to use the app in their normal work settings for a 1-month period. Staff were then asked to complete a 23-item questionnaire to provide feedback on the app. The participants rated the app and the videos positively and provided positive qualitative comments. The strengths and limitations of each study are discussed.

STIMULUS PREFERENCE ASSESSMENT

Table of Contents

Acknowledgements..... i

Abstract..... iii

Table of Contents..... iv

List of Tables..... ix

List of Figures..... x

Introduction..... 1

Stimulus Preference Assessments..... 3

 PS procedure..... 4

 MSWO procedure..... 5

 FO procedure..... 6

 Other preference assessment procedures..... 8

Preference Assessment Applications for Mobile Devices..... 8

 Assessment procedures..... 9

 Aid to select the appropriate procedure..... 9

 Procedural prompts..... 9

 Data logging capabilities..... 10

 Evidence-based tutorials..... 10

Development of Prefer..... 11

Preference Assessment Training Research..... 12

 Self-instructional manuals..... 13

 Training with observing and rating..... 15

STIMULUS PREFERENCE ASSESSMENT

Video training.....16

Development of Video Training for Prefer.....21

Clinical Utility of Stimulus Preference Assessments.....22

Rationale for the Present Studies.....23

STUDY 1: VIDEO TRAINING FOR PREFERENCE ASSESSMENTS.....24

Method.....25

Participants and Settings.....25

Materials.....26

Design.....26

Video Tutorials.....27

Behaviour Measures.....27

Baseline Simulated Assessment.....29

Video training.....31

Post Video Training Simulated Assessment.....31

Feedback Questionnaire.....31

Generalization Assessment with a Client.....32

Simulated Assessment with New Confederate.....33

Comparison Between Participant and Experimenter.....33

Interobserver Agreement on Participant’s Performance.....33

Confederate Adherence to the Script.....34

Results.....34

Paired-Stimulus Procedure.....34

STIMULUS PREFERENCE ASSESSMENT

Multiple-Stimulus-without-Replacement Procedure.....35

Free-Operant Procedure.....37

Feedback Questionnaire.....38

Comparison between Participant and Experimenter.....39

 PS Procedure.....39

 MSWO Procedure.....40

 FO Procedure.....41

Time Spent in Generalization Assessment with a Client.....42

Discussion.....42

STUDY 2: CLINICAL UTILITY OF PREFER.....45

Method.....45

Participants.....45

Material.....46

Procedure.....47

 Training.....47

 Using *Prefer*.....47

 Data collection.....47

 Help Me Choose.....47

 Video training.....48

 Prompts during assessment.....48

 Back-up.....48

 Future usage.....48

STIMULUS PREFERENCE ASSESSMENT

Reinforcers.....48

Results.....49

Help me Choose.....49

Video Training.....50

Prompts During Assessment.....50

Back-up.....51

Future Usage.....51

Reinforcers.....52

Discussion.....52

General Discussion.....54

References.....57

Appendix A: Screen Shoots of the *Prefer* App.....83

Appendix B: YouTube® Links, and Screen Shoots from the Multiple Stimulus without-
Replacement Video.....87

Appendix C: Procedural Integrity Checklist for the Paired-Stimulus Procedure.....92

Appendix D: Procedural Integrity Checklist for the Multiple-Stimulus-without-Replacement
Procedure.....93

Appendix E: Procedural Integrity Checklist for the Free-Operant Procedure.....94

Appendix F: Feedback Questionnaire.....95

Appendix G: Demographic Questionnaire.....96

Appendix H: Slides from tutorial on how to use *Prefer*.....99

STIMULUS PREFERENCE ASSESSMENT

Appendix I: Clinical Utility Questionnaire.....107

STIMULUS PREFERENCE ASSESSMENT

List of Tables

Table 1. Comparison of Preference Assessment Apps.....67

Table 2. Characteristics of Participants in Study 1.....68

Table 3. Feedback Questionnaire Results in Study 1.....69

Table 4. Participant Characteristics in Study 2.....70

Table 5. Summary of Results from Clinical Utility Questionnaire in Study 2.....71

List of Figures

Figure 1. Percent of correct responses during Paired-Stimulus simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 1 through 3. The horizontal dash line in each graph represents 85% correct. 74

Figure 2. Percent of correct responses during Paired-Stimulus simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 4 and 5. The horizontal dash line in each graph represents 85% correct. 75

Figure 3. Percent of correct responses during Multiple-Stimulus-without-Replacement simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 6 through 8. The horizontal dash line in each graph represents 85% correct. 76

Figure 4. Percent of correct responses during Multiple-Stimulus-without-Replacement simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 9 and 10. The horizontal dash line in each graph represents 85% correct. 77

Figure 5. Percent of correct responses during Free-Operant simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 11 through 13. The horizontal dash line in each graph represents 85% correct. 78

Figure 6. Percent of correct responses during Free-Operant simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a confederate for Participants 14 and 15. The horizontal dash line in each graph represent 85% correct. 79

Figure 7. Comparison of the results of the Paired-Stimulus assessments conducted by the researcher using paper-and-pencil, and by Participants 1, 2, 4, and 5 using *Prefer*. Items are ordered from the most to the least preferred on the horizontal axis based on the researcher’s assessment. 80

Figure 8. Comparison of the results of the Multiple-Stimulus-without-Replacement assessment conducted by the researcher using paper-and-pencil, and by Participants 8, 9, and 10 using *Prefer*. Items are ordered from the most to the least preferred on the horizontal axis based on the researcher’s assessment. 81

STIMULUS PREFERENCE ASSESSMENT

Figure 9. Comparison of the results of the Free-Operant assessment conducted by the researcher using paper-and-pencil, and Participants 12 and 13 using *Prefer*. Items are ordered from the most to the least preferred on the horizontal axis based on the researcher's assessment.....82

Development and Evaluation of a Mobile Device Application and a Video Training Component for Conducting Stimulus Preference Assessments

Introduction

Behavioural assessment involves collecting information about specific behaviours of a person usually through direct observation. The information may be used to assist making a diagnosis, quantify the level of a behaviour, identify environmental variables that may be maintaining the behaviour to assist in designing treatments, and to evaluate the effects of an intervention (Hawkins, 1979).

Behaviour assessments are useful for treating people with developmental disabilities (DD), which can be defined as “disabilities attributable to a mental or physical impairment, manifested before 22 years of age, likely to continue indefinitely, resulting in substantial limitation in three or more specified areas of functioning, and requiring specific and lifelong or extended care” (Developmental Disabilities Bill of Rights Act, Public Law 95-682, 1978). Developmental disabilities are estimated to occur in 1% of the world population (Maulik, Mascarenhas, Mathers, Dua, & Saxena, 2011), and may include challenges in communication, including to describe preferred items.

In the early 1990s, research had extended the behavioural assessment methodology to identifying preferences of individuals with developmental disabilities who are typically unable to describe to caregivers what they like (e.g., Fisher et al., 1992). There is now ample research demonstrating the effectiveness of this methodology, referred to as stimulus preference assessment, in identifying reinforcers for use in treatment programs (e.g., see review by Tullis et

al., 2011). Despite the usefulness and importance of stimulus preference assessments, uptake by practitioners who work with clients with developmental disabilities has been slow (Graff & Karsten, 2012a). Consequently, some recent research has focused on developing and evaluating the effectiveness of low-cost procedures to teach individuals to implement stimulus preference assessments (e.g., Hansard & Kazemi, 2018; Miljkovic, Kaminski, Yu, & Wishnowski, 2015; Ramon, Yu, Martin, & Martin, 2015). Although software programs have been developed to assist in data collection (Whiting & Dixon, 2012, 2014) and programs designed for clinicians and practitioners have tremendous potential to promote uptake, relatively little has been done in developing computer programs for stimulus preference assessments and in researching their effectiveness and clinical utility. The current research aimed to address this gap.

Computers have the potential to make collecting behavioural data simpler and faster. With the aid of specialised software, practitioners could allocate more time to the intervention relative to data collection. Several software programs have been developed or adapted to enhance behavioural data collection (Kahng & Iwata, 1998). However, excessive cost and low portability can limit uptake by behavioural practitioners. Moreover, with ever-changing technological standards, sophisticated equipment quickly becomes out-of-date, which increases the demand for software updates and development for new platforms.

Mobile devices are an appealing alternative for data collection due to their portability and popularity. Mobile devices that use the Apple Operating System such as iPhone®, iPod®, and iPad®, are commonly referred to as iOS® devices. These devices can serve a number of purposes, since many programs, called applications or “apps”, can be downloaded for use. In

applied behaviour analysis, there are several applications that have been designed to assist and teach people with developmental disabilities (e.g., Pro2loquo and Autism/DTT Numbers), or to help clinicians with assessment and data collection (e.g., VB-Mapp and Behavior Tracker Pro). There is also an article on how to develop iOS® applications to collect data on behaviours and their antecedents and consequences (Whiting & Dixon, 2012), information which could help to identify possible functions of problem behaviours. However, the literature does not show evaluations on the efficacy of most apps, including preference assessment apps.

In the ensuing sections of the Introduction, I provide an explanation of the importance of finding preferable items in behavioural interventions and a description of different preference assessment procedures. This is followed by a review of iOS® applications for conducting preference assessments that are currently available in the App Store® and a review of research on training procedures for preference assessments.

Stimulus Preference Assessments

Identifying preferred items in a population with developmental disabilities (DD) can be challenging, as the communication skills of these individuals may be limited (Ivancic & Bailey, 1996; Spevack, Wright, Yu, Walters, & Holborn, 2008; Spevack, Yu, Lee, & Martin, 2006). Stimulus preference assessment is a set of procedures that could help researchers and clinicians identify items and activities that children and adults with DD prefer. Those procedures are often used in behavioural interventions because preferred items often function as reinforcers, which can be used to promote desirable behaviour and learning (e.g., DeLeon, Frank, Gregory, & Allman, 2009; Hoffman, Samaha, Bloom, & Boyle, 2018; Kautz, DeBar, Vladescu, & Graff,

2018; Lee, Yu, Martin, & Martin, 2010; Paden, & Kodak, 2015; Toussaint, Kodak, & Vladescu, 2016). Preferred items can also be used to reduce challenging behaviours by promoting alternative behaviours. For example, Ahearn, Clark, DeBar, and Florentino (2005) evaluated the effects of presenting preferred items on the frequency of vocal and motor stereotypic behaviours that according to a functional analysis were automatically reinforced. The authors presented preferred items that provided similar and different sensory stimulation from the stereotyped behaviour, and found that preferred items decreased the frequency of the behaviour for both their participants, regardless of the type of sensory stimulation provided by the item.

Preference assessments can be divided in selection-based and engagement-based procedures. The three most common preference assessment procedures found in the literature are paired-stimulus (PS), multiple-stimulus-without-replacement (MSWO) and free-operant (FO) (Virués-Ortega et al., 2014). The PS and the MSWO procedures are selection-based and the FO procedure is engagement-based. Each procedure has different strengths and weaknesses. The selection of a procedure should be guided by the stimuli being assessed, the level of client functioning, motor skills, presence or absence of problem behaviours and the format of choices such as tangibles, pictures, verbal, and videos (Conyers et al., 2002; Hagopian, Long, & Rush, 2004; Higbee, Carr, & Harrison, 1999; Horrocks, & Higbee, 2008; Snyder, Higbee, & Dayton, 2012; Verriden & Roscoe, 2016; Virués-Ortega et al.).

PS procedure. For the PS procedure, after allowing the client to sample each item individually, the items are presented in pairs, in all possible combinations and in a randomized order. On each trial, the tester calls the client's attention to each item, which is placed at equal

distance to the client, and then asks the client to “pick one.” Once the client chooses an item (e.g., pointing to or touching an item), the tester removes the unselected item, records the selection, and allows the client to interact with the item for 30 s (or to consume it, if it is an edible). Attempts to reach for more than one item simultaneously are blocked gently by the tester and the trial is repeated. The degree of preference for each item is calculated by dividing the number of times an item was chosen by the number of trials that item was presented, and multiplying by 100. A preference hierarchy based on the degree of preference can be established among items.

The PS procedure was first proposed by Fisher et al. (1992). Compared to presenting the stimulus one at a time (single-stimulus procedure), Fisher et al. found that the PS procedure established a greater differentiation among the items assessed, and identified preferred items that were more powerful reinforcers. Findings by Fisher et al. have been replicated by others (e.g., Roscoe, Iwata, & Kahng, 1999). Moreover, Lee et al. (2010) have shown that the reinforcer values of different items, established a priori based on the rate of responding the items maintained, were correlated with preference values produced by a stimulus preference assessment using PS procedure.

MSWO procedure. For the MSWO procedure, after allowing the client to sample each item individually, usually six to eight items are displayed in a row in front of the client. The tester calls the client's attention to each item, and then asks the client to “pick one”. Immediately after a choice is made, the tester removes the unselected items from the table, and allows the client to consume the food item or interact with the leisure item for a brief period (e.g., 30 s).

The selected item on each trial is not presented in subsequent trials. Attempts to reach for two or more items simultaneously are blocked gently, and the trial is repeated. If the client does not make a choice within 30 s after being asked to do so, the tester records that as “no selection” and re-presents the trial once. The session ends if no item is selected again or until the last two items have been presented. The session is repeated a few times and the degree of preference for each item is reflected by the order in which each item was selected averaged across sessions. The MSWO procedure also yields a preference hierarchy among items being assessed. However, the client needs to be able to choose from an array of six to eight items in the MSWO procedure, rather than from only two items in the PS procedure.

The MSWO procedure was first proposed in 1996 by DeLeon and Iwata. These authors found that the procedure was effective in identifying items that could function as reinforcers and the procedure required less time than the PS procedure. Carr, Nicolson, and Higbee (2000) evaluated a shorter version of the MSWO procedure, where only three sessions were presented compared to five sessions that was originally proposed by DeLeon and Iwata. Carr et al. found that the highly preferred items identified with the brief MSWO procedure had reinforcer properties, and later studies replicated this finding (e.g., Higbee, Carr, & Harrison, 2000; Paramore & Higbee, 2005; Tullis, Cannella-Malone, & Fleming, 2012).

FO procedure. For the FO procedure, after allowing the participant to sample the items individually, all items are presented simultaneously to the client for the duration of the session. The tester calls the attention of the client to each item, and displays it in front of the client at approximately equal distance apart from each other and from the client. Next, the tester instructs

the client to interact with any item (e.g., “play with any of the toys on the table”). The client is allowed to interact with any item or with multiple items simultaneously. The tester records the time the client spent engaging with each item during the session, usually using an interval observation and recording method. For example, the tester records the presence of any interaction with each item during every 10-s interval of a session. The item with the highest percentage of intervals with interaction is designated as the most preferred (Roane, Vollmer, Ringdahl, & Marcus, 1998). The FO procedure is appropriate for clients who can engage with items. This procedure is superior to PS and MWSO in that it does not require retrieving items from the client before presenting each trial, which could be a problem for some clients (Kang et al., 2011). In addition, the duration of the session is pre-determined (typically 5 min). Hence, depending on the number of stimuli being presented, it can also be faster than other procedures. However, the FO procedure may not provide a clear hierarchy of preference among items being assessed. For example, if a client interacts with only their most preferred item during a session, the assessment yields no information about the relative preference of the other items. This procedure was first proposed by Roane et al. They showed that items ranked as preferred on the FO assessment were effective reinforcers, and that the client engaged with highly preferred items more frequently than less preferred items when both items were available. The authors also found that the FO procedure took less time to complete than the PS procedure, and that the client presented fewer problem behaviours. Other studies have replicated the findings for problem behaviours that were maintained by tangible reinforcers, but not for problem behaviour maintained by social attention (Kang et al., 2011; Tung, Donaldson, & Kahng, 2017).

Other preference assessment procedures. Less common preference assessment procedures are the single-stimulus, multiple-stimulus-*with*-replacement (MS), and response restriction. Single-stimulus (SS) preference assessment involves presenting each item individually for the same amount of time and recording the duration of engagement for each item (DeLeon, Iwata, Conners, & Wallace, 1999; Hagopian, Rush, Lewin, & Long, 2001) or the approaches (Pace, Ivancic, Edwards, Iwata, & Page, 1985; Fisher et al., 1992). However, this procedure can be more time consuming, or less effective than the PS procedure in producing a preference hierarchy since clients tend to approach all items (Fisher et al.). SS preference assessment may be an appropriate procedure for clients with profound impairments who may have difficulty discriminating between two items in a PS procedure (Hagopian et al., 2001). MS was similar to MSWO except that all stimuli are presented on every trial. A limitation of this procedure is that it often does not produce a preference hierarchy among assessed items because the client tends to always select their most preferred item on every trial (DeLeon & Iwata, 1996). The response restriction procedure is a variation of the FO procedure, which involves adding a series of rules to remove the availability of some items after they have been selected (Hanley, Iwata, Lindberg, & Conners, 2003). The authors have found that this procedure was effective in yielding more variability on the preference levels than the FO procedure. However, the assessment was more time consuming than the FO procedure.

Preference Assessment Applications for Mobile Devices

There are three preference assessment apps in the App Store®. One is named *Preference & Reinforcement Assessment* by Touch Autism, the second is named *myPref* by KV Adaptive

LLC, and the third is the application developed for this study, named *Prefer* (described later). In the sections below, I outline a list of desirable features of a preference assessment app, and review the features of *ABAPref* and *myPref* against this list (see Table 1).

Assessment procedures. For stimulus preference assessment procedures, it would be desirable for an app to incorporate at least the most commonly used procedures as described earlier (i.e., MSWO, PS, and FO). *Preference and Reinforcement Assessment* offers the options to conduct PS, MSWO, and MS. *MyPref* offers only the PS procedure (Table 1) and requires the user to always use six items for the assessment.

Aid to select the appropriate procedure. Given the variety of stimulus preference assessment procedures, it is important for testers to select the procedure that is most appropriate for the client based on their abilities to emit a selection response and to discriminate among the items being assessed (e.g., see review by Virués-Ortega et al., 2014). It would be highly desirable for an app to provide specific prompts to guide users to select the most appropriate procedure for a client. Neither app possesses this feature (Table 1).

Procedural prompts. It would also be desirable for an app to provide procedural prompts to the user at key points during the assessment to assist novice testers. *myPref* provides prompts for items (pictures) to be presented and placement of items on each trial. *Preference and Reinforcement Assessment* offers prompts for item sampling before the assessment, item placement on each trial, and to provide the selected item to the client (Table 1). It would be desirable if an app includes all of the above prompts plus providing a countdown prompt for item access after a selection.

Data logging capabilities. One of the advantages of a computer program is its ability to collect data on a number of potentially useful variables automatically, without extra human effort. It appears that neither of the existing apps has taken advantage of this capability. *Preference and Reinforcement Assessment* does not offer an option to back-up data, and *myPrefer* offers an option to e-mail the results. Beyond the primary dependent measure of percent preference reported by both apps, it would be desirable for a stimulus preference assessment app to record a number of other variables. For example, a record of the items that were presented on each trial, their locations on each trial, and the selection made by the client on each trial could be useful in determining whether there was a position bias on the selection (e.g., the client always chose the item on the left). Likewise, the latency of the selection response could be useful in determining preference among items that have been selected the same number of times, where shorter latency might indicate a stronger preference. Also, knowing the total assessment duration with a particular client can be useful to help plan future assessments. Moreover, it would be extremely desirable for an app to have the ability to back-up the data onto another computer for data protection, to maintain confidentiality in the case of shared iOS devices, and for data analysis across assessments or clients.

Evidence-based tutorial. Finally, neither app shown in Table 1 possesses any training for the user on how to use the app. It is not clear whether a novice tester could use the apps to conduct a stimulus preference assessment correctly.

In summary, the features of the above apps appear to be limited and there is much room for improvement. To address these limitations, I have designed an app for stimulus preference

assessments, named *Prefer*, and incorporated all of the desirable features suggested above. The application is available for free in the App Store®. The development of *Prefer* is described in more detail in the next section.

Development of *Prefer*

The iOS® application for conducting stimulus preference assessments, named *Prefer* (Julio, F. & University of Manitoba, 2016), was developed as part of a Canadian Institutes of Health Research-funded project on knowledge translation to address needs of special education teachers for managing challenging behaviours (Stoesz, McCombe, Martin, & Yu, 2016). The development of the application started in 2012 and was completed in 2015. Appendix A screenshots examples of *Prefer*'s main screen, prompts, recording screen and results.

The design of *Prefer* was initially done by consulting an iOS® developer from Brazil (Diogo Tridapalli). I sketched the content and the layout of all screens, and then I designed the screenshots of the application on OmniGraffle®. Design changes were made to accommodate programming. The computer code for *Prefer* was written by Paul Whyte from the Engineering Department of University of Manitoba. However, the app also contains codes developed by other professionals under open-source licensing. The app was developed on Xcode®.

The development of the application started with the selection-based procedures (i.e., PS, MSWO, and MS) as their codes share some similarities, followed by the engagement-based procedures (i.e., FO and SS). Later, I added a link in *Prefer* for a YouTube® playlist that consisted of video tutorials for the PS, MSWO, and FO procedures, and a tutorial on how to use *Prefer*. Also, a “Help Me Choose” feature was added to the app to guide users to select the most

appropriate procedure for their client from among the available procedures. This feature recommends the most appropriate procedure based on the user's responses to a series of yes-no questions about the client's engagement and selection abilities, the client's ability to choose among multiple items, the mode of choice presentation (e.g., object or picture), and type of stimuli to be assessed (e.g., food or leisure items). The procedure selection algorithm was based on the description provided by Virués-Ortega et al., 2014. Over the two years that the application was being developed, several versions were tested and revised.

Preference Assessment Training Research

Considering the cost of face-to-face individual or group instruction (i.e., instructor's time and learners' time), recent research has focused on several self-instructional strategies. Four general training strategies that do not require the presence of an instructor have been reported in the literature. One involved presenting the description of preference assessment procedures extracted from empirical articles published in the scientific literature. A second involved delivering self-instructional manuals describing preference assessment procedures (i.e., detailed instructions without technical jargons, with visual aids such as diagrams or pictures, examples, and review exercises). The third involved training modules with observing and rating a preference assessment procedure (Marano, Vladescu, Reeve, DiGennaro, & Reed, 2019). The fourth involved using a combination of video-modeling, voiceover or written instructions, task list, diagrams, and rehearsal.

Presenting the method section of a published article without further training with or without feedback has not been found to be an effective strategy to teach preference assessments

(Ramon et al., 2015; Roscoe, Fisher, Glover, & Volkert, 2006). However, there is evidence supporting the effectiveness of self-instructional manuals (Graff & Karsten, 2012b; Ramon et al.; Shapiro, Kazemi, Pogosjana, Rios, & Mendoza, 2016), training modules with observing and rating a preference assessment (Marano et al., 2019), and video training (Bovi, Vladescu, DeBar, Carroll, & Sarokoff, 2017; Deliperi, Vladescu, Reeve, Reeve, & DeBar, 2015; Hansard & Kazemi, 2018; Lipschultz, Vladescu, Reeve, Reeve & Dipsey, 2015; Miljkovic et al., 2015; Nottingham, Vladescu, Giannakakos, Schnell, Lipschultz, 2017; Rosales, Gongola, & Homlitas, 2015; Weldy, Rapp, & Capocasa, 2014). Research on self-instructional approach, training modules with observing and rating, and video training are reviewed below.

Self-instructional manuals. Ramon et al. (2015) evaluated a self-instructional manual for the MSWO procedure (Ramon & Yu, 2009). In this study, the authors compared the effects of a manual containing detailed instructions, checklists, examples, and review exercises with a method description adapted from published research articles on the same procedure. The study used an unbalanced crossover design, with random group assignment, where half the participants received the method description, and then the manual, and the other half received the manual, and then the method description. A multiple baseline across four participants was also used in each group. The study had nine participants in each group. For the participants who received the method section first, none mastered the procedure (i.e., 85% correct responses or above) with the method description alone, and after the manual was delivered, seven participants met the mastery criterion. The remaining two participants learned the procedure after watching a live demonstration. For the participants who received the manual first, four met the mastery criterion,

one reached the criteria after receiving the method section, and three after observing a live demonstration. Participants who achieved the learning criteria also showed generalization and retention of the learned skill. Graff and Karsten (2012b) studied the effects of a self-instructional package without feedback. The self-instructional package included enhanced instructions, pictures, diagrams, and examples. Participants were allowed to read the manual for a maximum of 30 min. The study design consisted of a multiple baseline across preference assessments methods. The authors monitored the percentage of trials implemented correctly during baseline, interventions, and generalization probes. During baseline, participants received adapted instructions for MSWO and PS procedures prior to each assessment with a confederate playing the role of a client. For the intervention, six participants were presented with the self-instructional package after the baseline, and five were also presented with the same instruction from baseline, in addition to data sheets. The results showed that the instruction or instruction with data sheets were insufficient to promote accurate performance (equal or above 90% of accuracy). After the enhanced instructions were provided, all participants met the criteria, including the performance on generalization probes with a client.

Shapiro et al. (2016) conducted two studies. First, they evaluated the effects of presenting the same self-instructional package from Graff and Karsten (2012b) or a version modified according to the participant's mistakes. The study used a multiple baseline design across participants. Participants were five behaviour technicians and eight undergraduate students (in psychology but without coursework in behaviour analysis). Unlike Graff and Karsten's study, only three of the five participants with behavioural analytical background met the mastery

criterion (equal or above 90% of accuracy) with only the self-instructional package. For the five undergraduate students who were exposed to only to a modified version of the self-instructional package, four achieved criterion after reading the manual, and one achieved it after also receiving the feedback and modeling session. For the three participants who received the manual followed by the modified version, two achieved criterion after receiving the modified version and one after receiving feedback and modeling. In the second experiment, Shapiro et al. (2016) evaluated the effectiveness of different feedback procedures, and found that task clarification (a list of the target response with their operational definition) was not effective for three participants, but that task clarification plus feedback on whether responses were performed correctly was effective.

In summary, it appears that self-instructional manuals alone were effective for teaching some to most participants to conduct MSWO assessment (Ramon et al., 2015; Graff & Karsten, 2012b; Shapiro et al., 2016). Participants who did not meet mastery after self-instruction alone met criterion with after feedback on errors (Shapiro et al.) and modeling (Ramon et al.).

Training modules with observing and rating. Marano et al. (2019) observed the effect of a computerized training program with 3 university graduates. The first module consisted of instructions to navigate the program. The second module consisted of 30-s videos showing each step of the PS procedure being performed correctly or incorrectly, where the participants were requested to answer if the behavior was being performed correctly or incorrectly. For the second module, participants received a feedback on the correct responses, or the instruction to try again on the incorrect responses. On the third module, participants were required to score novel clips

and rate multiple steps without receiving feedback on their answers. The training took a mean of 53 min (ranging from 48 to 59 min). All three participants achieved criterion (i.e., 100% correct responding) after completing the trainings, one participant performed immediately at criterion and the other two performed at criterion in the second session. All participants continued to perform above criterion with a client, or with more items on the assessment.

Video training. Miljkovic et al. (2015) evaluated the use of video modeling alone in a multiple probe design to teach six undergraduate university students with no prior experience with preference assessments to carry out the MSWO procedure. During the video modeling intervention, participants were asked to watch a video of the experimenter conducting an MSWO preference assessment with an actor. There were no instructions embedded in the video. The results of this study showed that video modeling alone, without additional instructions, produced large performance improvements for all participants, but it was insufficient to reach the mastery criterion of 85% accuracy. However, the addition of written instructions following video modeling resulted in criterion performance for all participants.

Bovi et al. (2017) evaluated the effects of presenting a video with voice-over instruction to teach the MSWO procedure in a multiple baseline across two participants. Participants were school staff with and without experience with children with ASD. The video modeled all components of the MSWO procedure with voice-over instruction describing each step, followed by a demonstration of the procedure without the voice-over. The video was 17 min and 13 s. Both participants achieved the mastery criterion immediately, which was defined as 90% correct responding on two probe sessions. However, the authors noted that participants watched the

video prior to each MSWO session (except in baseline), which might have influenced their performance during probe sessions. Another limitation of the study was that it was a multiple baseline across only two participants.

Weldy et al. (2014) developed and evaluated a video instructional program for MSWO and FO procedures. The training consisted of a 30-min video presentation for each assessment. The video presentation included instructions about each preference assessment procedure, video examples, and video modeling. The study used a non-concurrent multiple probe design across behaviours on different types of preference assessments. During baseline, the participants were asked to conduct a specific type of preference assessment. Then they were asked to watch the video in groups (i.e., with other participants). The mastery criterion was defined as performing correctly 90% or more of the items in a preference assessment task list. Seven participants met the mastery criterion after watching the presentation once. The remaining two participants achieved mastery after watching the video a second time.

Although the video training by Weldy et al. (2014) was effective, the authors pointed out several limitations. First, all participants had a background in behaviour analysis, and most of them had a bachelor degree in a related field. Therefore, further research is needed to determine whether the intervention would be effective with learners (e.g., caregivers) with no background in behaviour analysis. Second, the participants viewed the video in groups. It was possible that discussions or interactions among participants during or after training could have enhanced their performance. The authors also reported that the participants could have observed each other's sessions, which could have further improved performance. Third, all assessments were done with

a confederate whose responses during the assessment were not standardized across participants. This could have resulted in variability in the conditions under which each participant was tested.

Rosales et al. (2015) evaluated the effects of video training with embedded instructions for staff in a multiple baseline design across three participants for the PS, MSWO, and FO procedures. During the intervention, participants were permitted to take notes while watching the videos and use the notes when they performed the assessments. In addition, participants also received a sheet with the behaviours that were being scored (i.e., task list). All participants were first exposed to the PS video. The mastery criterion was 90% correct response in two mock assessments, and each assessment was preceded by watching the video. The three participants reached the mastery criterion after watching the PS video three, four and six times, respectively (the latter also received feedback on the performance). After learning the PS procedure, participants were exposed to MSWO video, and all three reached the mastery criterion immediately. Next, the participants were exposed to the FO video and two reached the learning criterion immediately and one participant, after watching the video three times. Retention was demonstrated for all preference assessment procedures.

Although Rosales et al.'s (2015) study showed promising results, the authors identified several limitations. First, all participants were exposed to the three preference assessments sequentially (i.e., PS, MSWO, and FO), and it is possible that learning the PS preference assessment could have impacted the learning of the MSWO procedure, and that learning the PS and MSWO procedures could have impacted the learning of the FO procedure. Second, the participants were not required to carry out the full preference assessments while being evaluated.

Therefore, the performance might not be representative of what might occur in a full assessment. Third, the confederate's responses were not standardized across participants. Fourth, the video training did not include skills for calculating and interpreting results, which is an important skill for practitioners. In addition, all participants had some experience working with individuals with autism. Further research with less experienced individuals (e.g., new staff members) would be desirable.

Lipschultz et al. (2015) evaluated the effectiveness of video modeling with voiceover instruction to simultaneously train staff to conduct MSWO, PS, and SS preference assessments. The authors also included the selection of a preference assessment procedure, and items to be used in the assessments as part of the training. The study used a concurrent multiple baseline design across four participants. All participants achieved the learning criteria (90% correct response in two assessments for all assessments procedures) after watching the videos between two and six times, one participant also received feedback before achieving the criterion. All participants maintained their performance above the mastery criterion during generalization with a client and during retention. The study showed that it was possible to train staff, with experience working with individuals with autism and developmental disabilities, to conduct three procedures simultaneously. However, some procedures required the participants to watch the video with the three procedures up to 6 times to achieve the criterion, and one of these participants also received feedback.

Hansard and Kazemi (2018) trained four undergraduate students to conduct the PS procedure. The training consisted of a video with voice-over and written instructions, video

modeling, and instruction to rehearse the whole assessment. The written instructions were the same as those presented by Graff and Karsten (2012b). The video was divided in setting up, implementation, and scoring. During the set-up phase, the video presented written instructions with voice-over, and then a picture of the assessment setting. The implementation session included vocal and written instructions, modeling, and a written request to practice each target response. The study used a non-concurrent multiple baseline across participants design, and all participants learned to perform the PS procedure (90% correct or more across two consecutive sessions), after watching the video once. One limitation of the study that the authors reported was that there was no assessment on whether this performance would generalize to clients. A second limitation was that that all participants were recruited from an upper division psychology class, and the finding could be different among population with a different educational background. In addition, although the confederate's responses in the mock assessment were scripted, there was no measure of whether the confederate adhered to the scripted responses during the simulated preference assessment.

Arnal Wishnowski, Yu, Pear, Chand, and Saltel (2018) also evaluated the self-instructional manual for MSWO procedure (Ramon & Yu, 2009), but the manual was delivered online using computer-aided personal system of instruction (Pear & Kinsner, 1988) and videos were added to the training program. The intervention was evaluated using a multiple probe design across six undergraduate students and four staff members. All participants showed immediate and substantial improvement after training, with all six students and two of the four staff members meeting the mastery criterion (85% correct or higher). A limitation of this study

was that generalization assessment was simulated. Other limitations were that social validity questionnaire was not administered for all participants, and that the study addressed only one preference assessment procedure.

In summary, only one study (Miljkovic et al., 2015) evaluated the effects of video modeling without any additional instructions. Participants in that study showed clear improvement after viewing the video although none met the mastery criterion. Adding instructions to the video resulted in all participants achieving the mastery criterion. In other studies, video plus instructions have been found to be quite effective in teaching behaviour analysts, staff, or undergraduate students to conduct PS, MSWO, or FO procedures (Arnal Wishnowski et al., 2018; Bovi et al., 2017; Hansard & Kazemi, 2018; Lipschultz et al., 2015; Rosales et al., 2015; Weldy et al., 2014).

Development of Video Training for *Prefer*

Considering the effectiveness of video training, its efficiency, and the ease with which it can be delivered in portable devices, I have developed self-instructional video tutorials to accompany *Prefer* to teach users to conduct stimulus preference assessments with the app. The videos showed two actors (volunteer graduate students), playing the roles of client and tester, demonstrating all aspects of each procedure (e.g., stimulus sampling, stimulus presentation, providing consequence for different client responses, recording responses, etc.). I used two recording devices, each focusing on the behaviour of one actor (i.e., tester or client). I used Final Cut Pro® to merge and edit the videos. The video included written instructions and diagrams, modeling, and picture-in-picture videos of the application. The PS video was pilot tested with a

few individuals not included in the study. Based on the common mistakes observed during pilot testing of the PS procedure, I added new written instructions to the video on how to perform these behaviours, requests for the participants to rehearse those behaviours, and a reminder at the end of the video of the common mistakes when implementing the procedure. The links for all three videos and screen shots of the MSWO procedure are provided in Appendix B.

Clinical Utility of Stimulus Preference Assessments

Despite ample evidence on the effectiveness and benefits of conducting preference assessments (e.g., Tullis et al., 2011; Weeden & Poling, 2011), uptake of the procedures has been slow. In an online survey of 406 participants who worked with individuals with autism and other special needs (Graff & Karsten, 2012a), although all participants reported using some form of indirect preference assessment (i.e., observing the client, asking caregivers or the individual), fewer than 60% of the respondents reported knowing the term “stimulus preference assessment”. The authors also found that only 49.5% of the respondents had received training on preference assessment procedures. For those who had received training, 13.5% were certified behaviour analysts and 66% were educators and practitioners. Approximately 60% of all respondents, including certified behaviour analysts, identified “lack of time” and approximately 50% of all respondents identified “lack of knowledge” as the main barriers to conducting stimulus preference assessments. These results suggest that research is much needed to develop both time-efficient and cost-effective training procedures to promote uptake of stimulus preference assessment.

Smart (2006) noted that the term “clinical utility” is typically used in health care, but

there is a lack of formal definition for this concept. He proposed that the term could be conceptualized as “a multi-dimensional judgment about the usefulness, benefits, and drawbacks of an intervention” (p. 379), which could ultimately indicate whether clients and practitioners would or would not use the intervention. In applied behaviour analysis, the concept of “social validity” has been proposed to examine clinical and social significance or benefits of an intervention (Kazdin, 1977; Wolf, 1978). For example, Kazdin and Wolf suggested that social validity may be assessed by measuring (a) the acceptability or importance of the behaviour being targeted for intervention; (b) the acceptability of the intervention procedure used to promote the behaviour change; and (c) the clinical significance or importance of the behaviour change for the client. Schwartz and Baer (1991) noted that the purpose of social validity is to evaluate the acceptability or viability of a programmed intervention. Therefore, social validity refers to the concern of behavioural researchers to find strategies and procedures that are not only effective, but also acceptable and useful for clients and practitioners, who will likely continue to use the procedures beyond the research evaluation. Although some research on teaching individuals to conduct stimulus preference assessments has included acceptability ratings of the training method by research participants (e.g., Arnal Wishnowski et al., 2018; Ramon et al., 2015), no study has examined the social validity of the stimulus preference assessment procedures based on the training and implementation of stimulus preference assessment procedures by practitioners in their daily jobs.

Rationale for the Present Studies

There is a dearth of iOS® apps that have been developed for stimulus preference

assessment and the features on the available app are limited (Table 1). *Prefer* was developed to address the limitations of previous apps and to promote the uptake of preference assessments by making it freely available in the App Store®. Specifically, *Prefer* includes five preference assessment procedures (MSWO, PS, FO, MS and SS), and its features (see Table 1) could help address the barriers to preference assessment identified by Graff and Karsten, 2012, such as the lack of time and training. For example, the app includes: (a) a built-in guide to help the user to select an appropriate preference assessment procedure; (b) picture prompts for item location; (c) a count-down timer for item access after a selection; (d) data logging with various information about the assessment (e.g., item location, latency, duration of the assessment); (e) a number of customization to each preference assessment (e.g., decreasing or increasing access time, session duration, and selection of up to 9 items); and (f) automatically calculating and graphing of percent preference for each item assessed.

An important limitation of currently available apps is the lack of research evidence to support their effectiveness. The present studies were conducted to address this shortcoming. In Study 1, I evaluated the effectiveness of the video tutorials to teach individuals with varied backgrounds and no experience in stimulus preference assessment to use the app to conduct PS, MSWO, and FO procedures. In study 2, I examined the social validity of *Prefer* with practitioners and clients in a clinical setting.

STUDY 1: VIDEO TRAINING FOR PREFERENCE ASSESSMENTS

The purpose of Study 1 was to evaluate the effectiveness of using video tutorials to teach individuals to use *Prefer* to conduct preference assessments with PS, MSWO, or FO procedures.

The procedural fidelity of conducting each procedure was evaluated in a concurrent multiple probe design across participants. The performance of each participant was observed during baseline, after training, and during generalization.

Method

Participants and Settings

Fifteen participants were recruited for Study 1. Participants included 10 university students and 1 university staff who had no or limited experience with behaviour analytic procedures. The other four participants were staff members who worked with individuals with DD. University participants were recruited from the University of Manitoba, through recruitment posters distributed throughout the campus. Staff participants were recruited through a community agency that served individuals with DD. All participants reported that they have not received prior training on any stimulus preference assessment procedure, and all were familiar with using a smartphone. Table 2 shows the age, gender, education, and experience of the participants with people with DD. The mean age was 25 years old (range, 18 to 47). Five participants were male and 10 were female. Educational experience ranged from first year university to completion of a Bachelor's degree to graduate studies. Each participant received an honorarium of \$60 regardless of their performance.

In addition, five children with a diagnosis of autism, according to their health records, were recruited to participate in the generalization test. Client participants were recruited from a community agency serving individuals with DD. The parents reported that all children were able to choose among two or more items, or engage with items. There was no demographic data

collected. Items used for client preference assessments were suggested by their parents. Each child received a toy, valued at approximately \$10-\$15, for their participation.

All training sessions and simulated assessments (described later) for student participants were conducted in a quiet session room at University of Manitoba, and all sessions for staff participants were conducted in a session room at St. Amant Research Centre. Client generalization assessments took place in a private room in the client's home.

Ethical approval was received from the University of Manitoba Psychology/Sociology Research Ethics Board and from St. Amant before the study began. Written informed consents were obtained from all participants and from the parents of the clients before participation started.

Materials

During all phases of the study, each participant was provided with an iPhone® with *Prefer* preconfigured for the assessment and six leisure items to be used during simulated assessments (e.g., car, tambourine, stuffed animal, etc.). A digital video recorder was used to record all sessions (with consent) for procedural and reliability checks.

Design

A concurrent multiple probe design (Cooper, Heron, & Heward, 2007) across participants was used for each preference assessment procedure. This design is a variation of the multiple baseline design, in which observations of the target behaviour are conducted continuously or frequently (Kazdin, 2011). In the multiple probe design, however, observations are conducted intermittently. This probe technique is applicable to behaviours that are expected to be stable

(Horner & Baer, 1978). Data evaluation in the multiple probe design relies on the same logic as a multiple baseline design and uses the same guidelines for visual inspection of single-case experimental designs (Kazdin; Martin & Pear, 2015).

The PS procedure was evaluated first with five participants (P1 through P5), the MWO procedure was evaluated next with P6 through P10, and the FO procedure was evaluated last with P11 through P15. The order of the three procedures was arbitrary and determined prior to participant recruitment.

Video Tutorials

A video was prepared for each procedure. The video consisted of diagrams showing the position of the items in relation to the tester, written instructions on how to conduct the preference assessment, actors demonstrating the tester's behaviours in carrying out various procedural components (e.g., item sampling, stimulus presentation, and providing appropriate consequences on each trial depending on the client's response). The length of the videos was 12.4 min for the PS procedure, 13.5 min for the MSWO procedure, and 10.6 min for the FO procedure. Links to all three videos and screen shots of the MSWO video are provided in Appendix B.

Behaviour Measures

The dependent variable was the participants' performance accuracy during stimulated assessments (described later), defined as the percentage of correctly implemented steps for each procedure. Participants' performance during each stimulated assessment was observed either live (75.9%) or from recorded videos (24.1%). All generalization assessment sessions with clients

were observed live.

The PS procedure consisted of presenting 6 items in pairs on each trial (15 trials per session). The participant's behaviours during each session were scored using a checklist as shown in Appendix C. The initial sampling of the items and session set up were scored once at the beginning of the session. On each trial, 6 to 11 target behaviours could be scored depending on the client's response. Lastly, the participant was asked to identify the preferred item(s) based on the results of the assessment at the end of the session. The percentage of performance accuracy for a session was calculated by dividing the number of correct responses by the sum of correct and incorrect responses, and multiplying this ratio by 100.

The MSWO procedure consisted of presenting an array of 6 items until all items were selected (or no choice was made) and repeating this presentation 3 times. The participant's behaviours during each session were scored using a checklist as shown in Appendix D. The initial sampling of the items and session set up were scored once at the beginning of the session. On each trial, 8 to 10 target behaviours could be scored depending on the client's response. Lastly, the participant was asked to identify the preferred item(s) based on the results of the assessment at the end of the session. The percentage of performance accuracy for a session was calculated by dividing the number of correct responses by the sum of correct and incorrect responses, and multiplying this ratio by 100.

The FO procedure consisted of allowing the client to engage with 6 items for 5 minutes, while all items were available to the client. The participant's behaviours during each session were scored using a checklist as shown in Appendix E. The initial sampling of the items and

session set up were scored once at the beginning of the session. The session was divided into 30 10-s intervals and each interval was scored according to the following rules. First, a participant could obtain 1 point for correctly recording the number of items that the client had interacted with. A fractional score would be given, if the participant was partially correct, by dividing the number of correctly recorded items by the total number of items the client had interacted with. For example, if the participant recorded interaction with one item but the client had interacted with two items, the participant would receive 0.5 point (1 divided by 2). Second, a participant could obtain 1 point for correctly maintaining all 6 items accessible to the client. A fractional score would be given, if the participant was partially correct, by dividing the number of items that were accessible to the client by the total number of items. For example, if only five items were accessible (one fell to the floor and was not returned to the table), the participant would receive 0.83 points (5 divided by 6). Lastly, a participant could lose 1 point if a prompt was given for the client to interact with an item. Therefore, a participant could receive a maximum of 2 points for an interval and a minimum of -1, although no participant had scored -1 in an interval in all sessions. The percentage of performance accuracy for a session was calculated by dividing the sum of scores for all 30 intervals plus 5 possible points (behaviours) for item sampling, session set up, and results interpretation by the maximum possible score of 65 (60 points for the 30 intervals and 5 points for correct item sampling, set up, and results).

Baseline Simulated Assessment

Simulated assessments using *Prefer* were conducted during baseline and after video training. A simulated preference assessment involved asking a participant to assess the

preference of a confederate (a graduate student), who played the role of a client. At the beginning of a simulated assessment, I gave the following oral instructions to the participant to conduct a preference assessment using one of the three procedures (instructions adapted from Ramon et al., 2015):

“Thank you for your help with my study. I would like you to find out what [*confederate’s name*] prefers using the procedure called [*Multiple-Stimulus-without-Replacement or Paired-Stimulus or Free-Operant*]. [*Confederate’s name*] will be playing the role of a client with intellectual or developmental disabilities who can hear but cannot speak. I cannot give you any information about the procedure, please do your best to find out what items [*confederate’s name*] likes or does not like. You may use this iPhone and the application that’s already set up to conduct the assessment. You can use the items in the box [*point to the box*] for the assessment. Let me know when you are ready. When you are ready to start you can touch “start” here [*show participant where in the application*].

During a simulated assessment, the confederate’s responses were scripted. For the PS and MSWO procedures, scripts consisted of how to make a selection (i.e., choose one item, choose two items simultaneously or successively, or reject one item and choose a second item) or making no selection. The scripts also included whether the confederate should or should not play with the item selected. For the FO procedure, the script consisted of which item(s) to interact with within each 10-s interval, or not to engage with any item. “The script also included whether

the confederate should drop an item on the floor. There was a total of nine scripts for each assessment procedure. The order of the scripts was randomized across participants. The scripts ensured that different aspects of the procedure were assessed and that each participant encountered the same level of complexity in the confederate's responses throughout the study.

Video Training

Following baseline simulated assessment, each participant was asked to watch the video. Participants were allowed to watch the video as many times as they wished and were asked to inform the researcher when they have finished. The total amount of time each participant spent watching the video was recorded.

Post-Video Training Simulated Assessment

Each participant was asked to conduct a simulated assessment immediately following the completion of the video training. The procedures were the same as the simulated assessment conducted during baseline, although different scripts for the confederate's responses were used. The learning criterion was defined as 85% correct responses in one session. This criterion was the same used in previous research (Arnal Wishnowski et al., 2018; Miljkovic et al., 2015; Ramon et al., 2015).

Feedback Questionnaire

Each participant completed an 8-item survey after completing the post-video simulated assessment. This survey was adapted from social validity questionnaires used in previous studies (Arnal Wishnowski et al., 2018; Ramon et al., 2015). Participants were asked to rate using a 5-point scale the importance of the goal of the study, whether the video was easy to follow, their

perceived confidence in applying the procedure, and the helpfulness of the *Prefer* app (Appendix F). In addition, participants were asked to report the number of times they have watched the video or parts of the video. The feedback questionnaire was applied immediately after the post video-training simulated assessment.

Generalization Assessment with a Client

Similar to previous research (e.g., Arnal Wishnowski et al., 2018; Miljkovic et al., 2015; Ramon et al., 2015), if a participant reached the mastery criterion (85% or higher) on a post-video simulated assessment. I invited them to conduct a preference assessment with a client with DD. Since the time between the training and generalization was not under control of the experimenter, and there was the potential for a great variability that could affect results on the maintenance of skills across participants. I also invited the participants to watch the video training one more time before applying the procedure to the client, therefore all participants were exposed to the procedure shortly before the generalization assessment.

The assessment was done in the client's home. The items used in the preference assessment were customized to each client based on the caregiver's suggestions. Therefore, the items used during generalization assessment were different from those used during training.

Simulated Generalization Assessment with a New Confederate

For two participants (Participants 14 and 15), there was no client available for generalization assessment. Therefore, these participants conducted a simulated generalization assessment with a confederate instead of a client. However, the generalization assessment involved a new confederate, a new script, and different items.

Comparison between Participant and Experimenter

On the day of each client generalization assessment conducted by a participant, the experimenter also conducted a direct stimulus preference assessment with the same client using the same procedure using paper and pencil (instead of *Prefer*). The two assessments were separated by at least 15 minutes, and the participant's assessment time was recorded.

The researcher's assessment was done before the participant's assessment for new clients or new procedures, and after the participant's assessment for clients who had already participated in the study. Therefore, allowing the researcher to confirm that the client could perform the assessment procedure, and evaluating if the client would engage in challenging behaviours by the frequent removal of items (for the PS and MSWO procedures) or tasks demands. Participants 1, 5, 8, 10, 12 and 13 did the assessment after the researcher, and Participants 2, 4, 8 and 10 did the assessment before the researcher.

Interobserver Agreement on Participant's Performance

An observer watched all simulated and generalization assessments and scored all sessions independently, for all participants. The observer recorded the participants responses live for 18.5% of the assessments, and from videos for 81.5% of the assessments. An agreement was defined as the experimenter and observer recording a given checklist item as correct or as incorrect on a particular trial for PS and MSWO procedures, or during a 10-s interval for the FO procedure. Percentage agreement was calculated by dividing the number of agreements by the sum of agreements plus disagreements, and multiplying this ratio by 100. The mean agreement scores were 95.6% (range, 84.1% to 100%) for the PS procedure, 96.7% (range, 87.1% to 100%)

for the MSWO procedure, and 92.8% (range, 88.5% to 100%) for the FO procedure.

Confederate's Adherence to the Script

The confederate's adherence to the script during all simulated assessments was evaluated. A percent adherence score was computed for each session by dividing the number of responses that were performed as scripted by the total number of applicable scripted responses and multiplying this ratio by 100. The mean adherence scores were 97.5% (range, 84.2% to 100%) for the PS procedure, 98.5% (range, 83.3% to 100%) for the MSWO procedure, and 98.9% (range, 90.3% to 100%) for the FO procedure.

Results

Paired-Stimulus Procedure

Figure 1 shows the percentage of correct responses during simulated assessments conducted before training (baseline), after training (post-video), and during generalization assessment with a client for the PS procedure for student Participants 1 through 3. Participant 1's correct response during baseline was 31.0%, and Participants 2 and 3 had no correct response during baseline. After watching the PS video once, all three participants met the mastery criterion, achieving a score of 89.9%, 85.9% and 92.2% correct at post-video, respectively. Participant 1 performed at 99.3% correct during the generalization assessment, conducted 18 days after training, and Participant 2 performed at 83.3%, just below the mastery criterion, conducted 12 days after training. Participant 2's performance was affected mainly by one type of error (not presenting the items in the locations specified by the app). She would have achieved a score of 95.6% if that error had not occurred. Participant 3 was not available to return to conduct

the generalization assessment.

Figure 2 shows the percentage of correct responses during simulated assessments conducted before training (baseline), after training (post-video), and during generalization assessment with a client for the PS procedure for staff Participants 4 and 5. Participant 4's score during baseline was 34.2% correct. Participant 5's mean score across baseline sessions was 64.8% correct, and the baseline performance appeared stable across the two sessions. After watching the video, both participants exceeded the mastery criterion during the simulated assessments, with scores of 95.1% and 98.5% correct, respectively. Participant 4 performed the preference assessment at 91.8% correct during the generalization assessment, conducted 24 days after training. Participant 5 performed the preference assessment at 91.9% during the generalization assessment, conducted 19 days after training.

Multiple-Stimulus-without-Replacement Procedure

Figure 3 shows the percentage of correct responses during simulated assessments conducted before training (baseline), after training (post-video), and during generalization assessment with a client for the MSWO procedure for Participants 6 through 8 (university staff and two students). During baseline, Participant 6's correct response was 10.3%, Participant 7's mean correct response was 0.5%, and Participant 8's mean correct response was 46.5%. Baseline performance appeared stable for Participants 7 and 8. After watching the MSWO video once, Participants 6 and 8 exceeded the mastery criterion during their simulated assessments, performing at 98.3% and 96.0%, respectively. Participant 7's score remained at baseline level for two sessions after the intervention. During these two sessions, the participant was

manipulating the toys herself, rather than applying the procedure to the confederate. Since English was her second language, a written translation of the oral instructions delivered at the beginning of the session was provided for this participant in her native language (Chinese). The instructions were translated by a psychology graduate student whose first-language was Chinese. At the beginning of the third session, the researcher delivered the written translation of the instructions for the participant, and allowed her to read it at her own pace and keep it. Once the participant indicated that she had finished reading, the researcher provided a verbal instruction (in English) to assess confederate's preference and requested the participant to start the assessment when ready. There was an immediate increase on the accuracy of the procedure implementation to 67.4%, which was still below the mastery criteria. After the third session, the participant received a task list of the behaviours being scored, verbal feedback from the experimenter (in English) on the errors observed in the previous session, and a demonstration of the correct behaviour. The demonstration included: (1) presenting the items in a straight line and in the order shown in the application; (2) removing the items after the client had selected one item; (3) retrieving or blocking selection of a second item, and (4) praising the client for selecting one item. The demonstration was presented twice at the participant's request. After the feedback and demonstration, Participant 7 met the mastery criteria, with a score of 99.2% correct.

Participants 7 and 8 were available for the generalization session. Participant 7 performed at 80.8% correct during the generalization assessment with a client, conducted 23 days after training. However, the session had to end early due to the client's aggressive behaviours.

Participant 7 correctly presented the items for the client to sample and completed 3 trials within that session. She made 5 errors out of a total of 26 scored behaviours. The errors included not praising the client after a selection (twice), not retrieving an item from the table that was not selected (twice), and not retrieving an item from the client at the end of the trial (once).

Participant 8 performed at 100% correct during the generalization session, which was conducted 47 days after training. Participant 6 was not available to conduct the generalization assessment.

Figure 4 shows the percentage of correct responses during simulated assessments conducted before training (baseline), after training (post-video), and during generalization assessment with a client for the MSWO procedure for staff Participants 9 and 10. Participant 9's baseline score was 63.3% correct and Participant 10's mean score during baseline was 69.8% correct. After watching the MSWO video once, both participants exceeded the mastery criterion, with scores of 92.9% and 97.0% correct, respectively. Participant 9 performed at 93.6% correct during generalization assessment, conducted 10 days after training, and Participant 10 performed at 97.7% correct during the generalization assessment, conducted 26 days after training.

Free-Operant Procedure

Figure 5 shows the percentage of correct responses during simulated assessments conducted before training (baseline), after training (post-video), and during generalization assessment with a client for the FO procedure for student Participants 11 through 13. Baseline scores were high for all participants. Participant 11's baseline score was 79.8 % correct. Participant 12's baseline performance showed an increasing trend with a mean of 77.3% correct across the two sessions. Participant 13's baseline mean scores were 84.3% across sessions. After

watching the FO video once, all three participants showed further improvement and exceeded the mastery criterion, with scores of 98.6%, 96.0%, and 90.3% correct, respectively. Participant 11 was not available for the generalization session. Participant 12 performed at 92.1% correct during the generalization assessment, conducted 31 days after training. Participant 13 performed at 93.0% during the generalization assessment, and it was also conducted 31 days after training.

Figure 6 shows the percentage of correct responses during simulated assessments conducted before training (baseline), after training (post-video), and during generalization assessment with a confederate for the FO procedure for staff Participants 14 and 15. Participant 14's baseline score was 64.5% correct. Participant 15's baseline scores showed an increasing trend across sessions, with a mean of 57.4%. After watching the FO video once both participants exceeded the mastery criterion with 91.4% and 99.5% correct, respectively. Participant 14 performed at 85.3% correct during the generalization assessment with a confederate, conducted 92 days after training. The errors from Participant 14 during the generalization assessment were related to touching "Start" after the assessment began, and not stopping the recording after the confederate stopped engaging with an item. Participant 15 performed 98.5% correct during generalization assessment with a confederate, conducted 105 days after training.

Feedback Questionnaire

Table 3 summarizes the results from the Feedback Questionnaire. When asked to rate the importance for individuals who work with people with intellectual or developmental disabilities to learn how to conduct preference assessment, all participants rated it as important or very important ($M = 4.6$, range 4 to 5). When asked to rate how easy or difficult it was to understand

the video, most participants found the video easy or very easy to understand ($M = 4.6$, range 3 to 5). When asked whether they found any part of the video confusing, three participants identified the following areas as confusing or difficult: (1) learning how to respond when the client has rejected both items for the PS procedure, (2) understanding what to do when a session has ended for the MSWO procedure, and (3) trouble viewing the tutorial video. Overall, participants felt confident about applying the preference assessment procedure ($M = 4.1$, range 3 to 5). Most participants found the app helpful or very helpful ($M = 4.5$, range 3 to 5). Two participants reported watching the full MSWO video more than once, and seven participants reported to have re-watched parts of the training video for the three procedures (4 for PS, 2 for MSWO, and 1 for FO). Parts that were re-watched included how to respond when the client behaves in an unexpected way for the PS procedure, demonstration of the MSWO procedure, sampling the items, and some of the readings for the FO procedure. Nine participants reported to have practiced presenting the items while watching the videos.

Comparison between Participant and Experimenter.

PS procedure. Figure 7 shows the results of the PS procedure during the generalization assessments for Participants 1, 2, 4, and 5 (using the *Prefer* app), and results obtained by the researcher for the same clients on the same day using paper-and-pencil. Participant 1 conducted the assessment with Client 1. The researcher's assessment showed that Tambourine (ranked 1) as the most preferred and the Turtle (ranked 6) as the least preferred. Participant 1's assessment showed the Tambourine tied with Car (mean ranked 1.5) as the most preferred and Turtle ranked 5 (Board was the least preferred) (see top left graph in Figure 7).

Participant 2 completed the assessment also with Client 1 (on a different day after Participant 1). The researcher's assessment showed that the Board (ranked 1) was the most preferred, and Alligator and Car (mean ranked 5.5) tied as the least preferred. The participant's assessment showed that Board was ranked 2 and Alligator tied with Turtle (mean ranked 5.5) as the least preferred. However, Car, ranked 5.5 by the researcher's assessment, was the most preferred (ranked 1) based on the participant's results (see top right graph in Figure 7).

Participant 4 completed the assessment with Client 2. The researcher's assessment showed that the Hexbug (ranked 1) was the most preferred and Drawing Board and Alphabet Board tied (mean ranked 5.5) as the least preferred. Participant 4's assessment showed that Hexbug (mean rank 2) was tied with two other items as the most preferred, and Alphabet Board was the least preferred (ranked 6). Note that Drawing Board, ranked 5.5 by the researcher, was one of the most preferred items (mean rank 2) based on the participant's results (see bottom left graph in Figure 7).

Participant 5 completed the assessment also with Client 2 (on a day different from Participant 4). The researcher's assessment showed that the Hexbug was the most preferred item (ranked 1) and Drawing Board was the least preferred (ranked 6). The participant's assessment showed that Hexbug was also the most preferred item (ranked 1) and Drawing Board ranked 5 (Lego was the least preferred) (see bottom right graph in Figure 7).

MSWO procedure. Figure 8 shows the results of the MSWO procedure during the generalization assessments for Participants 8, 9, and 10 (using the *Prefer* app), and results obtained by the researcher for the same clients on the same day using paper-and-pencil.

Participants 8 completed the assessment with Client 3. The researcher's assessment showed that Car was the most preferred item (ranked 1) and Slinky was the least preferred (ranked 6). The participant's assessment showed that Car was ranked 4 and Slinky was ranked 1 (see top left graph in Figure 8). However, difference between the most and least preferred items was small in both the researcher and participant's assessments, suggesting that there was no clear preferred item for this client.

Participant 9 completed the assessment with Client 4. Both the researcher and participant's assessments showed that McQueen was the most preferred (ranked 1) and Clear Car was the least preferred (ranked 6) (see bottom left graph in Figure 8). Participant 10 completed the assessment also with Client 4. The researcher's assessment showed that McQueen was the most preferred item (ranked 1) and Board was the least preferred (ranked 6). The participant's assessment showed that McQueen was also the most preferred item (ranked 1) and Board was ranked 3 (Alligator was the least preferred) (see right graph in Figure 8).

FO procedure. Figure 9 shows the results of the FO procedure during the generalization assessments for Participants 12 and 13 (using the *Prefer* app), and results obtained by the researcher for the same clients on the same day using paper-and-pencil. For Participant 12, the researcher's assessment showed that Bear was the most preferred item (ranked 1) and preference for all other items was quite low. The participant's assessment showed similar results with Bear being the most preferred item (ranked 1) and preference for all other items was low (see left graph in Figure 9).

Participant 13 completed the assessment with Client 5. The researcher's assessment

showed that Bear was the most preferred item (ranked 1), but only at 36.7 %, and preference for all other items was low. The participant's assessment showed similar results with Bear being the most preferred item (ranked 1), but only at 16.0%, and preference for all other items was low (see right graph in Figure 9).

Time Spent in Generalization Assessment with a Client

For the PS and the MSWO procedures, the time spent in the assessment by the researcher (with paper and pencil) and by the participants (with the iOS application) were overall similar. The average time per PS assessment conducted by participants and researchers were both 21 minutes. The average time per MSWO assessment conducted by participants and researchers were 24 and 22 minutes, respectively. The average time per FO assessment, conducted by participants and researcher, were 9 and 10 min, respectively.

Discussion

Fourteen out of 15 participants learned to conduct the PS, MSWO and FO preference assessments using video. Except for P12 and P15 in the FO procedures, baseline performance was relatively stable for those with more than one probe. Large and immediate performance improvement was observed following video training and this behaviour change was replicated across participants. For all participants who were available for the generalization phase, all except P2 and P7 performed at or above the mastery criterion. Hence, overall the results are similar to previous research involving video-training to conduct preference assessments with paper and pencil (Bovi et. al., 2017; Deliperi et. al., 2015; Hansard & Kazemi, 2018; Lipschultz et. al., 2015; Miljkovic et al., 2015; Nottingham et. al., 2017; Rosales et. al., 2015; Weldy et.al,

2014).

Baseline performance for the four staff participants were higher than the university participants, although all participants reported that they have not received prior training in preference assessment procedures. The higher baseline scores for staff could have been a result of generalization from their experience working with clients with DD in other contexts. Previous research did not replicate the experiment across different populations. Hence, more research is needed to determine the generality of this finding.

Overall, participants responded positively on the feedback questionnaire, with mean ratings above 4 on each item. Seven of the 15 participants reported to have re-watched a portion of the video. The portions that were re-watched most were related to how to respond when the client did not make a straightforward selection. Concerning agreements between the researcher and participant's PS assessments on the most preferred item, they ranged from exact agreement (P5) to a difference of 0.5 rank (P1) and 1 rank (P2 and P4). For the agreement comparisons in MSWO assessments on the most preferred item, they ranged from exact agreement (P9 and P10) to a difference of 3 ranks (P8). For the agreement comparisons in FO assessments on the most preferred item, both were exact agreements (P12 and P13). Overall, agreement on the least preferred item or ranking of other items was moderate to low. Times spent in conducting the assessments by the researcher and by the participants were comparable.

The methodology of Study 1 has several strengths. First, carryover effects and observational learning could not be ruled out in some previous research because participants learned to perform multiple procedures in succession or in groups. Both were avoided in the

present study since participants were taught individually to perform only one preference assessment procedure. Second, all responses for the confederates during simulated assessments were scripted to ensure that the test conditions were similar across phases and participants. Moreover, scripted simulated assessments ensured that all aspects of the procedures were evaluated. Third, generalization of learned skills to clients with DD was evaluated. Fourth, interobserver agreement on behavioural measures and procedural fidelity on confederates' script adherence were evaluated and found to be high. Fifth, participants included individuals without experience working with people with DD. Overall, the replication of skill acquisition across participants within each procedure and across procedures provided strong support for the effectiveness of video training and objective behavioural measures were corroborated by participants' subjective reports on the feedback questionnaire for each procedure.

Several limitations should be noted. First, the participants' baseline performance for the FO procedure was high for all five participants (P11 through P15) and two (P12 and P15) showed an increase in baseline. This could have been due to the fact that the FO procedure is simpler than the PS and MSWO procedures and the prompts embedded in *Prefer* were sufficient to yield more correct responses during baseline compared to the other procedures. Second, when comparing the participant's and researcher's preference assessment for the same client, although there was moderately strong agreement on the most preferred items, agreement on the less preferred items was weak or poor, particularly for Clients 3 and 5. It should be noted that these two clients did not show a strong preference even for their most preferred items and this could have contributed to the weak agreements on less preferred items (see Figures 8 and 9). Third, the

video was available in English, and participants did not have access to translation tools. For individuals whose first language was not English, future studies could make available translation software or dictionaries. Fourth, participants re-watched the video training prior to the generalization. This was done to avoid differences related to time elapsed between the video-training and generalization. As a result, participant's performance was likely higher than what it would have been without watching the video a second time. Future research could evaluate the generalization with a client, without asking the participants to re-watch the video training, and only requesting the participant to re-watch the video if they do not achieve the criterion, therefore evaluating maintenance and stimulus generalization.

STUDY 2: CLINICAL UTILITY OF PREFER

The purpose of Study 2 was to evaluate the utility of *Prefer* in a clinical setting. Specifically, clinicians working with clients with DD were asked to use the app for a period of time and then respond to a detailed questionnaire on the acceptability and usefulness of the app.

Method

Participants

Participants for Study 2 were six staff who worked with clients with DD. The only inclusion criterion was that they should be likely to have the opportunity to conduct stimulus preference assessments with clients as part of their normal duties during their participation. Two male and four female staff were recruited from four community agencies that provided services to individuals with DD. Three agencies were located in Manitoba, and one was located in Ontario.

Participants were asked to complete a demographics questionnaire (Appendix G). Table 4 shows the demographic information for five participants (one did not respond to the questionnaire). The mean age of the respondents was 29 years (range, 20 to 41 years) and educational level ranged from High School ($n = 2$) to undergraduate ($n = 3$). Three participants have completed at least 1 course in Behaviour Analysis; one participant also completed 6 graduate courses in Behaviour Analysis. Years of experience working with individuals with DD averaged 8.6 (range, 0.2 to 15). Participants reported to have contact with 14 to 130 clients per month. Only one participant worked with children; all others worked with adults. Three participants had training in stimulus preference assessments and had conducted assessments with clients. Participants reported to have conducted a mean of 4.2 preference assessments during the month before participant began (range, 0 to 20), and they anticipated that they would be conducting a mean of 5.4 preference assessments in the following month (range, 1 to 35).

Materials

Five participants downloaded *Prefer* from the App Store® and installed it on their own iOS® device, and one borrowed a device from the experimenter with the application installed. Slides with written instructions and pictures on how to use *Prefer* were available to participants throughout the study in the “Tutorials” icon in the app. The YouTube® playlist included a video on “How to Use Prefer” that contains a link in the description of video to download a document with the slides. (Appendix H shows the slides contained in the video and the document). All participants also had access to the training videos used in Study 1 for all three procedures, as the link to the YouTube® playlist was provided within the app. However, participants were not

asked or required to watch the videos.

Procedure

Training. I met with each participant to demonstrate how to use *Prefer*. I met with P1 and P2 individually. I met P3 with her supervisor. And I met with P4, P5, P6 and their supervisor as a group. Meetings with P2 and P3 occurred online. During each meeting, I showed how to use *Prefer* using the slides shown on Appendix H, or I asked the participants to watch the video on “How to Use Prefer” that showed the same slides, and then answered any questions they might have. The meetings ranged from 7 minutes to 30 minutes, depending on the number of questions the participants had.

Using *Prefer*. Each participant was asked to use the app with their clients for 4 weeks, assigning a different code for each client (so that the client’s identity would not be revealed). For P2 through P6, I sent an e-mail after two weeks reminding them to contact me if they have any question.

Data collection. After 4 weeks, I asked each participant to send the assessment data from their iOS device by e-mail using the back-up option in the app, and to complete a detailed Clinical Utility Questionnaire on the app. The questionnaire was developed for this study and contained 23 questions about the app and the video tutorials (Appendix I) covering the following features.

Help Me Choose. This section contained one question on the helpfulness of the guide (a series of yes-no questions) to help users select the appropriate preference assessment procedure based on various considerations (e.g., client characteristics, items to be assessed, etc.).

Video training. This section contained 9 questions. Participants were asked the number of times they have watched the videos, whether the videos were easy to understand, and whether they believed they conducted the assessments as shown on the video for each procedure (PS, MSWO, and FO).

Prompts during assessment. This section contained 6 questions about the prompts that the app offered during the assessments. Specifically, *Prefer* provided prompts to allow the client to sample each item, prompts for which items to present and their locations, prompts for which items to present on the next trial, a timer during the item access interval, and a graphical summary of the assessment results.

Back-up. This section contained 7 questions. Participants were asked about how often they used this feature, the importance of being able to back-up the assessment, and the usefulness of having a record of the items presented on each trial/session, the items selected on each trial or interval, the total time engaging with an item, the latency to choosing an item, and the total duration of the assessment.

Future usage. This section contained 3 questions. The first asked the likelihood of the participant to continue using the app, the second asked how likely the participant would recommend the app to a colleague, and the third asked how likely the participant would recommend the video tutorials to a colleague.

Reinforcers. This section contained one question that asked whether the assessments were useful in finding reinforcers.

Participants rated each question using a 5-point-scale, with 1 being the worst rating (e.g.,

“not helpful”, “not useful”, “not important”) and 5 being the best rating (e.g., “very helpful”, “very useful”, “very important”). Each question also included the option “unable to rate” (e.g., a participant had not used a feature). Participant could also provide written comments after each question.

Results

During the 1-month period, the six participants completed a combined total of 5 MSWO, 6 PS, and 6 FO assessments with 9 clients based on the data returned to the researcher. Participant 1 conducted one MSWO assessment and one FO assessment with one client, Participant 2 conducted two MSWO assessments with two clients, Participant 3 conducted one PS assessment and three FO assessments with one client, and Participant 4 conducted two PS assessments with one client. Participants 5 and 6 reported that they had completed their assessments together, totalling two MSWO assessments, three PS assessments, and two FO assessments.

Table 5 shows the mean ratings and range across participants for each question of the Clinical Utility Questionnaire. The number of respondents for each question and the number of positive or negative comments, if provided, are also shown in the table. The results of each feature are summarized below.

Help Me Choose

All six participants answered this question, and the mean rating for this feature was 4 (range, 1 to 5). One participant commented that the feature would be “extremely helpful for caregivers and direct care staff who might not understand the rationale of ‘why’ to choose a

specific assessment, and increase usability of the app”.

Video Training

Participants reported that they watched the PS video a mean of 1.7 times (range, 0 to 3). The five participants who watched the video found it easy to understand ($M = 4.2$, range 3 to 5). All four participants who used the procedure believed that they conducted the preference assessment quite closely to what they saw on the video ($M = 4.8$, range 4 to 5).

Participants watched the MSWO video a mean of 1.6 times (range, 0 to 3). The five participants who watched the video found it easy to understand ($M = 4.2$, range 3 to 5), and the four participants who used the procedure believed that they conducted the preference assessment quite closely to what they saw on the video ($M = 4.5$, range 4 to 5).

Participants watched the FO video a mean of 1.5 times (range, 0 to 3). The four participants who watched the video found it easy to understand ($M = 4.5$, range 4 to 5), and the three participants who used the procedure believed that they conducted the preference assessment quite closely to what they saw on the video ($M = 4$, range 3 to 5).

Prompts during Assessment

Prompts for item sampling, which items to present and their locations, items to present on the next trial, item access interval, and graphical summary of the results were all rated highly. Mean ratings across those questions ranged from 4.2 to 5, with rating below 3. At least one positive comment was provided for each question. In general, comments noted that the prompts were very important or very helpful especially for new staff members. For the graphical display of results, one participant wrote that he shared the chart with a colleague, who was “very

surprised by the results”, and said that “she would use the identified reinforcer more frequently in her sessions”.

Back-Up

Only two participants reported to have used the back-up feature. One did so for all the assessments and the other used it once. One participant justified that she did not use the feature because the e-mail was not set up on the agencies’ device. All six participants rated the importance of being able to back-up the data, the mean rating was 4 (range, 1 to 5). There were two positive comments: one participant acknowledged that he did not use the feature during the study, but noted that he could see himself using it more frequently in the future; the other noted that keeping records was very important in her agency. Two participants rated the usefulness of having a record of information. Specifically, the importance of being able to see the items presented on each trial had a mean rating of 4 (range, 3 to 5); the usefulness of having a record of the client’s selection on each trial had a mean rating of 4.5 (range, 4 and 5); the usefulness of having a record of the client’s engagement time with an item had a mean rating of 4.5 (range, 4 and 5); the usefulness of having a record of response latency on each trial had a mean rating of 3 (range, 1 and 5); and the usefulness of have a record of the duration of the assessment had a mean rating of 4 (range, 3 and 5).

Future Usage

When asked about the likelihood of continuing to use the app, participants rated it an average of 3.5 (range, 2 to 5). There was one positive comment indicating that the application “is very user- friendly and handy”. When asked about the likelihood of recommending the app to a

friend, participants rated it an average of 3.8 (range, 3 to 5), and there was one positive comment indicating that the app would be beneficial for all direct-care staff for the participant's agency.

When asked about the likelihood of recommending the videos to a colleague that did not have training on preference assessment procedures, participants rated it an average of 3.3 (range, 1 to 5). There were two positive comments: one stressed that she would, and the other indicated that he would likely to use the videos as a reminder to staff when he trains them.

Reinforcers

When asked about whether the app was helpful to find reinforcers for their interventions four participants rated average of 3.8 (range, 2 to 5). The remaining participants rated it as not applicable. There was one negative comment indicating that the participant was not sure on how to assess reinforcers that are not table activities and manipulative items.

Discussion

In general, participants rated the application features and the video training component positively. The features that received high ratings were prompts for item location on the next trial ($M = 5$), and prompts for item location on each trial ($M = 4.8$). The app also received high ratings for the pictorial prompts of the items to be presented ($M = 4.5$). The six participants had varied amount of education, training in behaviour analysis, and experience in DD, and they worked with clients of various age (see Table 4). That all six participants rated the video tutorials and the app's features positively supported the applicability of the app across a range of users and clients.

The FO video received slightly higher rating on being easy to understand ($M = 4.5$) than

the MSWO and PS videos ($M = 4.2$). The three videos were shot and edited in the same manner. However, the FO procedure is simpler to implement than the PS and MSWO procedures, requiring less intervention from the user.

Only two participants used the backup feature, although five of the six participants rated the feature as important. Moreover, two of the participants who did not use the feature provided positive comments about it. The items that received the highest rating on data back-up were: the record of the item selected during each trial or interval, and engagement time ($M = 4.5$ each).

One participant commented that he did not know how to assess non-tangible items (e.g., activities). This limitation was brought up during the initial meeting on how to use the app. The experimenter explained the possibility of using pictures to represent activities (or items), and the possibility of using the SS procedure for individuals with profound disabilities. However, the *Prefer* app did not have videos on how to use the pictorial preference assessments for activities or how to use SS procedure. Therefore, future improvement of the app should include video tutorials for SS procedure and for picture presentation of activities. The small number of participants in Study 2 also limits the generality of the findings. For example, only one of the five participants worked with children and three of the five participants already had some training or familiarity with stimulus preference assessments (see Table 4). Future utility studies could benefit from a larger sample such that results could be examined by subgroups with different characteristics (e.g., different amounts of training and experience). A second limitation of Study 2 was that participants received a 15-min to 1-hour session with the experimenter to demonstrate the use of the app. Future research should consider evaluating the effectiveness of

the app and video tutorials to use the app without any face-to-face training by the experimenter.

GENERAL DISCUSSION

The present studies join others in advancing the application and improving the use of computing technology to benefit individuals with DD. Study 1 showed that video training for *Prefer* was successful in teaching participants with and without a background in behaviour analysis to conduct PS, MSWO, and FO procedures. Study 2 showed that practitioners who worked with children and adults with DD reported that *Prefer* and its features were helpful after using it for one month in their normal work setting.

In evaluating psychosocial interventions, Lonigan, Elbert, and Johnson (1998) distinguished between efficacy and effectiveness studies. Efficacy studies are conducted under more controlled conditions, whereas effectiveness studies are conducted in real-world conditions. Along this continuum, Study 1 investigated the efficacy of the 3 video trainings to conduct preference assessments with *Prefer*, and Study 2 evaluated the effectiveness of the *Prefer*, including the video trainings. In Study 2, no attempts were made to control the characteristics of the participants, and participants were asked to use the app with their clients in their usual work settings.

Both studies showed consistent results with regards to the video trainings. Study 1 showed that most participants learned the procedures after watching the video trainings just once, while Study 2 showed that participants rated the tutorials as easy to understand. In addition, the FO procedure was rated as easier to understand than the other two procedure (i.e., PS and MSWO), and Study 1 showed that conducting this assessment procedure on *Prefer* indeed

seemed to be easier, as participants had high baseline levels during baseline.

Though the high baselines on FO procedure from Study 1 may suggest that if user familiarize themselves with the app, they might be able to apply the procedure without further need for trainings, Study 2 showed that users used only the procedures that had a training component to it. Moreover, it showed that one participant used the FO procedure to record the engagement with only one item successively (which could be done with the SS procedure in the app). Therefore, future research should evaluate the performance after familiarizing with the app, and how comfortable the participants feel about applying the procedure with an app before watching FO video trainings.

The time spent in training was typically a little higher than the video-training length, with the exception of P7 who watched the video twice in the training session. The mean time spent watching the tutorials were 12.7 min, 15.5 min, and 11.5 min for the PS, MSWO and FO procedures, respectively. The time spent in training seems lower than what the literature shows. For instance, Weldy et al.'s (2014) training session for the MSWO procedure was 30 min, and Bovi et al.'s (2016) MSWO video training was 17.2 min. Rosales et al. (2015) had shorter training time (i.e., 6.5, 9, and 9.43 min for the PS, MSWO and FO videos, respectively), but more than one video was evaluated with the same participant hence there could have been some carry-over effects.

Although possible carryover effects for training more than one procedure can hamper the appraisal of individual trainings, some studies have evaluated training package that consisted of videos that includes more than one procedure, taking advantage of the similarity between the

procedures and possible generalization (e.g., Lipschultz et al., 2016). Future research can evaluate training for more than one preference assessment procedure with an app with the same participants to increase training efficiency.

References

- Ahearn W. H., Clark K. M., DeBar, R., & Florentino, C. (2005). On the role of preference in response competition. *Journal of Applied Behavior Analysis, 38*, 247-250.
doi:10.1901/jaba.2005.36-04
- Arnal Wishnowski, L., Yu, C. T., Pear, J. J., Chand, C., & Saltel, L. (2018). Effects of computer-aided instruction on the implementation of the MSWO stimulus preference assessment. *Behavioral Interventions, 33*, 56-68. doi:10.1002/bin.1508
- Bovi, G. M. D., Vladescu, J.C., DeBar, R. M., Carroll, R. A., & Sarokoff, R. A. (2017). Using video modeling with voice-over instruction to train public school staff to implement a preference assessment. *Behavior Analysis in Practice, 10*, 72–76. doi:10.1007/s40617-016-0135-y
- 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. doi:10.1901/jaba.2000.33-353
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis*, 2nd edition. Pearson, Upper Saddle River, NJ.
- Conyers, C., Doole, A., Vause, T., Harapiak, S., Yu, D. C. T., & Martin, G. L. (2002). Predicting the relative efficacy of three presentation methods for assessing preferences of persons with developmental disabilities. *Journal of Applied Behavior Analysis, 35*, 49-58.
doi:10.1901/jaba.2002.35-49
- DeLeon, I. G., Frank, M. A., Gregory, M. K., & Allman, M. J. (2009). On the correspondence

between preference assessment outcomes and progressive-ratio schedule assessments of stimulus value. *Journal of Applied Behavior Analysis*, 42, 729–733.

doi:10.1901/jaba.2009.42-729

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, 29, 519-533.

doi:10.1901/jaba.1996.29-519

DeLeon, I. G., Iwata, B. A., Conners, J., & Wallace, M. D. (1999). Examination of ambiguous stimulus preferences with duration-based measures. *Journal of Applied Behavior*

Analysis, 32, 111–114. doi:10.1901/jaba.1999.32-111

Deliperi, P., Vladescu, J. C., Reeve, K. F., Reeve, S. A., & DeBar, R. M. (2015). Training staff to implement a paired-stimulus preference assessment using video modeling with voiceover

instruction. *Behavioral Interventions*, 30, 314-332. doi:10.1002/bin.1421

Developmental Disabilities Bill of Rights Act (1978). Public Law 95-682, United States of America.

Fisher, W. 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 to profound disabilities. *Journal of Applied Behavior Analysis*, 25, 491–498.

doi:10.1901/jaba.1992.25-491

Graff, R. B., & Karsten, A. M. (2012a). Assessing preferences of individuals with developmental disabilities: A survey of current practices. *Behavior Analysis in Practice*, 5, 37-48.

doi:10.1007/BF03391822

- Graff, R. B., & Karsten, A. M. (2012b). Evaluation of a self-instruction package for conducting stimulus preference assessments. *Journal of Applied Behavior Analysis, 45*, 69–82. doi:10.1901/jaba.2012.45-69
- Hagopian, L. P., Long, E. S., & Rush, K. S. (2004). Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28*, 668–677. doi:10.1177/0145445503259836
- Hagopian, L. P., Rush, K. S., Lewin, A. B., & Long, E. S. (2001). Evaluating the predictive validity of a single stimulus engagement preference assessment. *Journal of Applied Behavior Analysis, 34*, 475–485. doi:10.1901/jaba.2001.34-475
- Hanley, G. P., Iwata, B. A., Lindberg, J. S., & Conners, J. (2003). Response restriction analysis: I. Assessment of activity preferences. *Journal of Applied Behavior Analysis, 36*, 47–58. doi:10.1901/jaba.2003.36-47
- Hansard, C., & Kazemi, E. (2018). Evaluation of video self-instruction for implementing paired-stimulus preference assessments. *Journal of Applied Behavior Analysis, 51*, 675–680. doi:10.1002/jaba.476.
- Hawkins, R.P. (1979). The functions of assessment: Implications for selection and development of devices for assessing repertoires in clinical, educational, and other settings. *Journal of Applied Behavior Analysis, 12*, 501–516. doi: 10.1901/jaba.1979.12-501
- 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(1)*, 63–72. doi:10.1016/S0891-4222(98)00032-8

- 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.
doi:10.1016/S0891-4222(99)00030-X
- Hoffman, A. N., Samaha, A. L., Bloom, S. E. & Boyle, M. A. (2018). Preference and reinforcer efficacy of high- and low-tech items: A comparison of item type and duration of access. *Journal of Applied Behavior Analysis, 50*, 222-237. doi:10.1002/jaba.383
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis, 11*, 189-196. doi:10.1901/jaba.1978.11-189
- Horrocks, E. & Higbee, T.S. (2008). An evaluation of a stimulus preference assessment of auditory stimuli for adolescents with developmental disabilities. *Research in Developmental Disabilities, 29*, 11-20 doi:10.1016/j.ridd.2006.09.003
- Ivancic, M. T., & Bailey, J. S. (1996). Current limits to reinforcer identification for some persons with profound multiple disabilities. *Research in Developmental Disabilities, 17*, 77-92.
doi:10.1016/0891-4222(95)00038-0
- Julio, F., & University of Manitoba (2016). *Prefer* (Version 1.1) [iOS application]. Retrieved from: <https://itunes.apple.com/ca/app/prefer/id944700346?mt=8>
- Kahng, S., & Iwata, B. (1998). Computerized systems for collecting real-time observational data. *Journal of Applied Behavior Analysis, 31*, 253-261. doi:10.1901/jaba.1998.31-253
- Kang, S., O'Reilly, M. F., Fragale, C. L., Aguilar, J. M., Rispoli, M., & Lang, R. (2011). Evaluation of the rate of problem behavior maintained by different reinforcers across

- preference assessments. *Journal of Applied Behavior Analysis*, 44, 835–846.
doi:10.1901/jaba.2011.44-835
- Kautz, M. E., DeBar, R. M., Vladescu, J. C., & Graff, R. B. (2018). A further evaluation of choice of task sequence. *Journal of Special Education*, 52, 16-28.
doi:10.1177/0022466917735655
- Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, 1, 427-452. doi:10.1177/014544557714001
- Kazdin, A. E. (2011). *Single-case research designs (2nd ed.)*. New York, NY: Oxford University Press, Inc.
- Lee, M. S. H., Yu, D. 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.
doi:10.1901/jaba.2010.43-95
- Lipschultz, J. L., Vladescu, J. C., Reeve, K. F., Reeve, S. A., & Dipsey, C. R. (2015). Using video modeling with voiceover instruction to train staff to conduct stimulus preference assessments. *Journal of Developmental and Physical Disabilities*, 27, 505–532.
doi:10.1007/s10882-015-9434-4.
- Lonigan, C. J., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: an overview. *Journal of Clinical Child Psychology*, 27(2), 138-145. doi: 10.1207/s15374424jccp2702_1
- Marano, K.E., Vladescu, J. C., Reeve, K.F. & DiGennaro Reed, F. (in press). Effects of conducting behavioral observations and ratings on staff implementation of a paired-

- stimulus preference assessment. *Journal of Applied Behavior Analysis*. Advanced online publication doi:10.1002/jaba.584
- Martin, G. L., & Pear, J. J. (2015). *Behavior Modification: What it is and how to do it (10th ed.)*. Upper Saddle River, NJ: Pearson Education, Inc.
- Maulik P. K., Mascarenhas M. N., Mathers, C. D. , Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: a meta-analysis of population-based studies. *Research in Developmental Disabilities*, 32, 419–436. doi: 0.1016/j.ridd.2010.12.018
- Miljkovic, M., Kaminski, L., Yu, D. C. T., & Wishnowski, L. (2015). Evaluation of video modeling and self-instructional manual to teach students to conduct a preference assessment. *Journal on Developmental Disabilities*, 21(1), 3-10.
- Nottingham, C. L., Vladescu, J. C., Giannakakos, A. R., Schnell, L. K., Lipschultz, J. L. (2017). Using video-modeling with voice-over instruction plus feedback to train implementation of stimulus preference assessments. *Learning and Motivation*, 58, 37-47. Doi: 10.1010/j.lmot.2017.01.008
- 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. doi:10.1901/jaba.1985.18-249
- Paden, A. R. & Kodak, T. (2015). The effects of reinforcement magnitude on skill acquisition for children with autism. *Journal of Applied Behavior Analysis*. 48, 924-929. doi:10.1002/jaba.239.
- Paramore, N. W., & Higbee, T. S. (2005). An Evaluation of a Brief Multiple-Stimulus Preference

- Assessment with Adolescents with Emotional-Behavioural Disorders in an Educational Setting. *Journal of Applied Behavior Analysis*, 38, 399–403. doi:10.1901/jaba.2005.76-04
- 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. Retrieved from:
https://journals.sagepub.com/doi/abs/10.1207/s15328023top2302_14
- Ramon, D., & Yu, C. T. (2009). *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. University of Manitoba and St. Amant Research Centre, Canada.
- Ramon, D., Yu, C. T., Martin, G. L., & Martin, T. (2015). Evaluation of a self-instructional manual to teach multiple-stimulus without replacement preference assessments. *Journal of Behavioral Education*, 24, 289-303. doi:10.1007/s10864-015-9222-3
- Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, 31, 605-620. doi: 10.1901/jaba.1998.31-605
- 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. doi:10.1901/jaba.2006.7-05
- Roscoe, E. M., Iwata, B. A., & Kahng, S. (1999). Relative versus absolute reinforcement effects:

- implications for preference assessments. *Journal of Applied Behavior Analysis*, 32, 479–493. doi:10.1901/jaba.1999.32-479
- Rosales, R., Gongola, L. & Homlitas, C. (2015). An evaluation of video modeling with embedded instructions to teach implementation of stimulus preference assessments. *Journal of Applied Behavior Analysis*, 48, 209–214. doi:10.1002/jaba.174
- Schwartz, I. S., & Baer, D. M. (1991). Social Validity Assessments: Is Current Practice State of the Art? *Journal of Applied Behavior Analysis*, 24, 189-204. 794-808. doi:10.1901/jaba.1991.24-189
- Shapiro, M., Kazemi, E., Pogosjana, M., Rios, D., & Mendoza, M. (2016). Preference assessment training via self-instruction: A replication and extension. *Journal of Applied Behavior Analysis*, 49, doi:10.1002/jaba.339
- Smart, A. (2006). A multi-dimensional model of clinical utility. *International Journal for Quality in Health Care*, 18, 377-382. doi:10.1093/intqhc/mzl034
- Snyder, K., Higbee, T.S., & *Dayton, L. (2012). Preliminary investigation of a video-based stimulus preference assessment. *Journal of Applied Behavior Analysis*, 45, 13-18.
- Spevack, S., Wright, L., Yu, C.T., Walters, K., & Holborn, S. (2008). Passive and active approach responses in preference assessment for children with profound multiple disabilities and minimal movement. *Journal on Developmental Disabilities*, 14, 61-68. Retrieved from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3608576/>
- Spevack, S., Yu, C. T., Lee, M. S., & Martin, G. L. (2006). Sensitivity of passive approach during preference and reinforcer assessments for children with severe and profound

- intellectual disabilities and minimal movement. *Behavioral Interventions*, *21*, 165-175
doi:10.1002/bin.216
- Stoesz, B., McCombe, L., Martin, T., & Yu, C. T. (2016). A knowledge brokering process for challenging behaviours in special education. *Journal on Developmental Disabilities*, *22*(2), 47-63.
- Toussaint, K. A., Kodak, T., & Vladescu, J. C. (2016). An evaluation of choice on instructional efficacy and individual preferences among children with autism, *Journal of Applied Behavior Analysis*, *49*, 170-175. doi:10.1002/jaba.263
- Tullis, C. A., Cannella-Malone, H. I., Basbigill, A. R., Yeager, A., Fleming, C., 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.
- Tullis, C. A., Cannella-Malone, H. I., & Fleming, C. V. (2012). Multiple stimulus without replacement preference assessments: An examination of the relation between session number and effectiveness. *Journal of Developmental and Physical Disabilities*, *24*, 337-345. doi:10.1007/s10882-012-9273-5
- Tung, S. & Donaldson, J., & Kahng, S. W. (2017). The effects of preference assessment type on problem behavior. *Journal of Applied Behavior Analysis*, *50*, 861-866.
doi:10.1002/jaba.414.
- Verriden, A. L., & Roscoe, E. M. (2016). A comparison of preference-assessment methods. *Journal of Applied Behavior Analysis*, *49*, 265-285. doi:10.1002/jaba.302

- Virués-Ortega, J. V, Pritchard, K., Grant, R. L., North, S., Hurtado-Parraro, C., Lee, M., Temple, B., Julio, F. M., & Yu, C.T. (2014). Clinical decision-making and preference assessment for individuals with intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disability, 119*, 157 -170. doi:10.1352/1944-7558-119.2.151
- Weeden, M., & Poling, A. (2011). Identifying reinforcers in skill acquisition studies involving participants with autism: Procedures reported from 2005 to 2009. *Research in Autism Spectrum Disorders, 5*, 388–391. doi:10.1016/j.rasd.2010.05.004
- Weldy, C. R., Rapp, J. T., & Capocasa, K. (2014). Training staff to implement brief stimulus preference assessments. *Journal of Applied Behavior Analysis, 47*, 1–5. doi:10.1002/jaba.98
- Whiting, S. W., & Dixon, M. R. (2012). Creating an iPhone application for collecting continuous ABC data. *Journal of Applied Behavior Analysis, 45*, 643–56. doi:10.1901/jaba.2012.45-643
- Whiting, S. W., & Dixon, M. R. (2014). Portable data assistants for data collection in autism treatment. In V. B. Patel, V. R. Preedy, & C. R. Martin (Eds.), *Comprehensive Guide to Autism* (pp. 2835-2852). New York. Springer
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214. doi:10.1901/jaba.1978.11-203

Table 1

Comparison of Preference Assessment Apps

Features	<i>myPref</i> ¹ by KV Adaptive LLC, 2016	<i>Preference & Reinforcement Assessment</i> by Touch Autism, 2011
<i>Assessment procedures</i>		
Paired-Stimulus	✓	✓
Multiple-Stimulus-without-Replacement	-	✓
Multiple-Stimulus	-	✓
Free-Operant	-	-
Single-Stimulus	-	-
<i>Aid to select the appropriate procedure</i>	-	-
<i>Procedural Prompts</i>		
Pre-assessment prompt (individual sampling)	-	✓
Location of the items on each trial	✓	✓
Photo of the item as prompt	✓	-
Present appropriate item after selection	-	✓
Consequence duration timer	-	-
<i>Graphical results</i>	-	✓
<i>Data logging capabilities</i>		
Percent preference	✓	✓
Items presented on each trial	-	-
Prompt for item location	-	-
Engagement time	-	-
Latency of selection response	-	-
Assessment duration	-	-
Data back-up	✓	-
<i>Evidence-based tutorial</i>	-	-
<i>Open source</i>	-	-

¹ABA*pref* was replaced with *myPref* in 2016

Table 2

Characteristics of Participants in Study 1

Participant	Age (Yrs)	Gender	Student/Staff	Education	Experience with DD
<i>Paired-Stimulus</i>					
P1	23	F	Student	Microbiology (MSc 1 st year)	None
P2	29	M	Student	Microbiology (PhD 4 th year)	Some
P3	35	F	Student	Biology (BSc 5 th year)	Some
P4	24	F	Staff	Psychology (BA completed)	A lot
P5	25	F	Staff	Microbiology (BSc 1 st year)	Some
<i>Multiple-Stimulus-without-Replacement</i>					
P6	47	F	University Staff	English (BA completed)	None
P7	23	F	Student	Economics (BA 3 rd year)	None
P8	19	M	Student	Engineering (BSc 2 nd year)	Some
P9	25	F	Staff	Psychology (BA 4 th year)	A lot
P10	22	F	Staff	Psychology (BA completed)	Some
<i>Free-Operant</i>					
P11	20	M	Student	Biotechnology (BSc 1 st year)	Some
P12	18	F	Student	Science (BSc 1 st year)	Some
P13	21	M	Student	Science (BSc 3 rd year)	None
P14	20	F	Student	International Business & Human Resources (BComm 3 rd year)	Some
P15	23	M	Student	Psychology (BA 2 nd year)	Some

Table 3

Feedback Questionnaire Results in Study 1

<i>Questionnaire Items</i>	<i>PS (n=5)</i>	<i>MSWO (n=5)</i>	<i>FO (n=5)</i>	<i>Total (n=15)</i>
1. In your opinion, how important it is for individuals who work with people with intellectual and developmental disabilities to learn how to conduct preference assessments?	$M = 4.6$ (range, 4 to 5)	$M = 4.6$ (range, 4 to 5)	$M = 4.6$ (range, 4 to 5)	$M = 4.6$ (range, 4 to 5)
2. Did you find the video tutorial easy or difficult to understand?	$M = 4.8$ (range, 4 to 5)	$M = 4.4$ (range, 3 to 5)	$M = 4.6$ (range, 4 to 5)	$M = 4.6$ (range, 3 to 5)
3. Did you find any part of the video tutorial confusing?	Yes, $n = 1$; No, $n = 4$	Yes, $n = 1$; No, $n = 4$	Yes, $n = 1$ No, $n = 4$	Yes, $n = 3$ No, $n = 12$
4. How confident do you feel about applying this procedure with real clients?	$M = 4.0$ (range, 3 to 5)	$M = 4.0$ (range, 3 to 5)	$M = 4.2$ (range, 4 to 5)	$M = 4.1$ (range, 3 to 5)
5. How helpful was having the iOS application for conducting the assessment?	$M = 4.8$ (range, 4 to 5)	$M = 4.0$ (range, 3 to 5)	$M = 4.6$ (range, 3 to 5)	$M = 4.5$ (range, 3 to 5)
6. How many times did you watch the full video tutorial?	1, $n = 5$	1, $n = 3$ >1, $n = 2$	1, $n = 5$	1, $n = 13$ >1, $n = 2$
7. Was there any specific part of the video that you watched more times?	Yes, $n = 4$ No, $n = 1$	Yes, $n = 2$ No, $n = 3$	Yes, $n = 1$ No, $n = 3$	Yes, $n = 7$ No, $n = 8$
8. While watching the video, did you practice presenting the items?	Yes, $n = 3$ No, $n = 2$	Yes, $n = 4$ No, $n = 1$	Yes, $n = 2$ No, $n = 3$	Yes, $n = 9$ No, $n = 6$

PS = paired-stimulus, MSWO = multiple-stimulus-without-replacement, FO = free-operant

Table 4

Participant Characteristics in Study 2

	P1	P2	P3	P4	P5
Gender, Age (years)	Male, 29	Female, 35	Male, 41	Female, 20	Female, 21
Education	Undergrad	Undergrad	Undergrad	High school	High school
Courses in behaviour analysis	2 undergrad	2 undergrad, 6 grad	1 undergrad	Unsure	None
Years working in DD or autism field	12	14.6	15	1.3	0.2
Number of client contacts in a month	14	> 25	Approx.130	25	35
Age range (yrs) of the clients	3 to 5	18 to 85	18 to 65	21 to 65	21 to 65
Current position and years	Senior Tutor, 7.2	Outreach Behaviour Consultant, 10	Literacy and Communication Instructor, 9	Support Worker, 0.5	Direct Support Professional, 0.2
Prior training in SPAs	Yes (MSWO, PS)	Yes	Yes (YouTube® videos)	No	No
Number of SPAs administered previously	MS 50; PS 50	MSWO 10; MS 10; PS 20; FO 25; SS 10	MSWO 4; PS 2; FO 10	0	0

Notes: One participant (P6) did not respond to the demographics questionnaire. MS = multiple-stimulus; MSWO = multiple-stimulus-without-replacement; PS = paired-stimulus; SS = single-stimulus; FO = free-operant; SPA = stimulus preference assessment.

Table 5

Summary of Results from Clinical Utility Questionnaire in Study 2 (n = number of respondents)

Question/Statement	<i>n</i>	Mean Rating (range)	Comments +/-
<i>Help me Choose</i>			
Rate the helpfulness of this feature.	6	4 (1-5)	1+
<i>Video Training</i>			
How many times did you watch the PS video tutorial?	6	1.7 (0-3)	
Was the PS video difficult or easy to understand?	5	4.2 (3-5)	1+
Do you believe that you conducted the PS preference assessment with your client(s) in the same way as you saw in the video tutorial?	4	4.8 (4-5)	
How many times did you watch the MSWO video tutorial?	6	1.6 (0-3)	
Was the MSWO video difficult or easy to understand?	5	4.2 (3-5)	
Do you believe that you conducted the MSWO preference assessment with your client(s) in the same way as you saw in the video tutorial?	4	4.5 (4-5)	
How many times did you watch the FO video tutorial?	6	1.5 (0-3)	
Was the FO video difficult or easy to understand?	4	4.5 (4-5)	
Do you believe that you conducted the FO preference assessment with your client(s) in the same way as you saw in the video tutorial?	3	4.0 (3-5)	
<i>Prompts During Assessment</i>			
Sample Each Item: The app prompts you to let the client sample the items before an assessment, did you find the reminder helpful?	6	4.3 (3-5)	1+
Pictorial Representation: During the assessment, <i>Prefer</i> shows a picture to represent the items that should be presented on each trial. How helpful was it to have a pictorial representation of these items (i.e., having the picture of the item as opposed to have the name)?	6	4.5 (3-5)	1+
Item Location: When applicable, during an assessment, <i>Prefer</i> will rotate the pictures after a trial, and show a	5	4.8 (4-5)	1+

number below the pictures representing the location that they should be presented. How helpful was it to be able to see the location of the items in *Prefer* (i.e., in addition to which item to present)?

Next Trial Item Location: For the PS procedure, a box on the lower half of the screen would show which items should be presented on the following trial, and its location. How helpful was it to be able to see the location of the items for the next trial?	2	5.0	1+
---	---	-----	----

Access Time. For the PS and MSWO procedures, after a selection is made, a window pops-up prompting the user to allow the client to have access to the selected item, and offers a timer. How helpful was this prompt?	5	4.2 (3-5)	1+
---	---	-----------	----

Results. After a preference assessment is concluded, <i>Prefer</i> shows a graph with the results. How helpful was this graph?	5	4.3(3-5)	2+
--	---	----------	----

Back-Up

How often did you use this feature?	6	1.5 (0-5)	
-------------------------------------	---	-----------	--

How important do you think it is to be able to back-up the results of the assessment (i.e., to be able to export the result to a different device/computer)?	6	4.0 (1-5)	2+
--	---	-----------	----

Items presented on each trial or session with the location	2	4.0 (3-5)	
--	---	-----------	--

Item selected on each trial or interval (i.e., item with the name capitalized)	2	4.5 (4-5)	
--	---	-----------	--

Engagement time	2	4.5 (4-5)	
-----------------	---	-----------	--

Latency to choose an item (i.e., how long it took tester to record the selection on each trials)	2	3.0 (1-5)	
--	---	-----------	--

Duration of the assessment	2	4.0 (3-5)	
----------------------------	---	-----------	--

Future Usage

Do you plan to continue to use the app to conduct preference assessment in the future if an iOS device is available?	6	3.5 (2-5)	1+
--	---	-----------	----

Would you recommend <i>Prefer</i> to a colleague?	6	3.8 (3-5)	1+
---	---	-----------	----

Would you recommend the video tutorial to colleagues who have not received training yet?	6	3.3 (1-5)	2+
--	---	-----------	----

Reinforcers

Were the assessment results useful in helping you to identify reinforcer(s) for your interventions?	4	3.8 (2-5)	1-
---	---	-----------	----

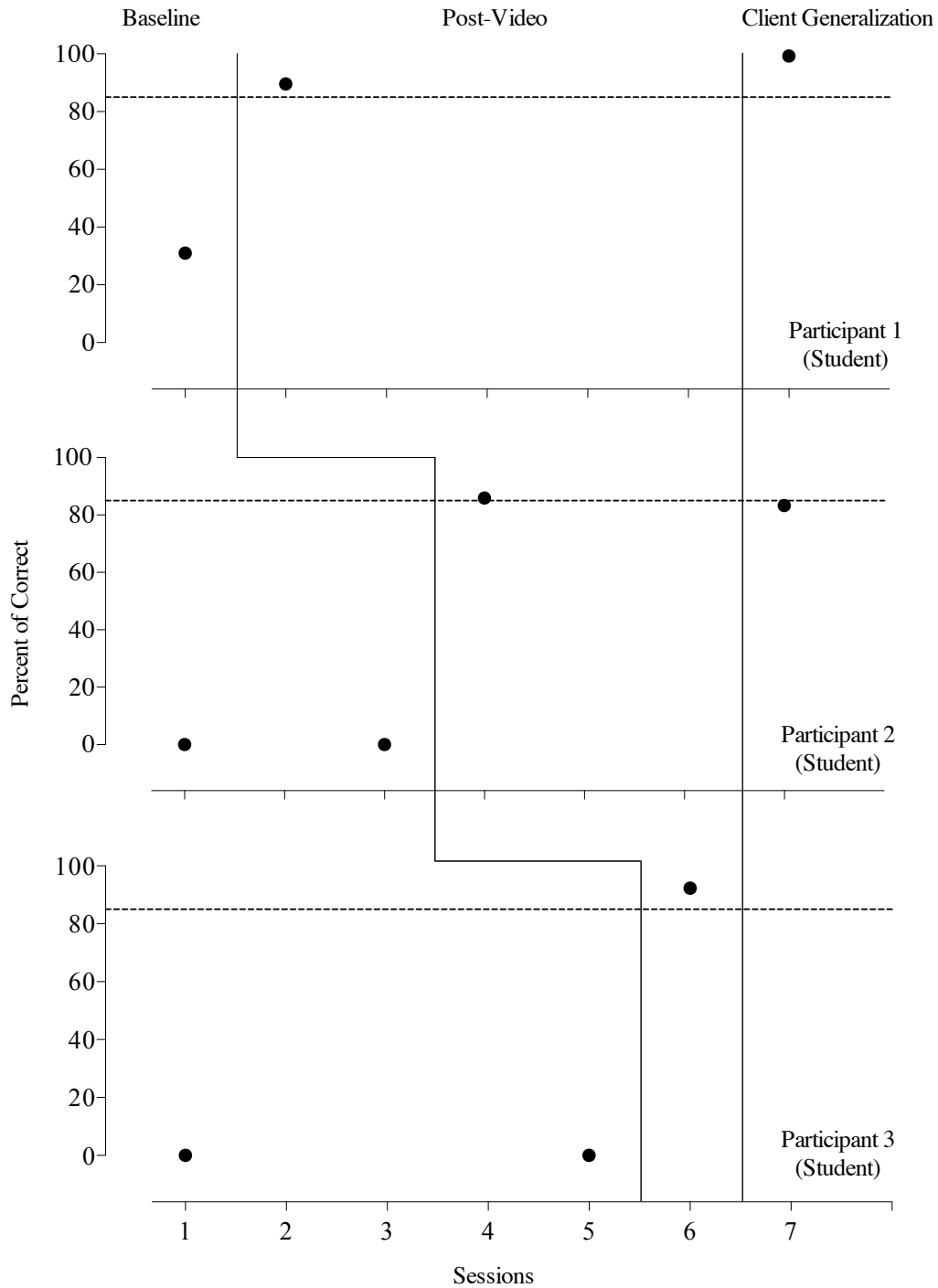


Figure 1. Percent of correct responses during Paired-Stimulus simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 1 through 3. The horizontal dash line in each graph represents 85% correct.

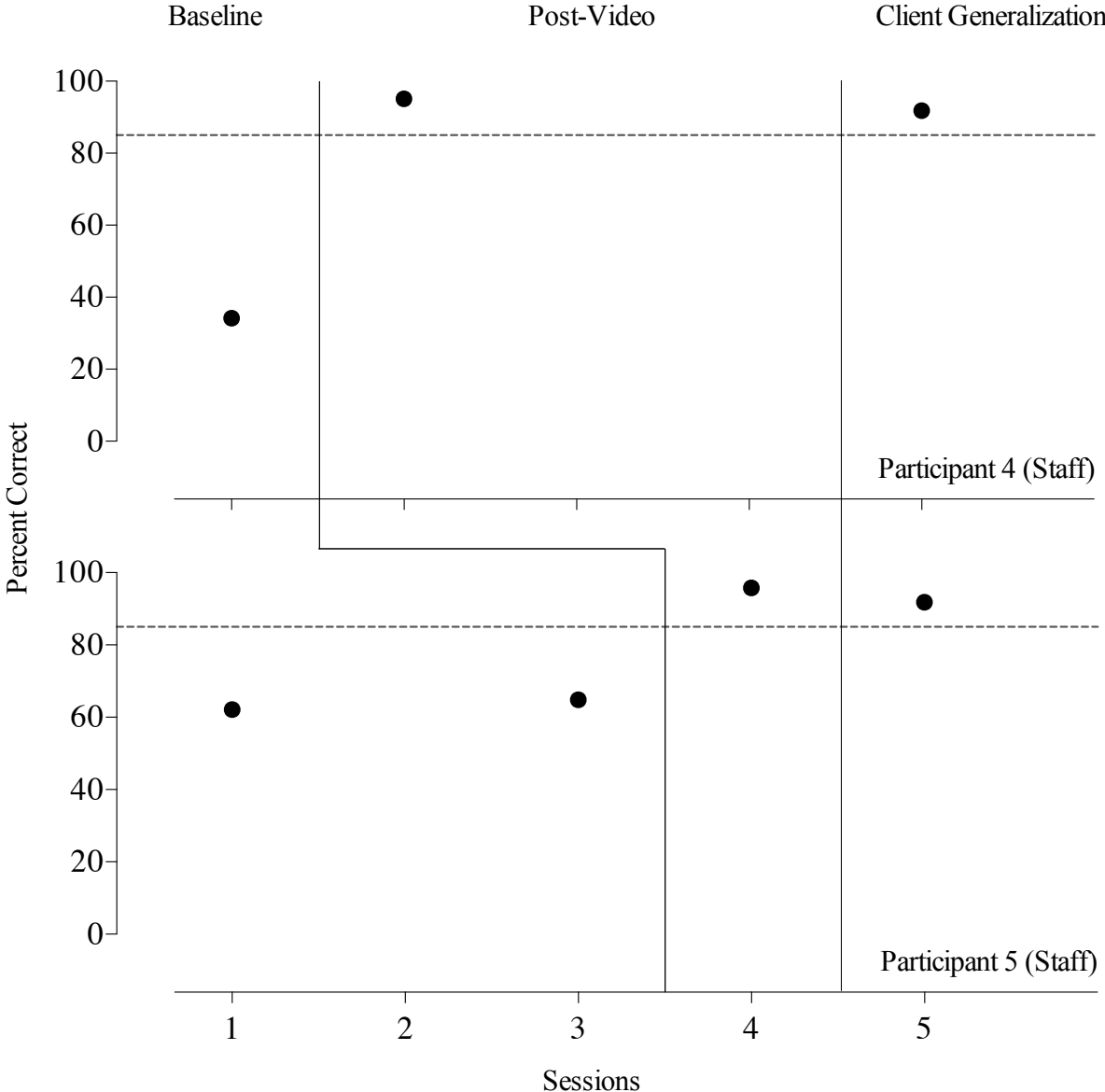


Figure 2. Percent of correct responses during Paired-Stimulus simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 4 and 5. The horizontal dash line in each graph represents 85% correct.

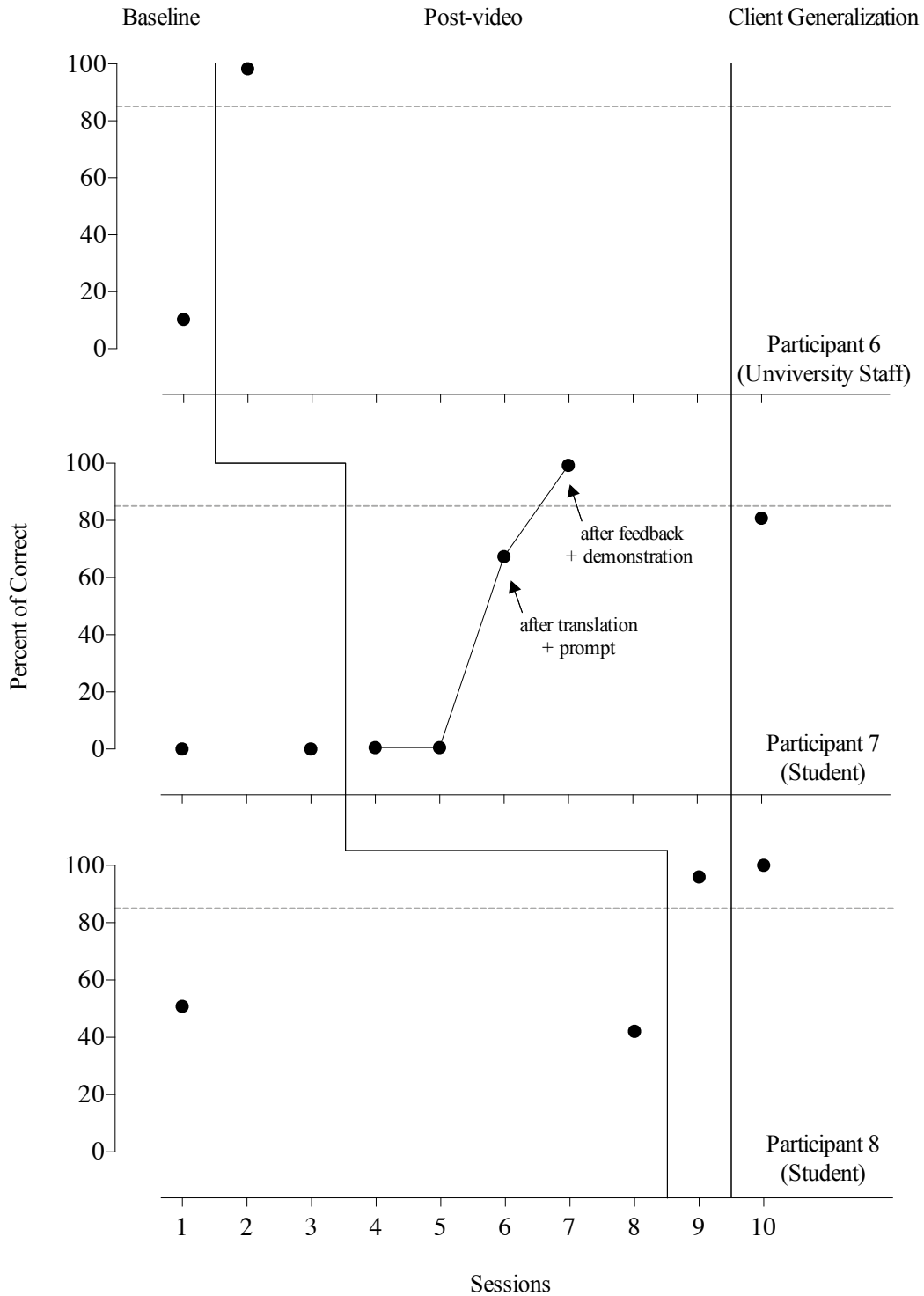


Figure 3. Percent of correct responses during Multiple-Stimulus-without-Replacement simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 6 through 8. The horizontal dash line in each graph represents 85% correct.

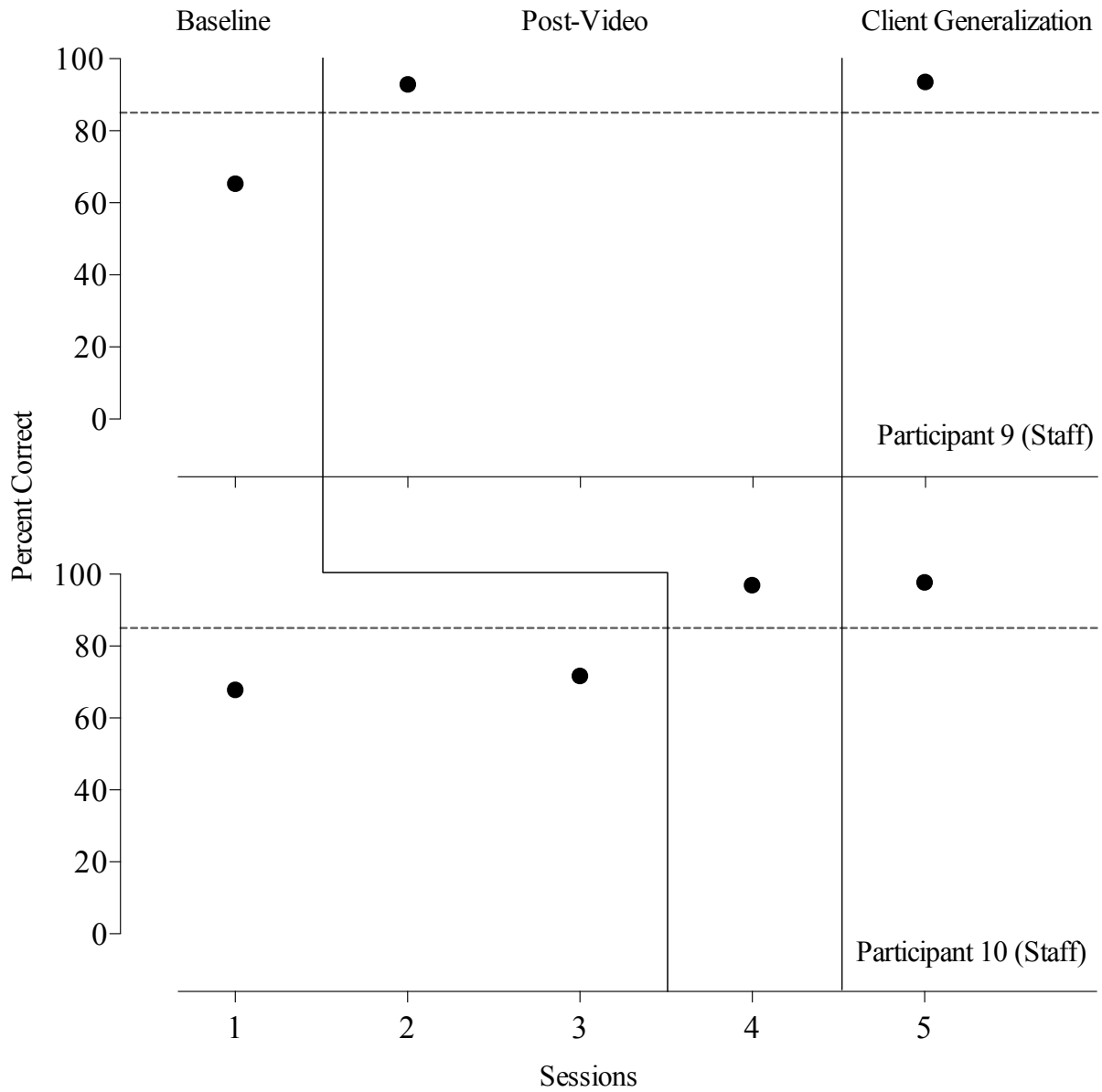


Figure 4. Percent of correct responses during Multiple-Stimulus-without-Replacement simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 9 and 10. The horizontal dash line in each graph represents 85% correct.

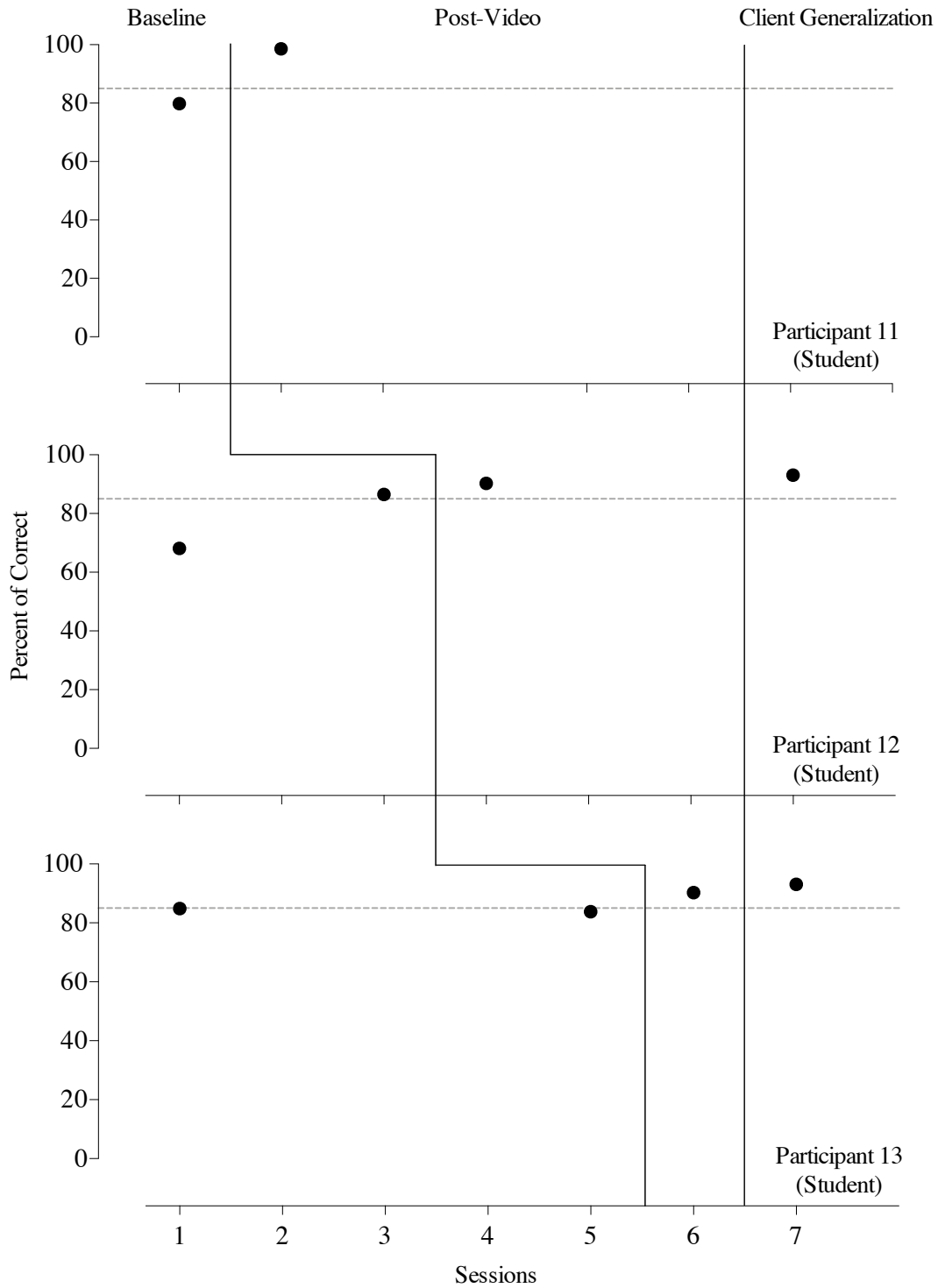


Figure 5. Percent of correct responses during Free-Operant simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a client for Participants 11 through 13. The horizontal dash line in each graph represents 85% correct.

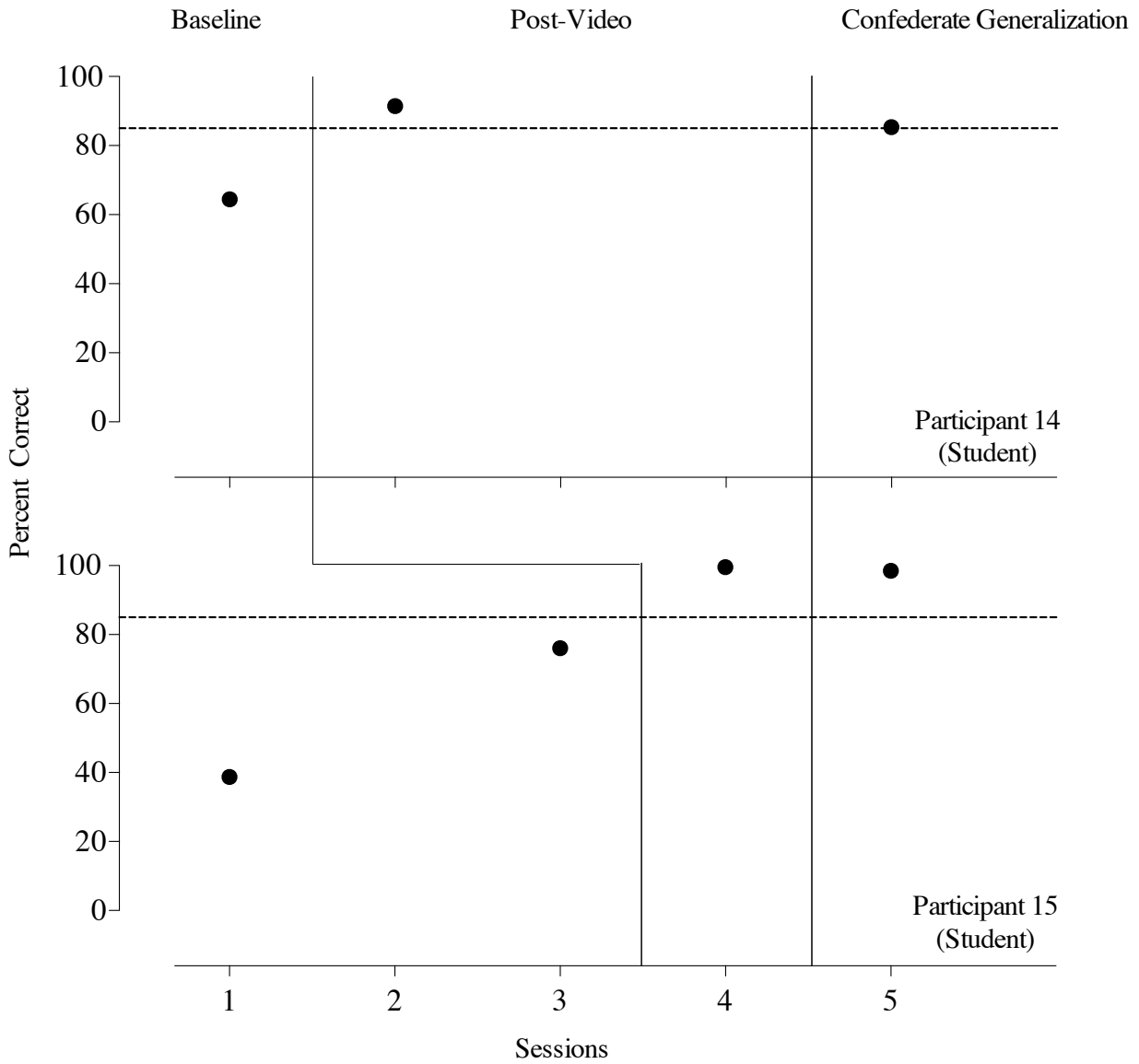


Figure 6. Percent of correct responses during Free-Operant simulated assessments before (baseline) and after training (post-video), and during generalization assessment with a confederate for Participants 14 and 15. The horizontal dash line in each graph represent 85% correct.

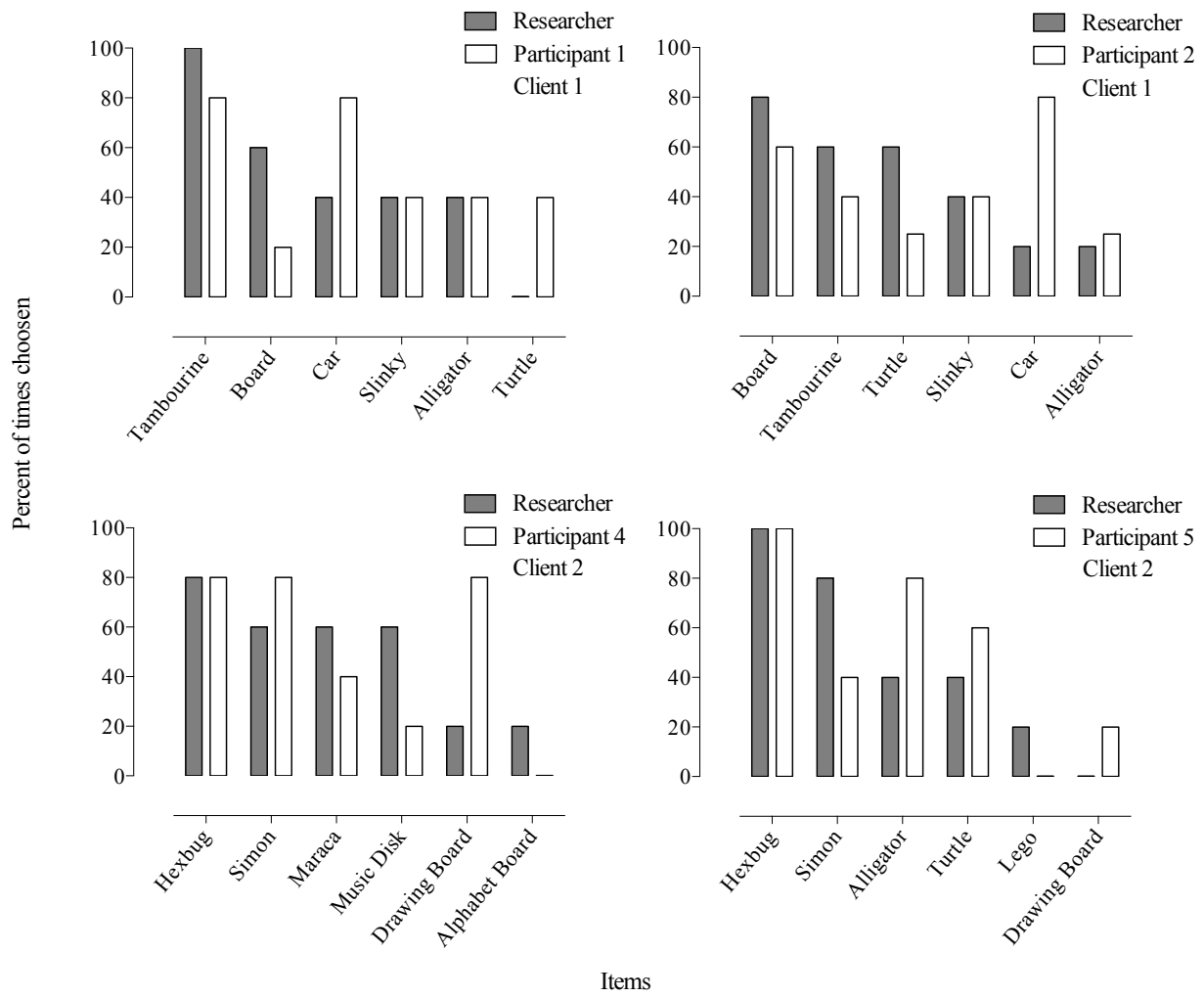


Figure 7. Comparison of the results of the Paired-Stimulus assessments conducted by the researcher using paper-and-pencil, and by Participants 1, 2, 4, and 5 using *Prefer*. Items are ordered from the most to the least preferred on the horizontal axis based on the researcher's assessment.

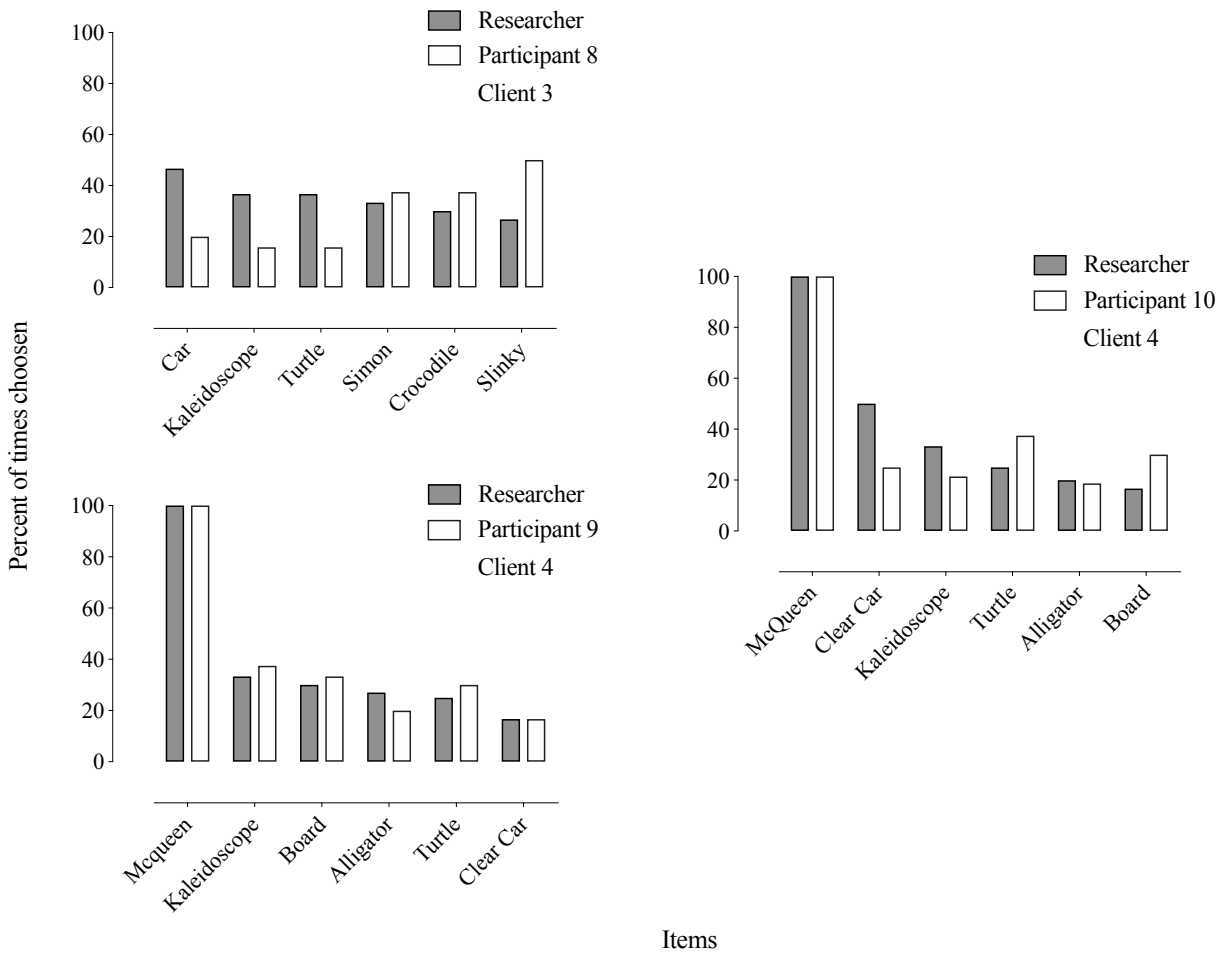


Figure 8. Comparison of the results of the Multiple-Stimulus-without-Replacement assessment conducted by the researcher using paper-and-pencil, and by Participants 8, 9, and 10 using *Prefer*. Items are ordered from the most to the least preferred on the horizontal axis based on the researcher’s assessment.

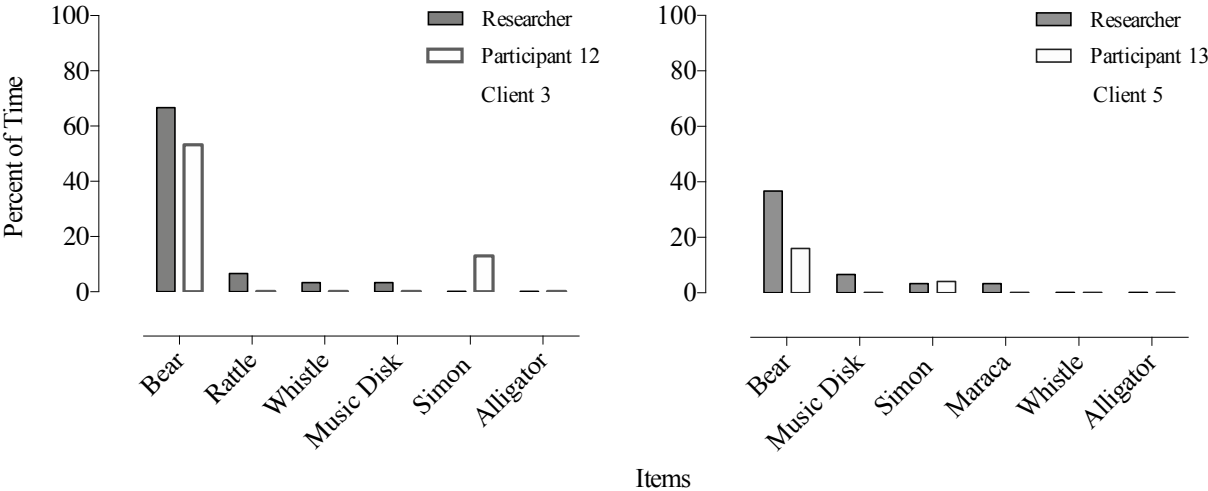


Figure 9. Comparison of the results of the Free-Operant assessment conducted by the researcher using paper-and-pencil, and Participants 12 and 13 using *Prefer*. Items are ordered from the most to the least preferred on the horizontal axis based on the researcher’s assessment.

Appendix A

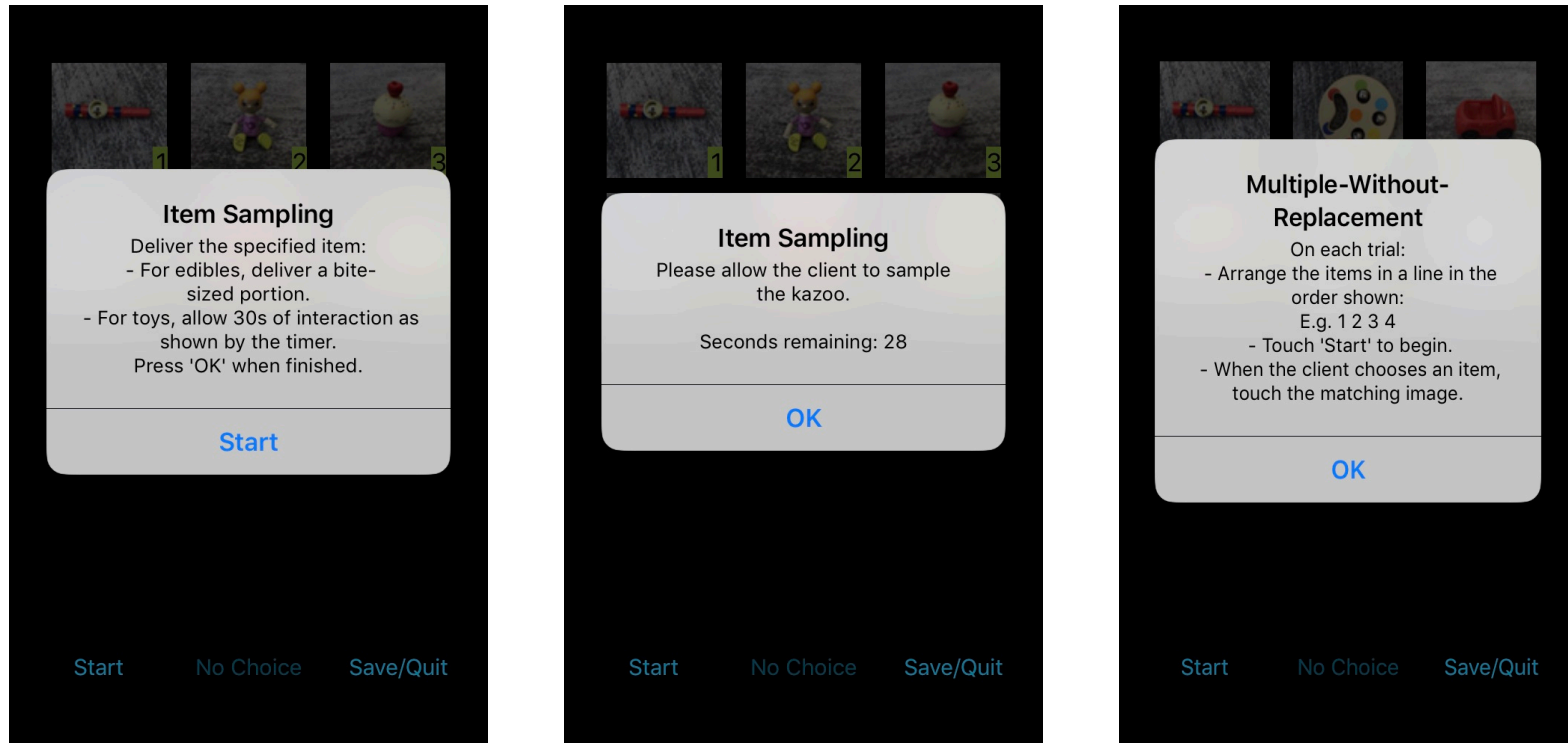
Screen Shots of the *Prefer* App

This appendix shows the main menu of *Prefer*, sample screens with prompts for the user to conduct item sampling and how to arrange the items for presentation, sample screens for recording, sample graphs of the assessment results.

Prefer main screen

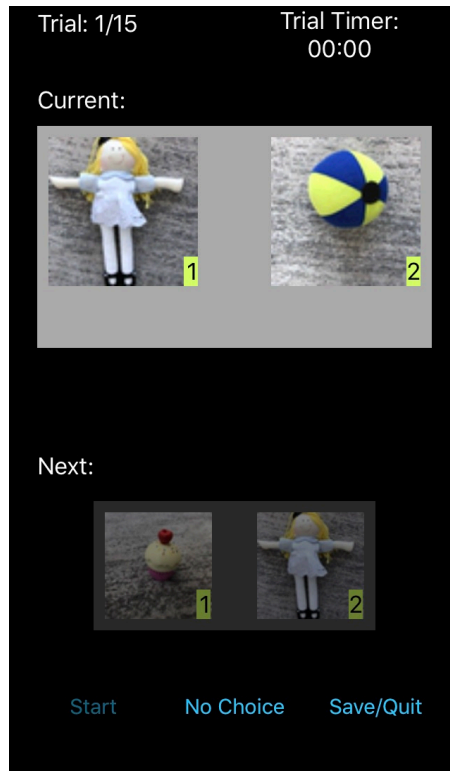
Main Menu	
PROTOCOLS	
Free-Operant	>
Multiple-Without-Replace...	>
Multiple-With-Replacement	>
Paired-Stimulus	>
Single-Stimulus-Engageme...	>
LIBRARY	
Stimulus Library	>
SUPPORT	
Help Me Choose	>
Tutorials	

Assessment set up for MSWO procedure

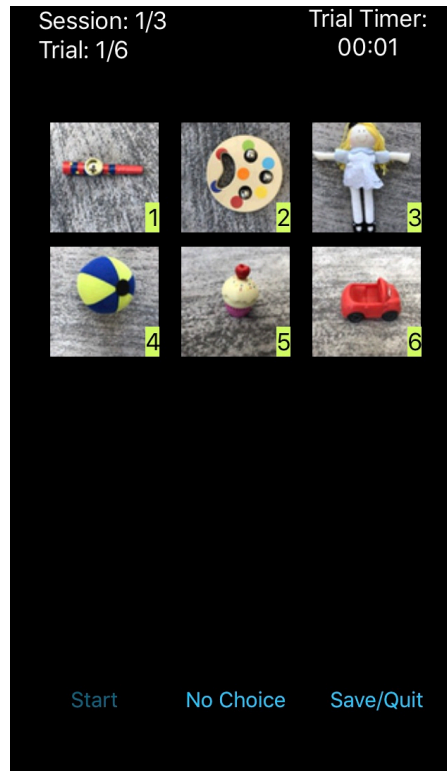


All assessment procedures start with prompts to allow the client to sample the items to be assessed for a user-specified interval. Later, the user receives a prompt on how to arrange the items for the assessment.

Preference assessment recording screen for the three procedures:



Paired-Stimulus (PS)



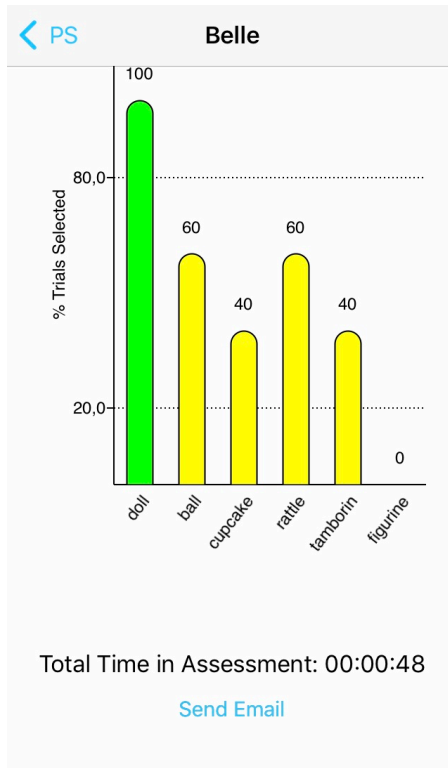
Multiple-Stimulus-without-Replacement (MSWO)



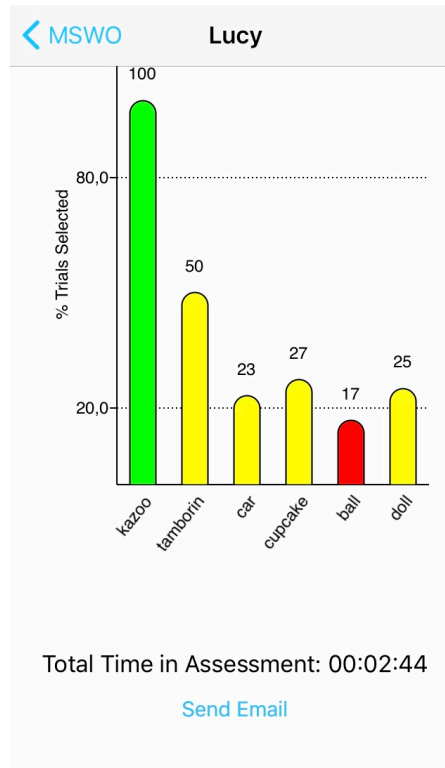
Free-Operant (FO)

The user should touch the matching picture to record the selection (PS and MSWO) or engagement (FO). For the PS procedure, there is also a prompt for the items to be presented and their positions on the next trial. For the FO procedure, the selected icon will be shaded blue.

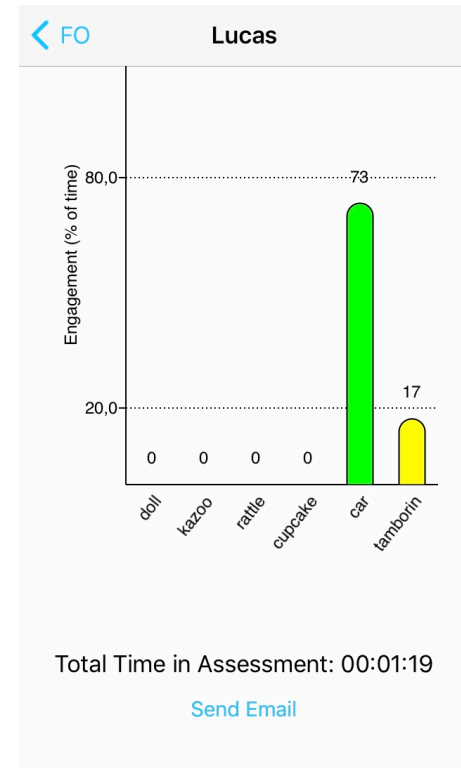
Screen shoots of the example results from the three procedures.



Paired-Stimulus (PS)



Multiple-Stimulus-without-Replacement (MSWO)



Free-Operant (FO)

Immediately after the assessment, a bar graph of the assessment results appears on the screen. The most preferred item is displayed in a green bar and least preferred item is in red.

Appendix B

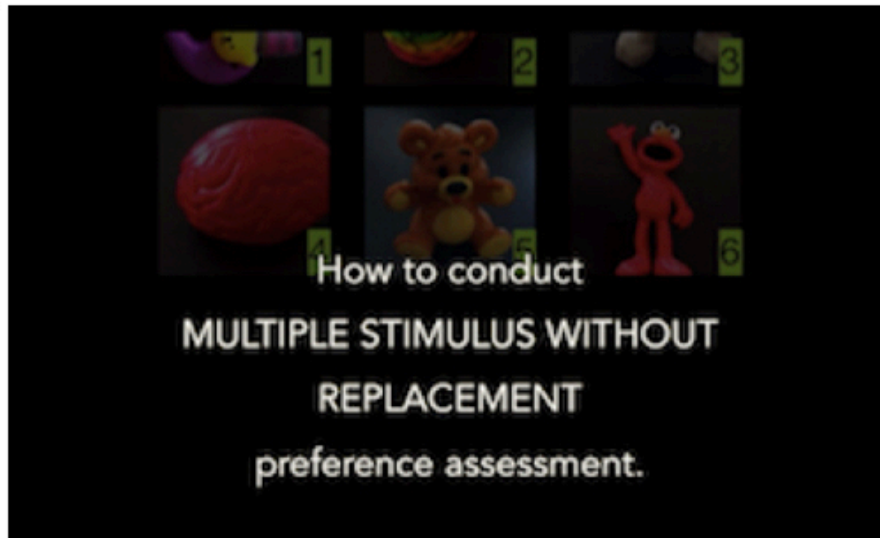
YouTube® Links, and Screen Shots of the MSWO Video

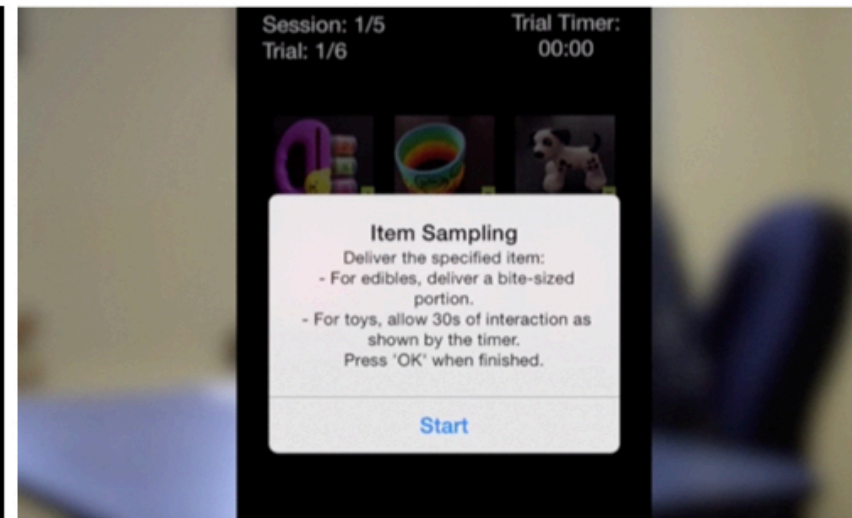
PS video tutorial: <https://youtu.be/EFf910FJIRs>

MSWO video tutorial: <https://youtu.be/OW4ZRihq6is>

FO video tutorial: <https://youtu.be/eCo2aMPEuag>


Examples of screen shots from the beginning of the MSWO video





MSWO Preference Assessment

After presenting the last item, you will receive a prompt to start the MSWO preference assessment.



Multiple no Replacement
 On each trial:
 - Arrange the items in a line in the order shown:
 E.g. 1 2 3 4
 - Touch 'Start' to begin.
 - When the client chooses an item, touch the matching image.

OK

Hold each item in front of the client and instruct him/her to look at it, ensuring that the client is paying attention to the task.

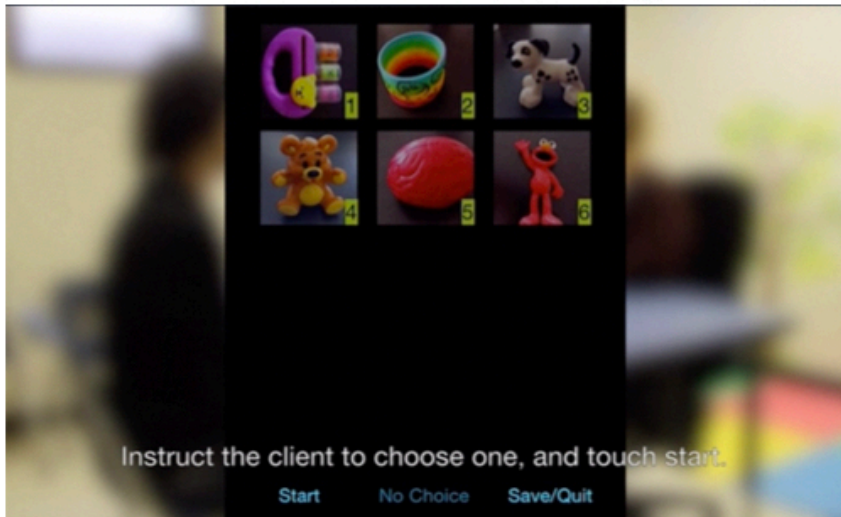


They should have the same distance from each other,





Now that you have practiced presenting the items, let's continue watching what should happen after a trial is presented.



Touch the matching icon to record the selection.

Start No Choice Save/Quit

You will immediately receive a prompt to allow the client to play with the item for 30 s.

Item Access
Seconds remaining: 28
OK

Item Access
Seconds remaining: 19
OK

Session: 1/3
Trial: 2/6
Trial Timer: 00:00

Remove the retrieved item, and present the next trial in the same way, as on Prefer

Appendix C

**Procedural Integrity Checklist for the Paired-Stimulus Procedure
(Study 1)**

PS PROCEDURE: PARTICIPANT SCORING SHEET

Participant: _____
Session date: _____

Researcher : _____
Phase: _____

PARTICIPANT SCORING: Mark correct as “1”, and incorrect as “0”.

Preparation before assessment															
Sampling: Allow client to interact with each item for 30 sec															
Trial presentation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Provide de instruction to “look” when presenting each item															
Present the correct items, in the correct location															
Provide the instruction to “pick one”															
Appropriate Selection															
Provide access to the selected item															
If client approaches a second item: Block / Retrieve item															
Remove unselected item															
Praise task engagement (if only one item was approached)															
Record selection in the app (check app file)															
End trial after 30 s / Retrieve item from client															
Selection of two items simultaneously															
Block approach to items (if possible)/ Retrieve items															
Re-present instruction to “pick one”															
Provide access to the selected item															
Remove unselected item															
Praise task engagement															
Record selection in the app (check app file)															
End trial after 30 s / Retrieve items from client															
No Choice after 15 sec															
Re-present instruction to pick one item (If client selects an item complete “Appropriate Selection”)															
If no selection is made record “No choice” (check app file)															
End trial after 30 s / Retrieve items															
Interpreting results/graph															
Name the correct item(s) as preferable according to the app’s results.															

TOTAL: _____ [correct / (correct + incorrect)]*100

Appendix D

Procedural Integrity Checklist for the Multiple-Stimulus-without-Replacement Procedure (Study 1)

MSWO PROCEDURE: PARTICIPANT SCORING SHEET

Participant: _____
 Session date: _____

Researcher : _____
 Phase: _____

PARTICIPANT SCORING: Mark correct as “1”, and incorrect as “0”.

Preparation before assessment															
Sampling: Allow client to interact with each item for 30 sec															
Trial presentation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Provide de instruction to “look” when presenting each item															
Present the correct items, in the correct location															
Provide the instruction to “pick one”															
Appropriate Selection															
Provide access to the selected item															
If client approaches a second item: Block / Retrieve item.															
Remove unselected items															
Praise task engagement (if only one item was approached)															
Record selection in the app (check app file)															
End trial after 30 s / Retrieve item from client															
Selection of two items simultaneously															
Block approach to items (if possible)/ Retrieve items															
Re-present instruction to “pick one”															
Provide access to the selected item															
Remove unselected items															
Praise task engagement															
Record selection in the app (check app file)															
End trial after 30 s / Retrieve items from client															
No Choice after 15 sec															
(1x)Record “ No choice” (check app file)															
Re-present instruction to “pick one”															
(2x)Record “ No choice” (check app file)															
Re-present instruction to “pick one”															
(3x)Record “ No choice” (check app file)															
End Session															
Selection followed by “no choice”															
Allow access to item															
Remove unselected items															
Praise task engagement															
Record selection in the app (check app file)															
End trial after 30 s / Retrieve items from client															
Interpreting results/graph															
Name the correct item(s) as preferable according to the app’s results															

TOTAL: _____ [correct / (correct + incorrect)]*100

Appendix E

Procedural Integrity Checklist for the Free-Operant Procedure (Study 1)

FO PROCEDURE: PARTICIPANT SCORING SHEET

Participant: _____ Researcher : _____ Session date: _____ Phase: _____

Preparation for Assessment & Results: Mark correct with "1", and incorrect with "0".	
Preparation for assessment	
Sampling: Allow client to interact with each item for 30 sec	
Session presentation	
Provide the instruction to "look" when presenting each item	
Present the correct items, in the correct location.	
Provide the instruction for the client to "Play with the toys."	
Results	
Name the correct item(s) as preferable according to the app's results.	

Engagement Record:
Engagement Record: Record session independently. Consider occurrence of any engagement within the 10-s interval. Check data recorded by the participant from *Prefer*, mark correct as "1", and incorrect as "0"

Engagement Score = correct record / (correct record + incorrect record)

Assessment:

Items Available: Record the fraction of items available

Prompts: If there was a prompt record as 1, if not 0

Assessment Score = Items available - prompts

Add the name of the items the were assessed:

Item 1: _____ Item 2: _____ Item 3: _____
 Item 4: _____ Item 5: _____ Item 6: _____

Intervals (5 min)	Engagement Record.						Engagement Score	Assessment		Assessment Score
	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6		Items Available	Prompts	
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										

TOTAL: _____ [(Assessment Score + Engagement Score + Preparation for Assessment & Results)/65]*100

Appendix F
Feedback Questionnaire
(Study 1)

Thank you for participating in this study. Please answer the questions below.

1. In your opinion, how important it is for individuals who work with people with intellectual and developmental disabilities to learn how to conduct preference assessments?

(Not important) 1 2 3 4 5 (Very important)

2. Did you find the video training easy or difficult to understand?

(Very difficult) 1 2 3 4 5 (Very easy)

3. Did you find any part of the video training confusing?

4. How confident do you feel about applying this procedure with clients?

(Not confident) 1 2 3 4 5 (Very confident)

5. How helpful was having the iOS application for conducting the assessment?

(Not helpful) 1 2 3 4 5 (Very helpful)

6. How many times did you watch the full video tutorial?

1 2 more than 2

7. Was there any specific part of the video that you watched more times?

Yes No

If yes, which one?

8. While watching the video, did you practice presenting the items?

Yes No

Appendix G
Demographic Questionnaire
(Study 2)

Clinical Utility Study of *Prefer*
Demographic and Experience Questionnaire

Researcher: Flavia Julio
Phone: 256-4301 ext 5434
Email: juliof@myumanitoba.ca

The purpose of this questionnaire is to gather some demographic information about your education and work experience. The information will help the researcher better interpret the results of the study.

Answer each question by filling in the blanks or selecting the best options among the choices provided. Although it would be most helpful if you could answer all questions, you should feel free to decline answering any of the questions.

We ask you to provide your name in this questionnaire so that we can relate your demographics and educational and work experience to feedback on the *Prefer* app that we will collect from you after you have used the app. We will **not** include your name in any subsequent reports or presentations.

If you have any questions, please contact the researcher listed above.

Date: _____
Name: _____
Age: _____
Sex: Female / Male (circle one)

1. Please check the highest education level completed

High school _____
Community college _____
Undergraduate degree _____
Graduate degree _____
Other (please describe): _____

2. How long have you worked with people with intellectual or developmental disabilities or autism spectrum disorders? _____years_____months

3. Indicate the number of courses on applied behaviour analysis or behaviour analytical procedures that you completed.

Community college courses: _____

University undergraduate courses: _____

University graduate courses: _____

Other (please specify): _____

3. What is your current position? _____

4. What is the name of your program or service or department? _____

4. How long have you worked in your current position? _____ Years _____ months

5. Typically, how many clients do you have contacts with in a one-month period? _____

6. What is the typical age range of the clients you work with?

From: _____ Years _____ months

To: _____ Years _____ months

7. Have you ever received training on how to conduct stimulus preference assessments? Yes / No
If yes, please describe the training and the procedure that you were trained on:

8. For each stimulus preference assessment procedure listed below, indicate how many times (best estimate) you have used that procedure. Enter zero for a procedure if you have not used it before

Multiple-Stimulus-Without-Replacement: _____ times

Multiple-Stimulus-With-Replacement: _____ times

Paired-Stimulus: _____ times

Free-Operant: _____ times

Single-Stimulus: _____ times

Other (please describe): _____

9. How many stimulus preference assessments have you conducted with clients during the last month? _____

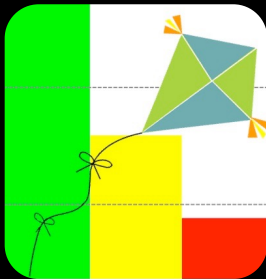
10. How many stimulus preference assessments do you expect to conduct during the next month?

Thank you for completing the questionnaire. Please return it in a sealed envelope to Flavia Julio, c/o St.Amant Research Centre or call her at 256-4301 ext. 5434 and she can pick it up.

Appendix H

Slides from tutorial on how to use *Prefer*
(Study 2)


HOW TO USE PREFER



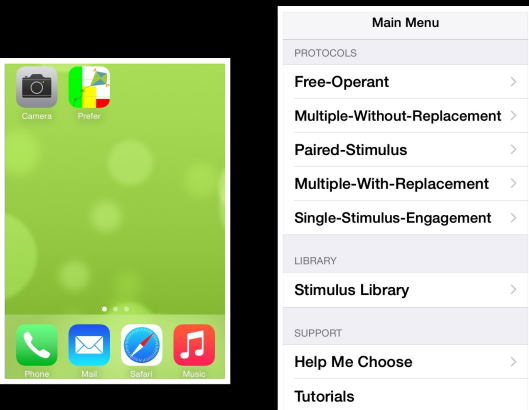
5 quick steps to start a preference assessment on Prefer

Step 1 - Take pictures of the items that will be assessed with the iOS device.

TIP: You may ask for parents and caregivers to name the items that the client might like.

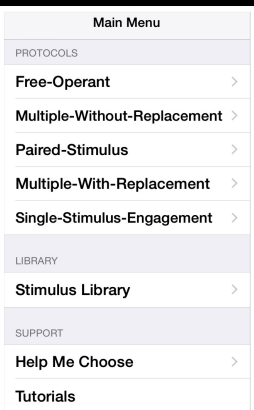


Step 2 - Open Prefer, and choose a preference assessment protocol



Main Menu	
PROTOCOLS	
Free-Operant	>
Multiple-Without-Replacement	>
Paired-Stimulus	>
Multiple-With-Replacement	>
Single-Stimulus-Engagement	>
LIBRARY	
Stimulus Library	>
SUPPORT	
Help Me Choose	>
Tutorials	

Step 2 - Open Prefer, and choose a preference assessment protocol

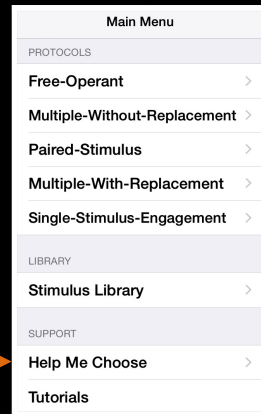


Main Menu	
PROTOCOLS	
Free-Operant	>
Multiple-Without-Replacement	>
Paired-Stimulus	>
Multiple-With-Replacement	>
Single-Stimulus-Engagement	>
LIBRARY	
Stimulus Library	>
SUPPORT	
Help Me Choose	>
Tutorials	

Step 2 - Open Prefer, and choose a preference assessment protocol

Don't know which protocol to choose?

Touch "Help Me Choose", and answer the yes-or-no questions.



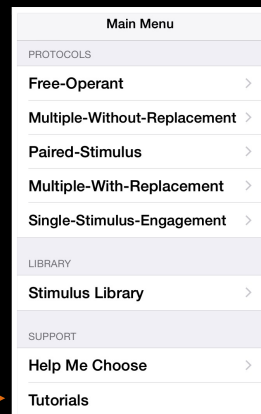
2 - Open Prefer, and Choose a preference assessment protocol



2 - Open Prefer, and Choose a preference assessment protocol

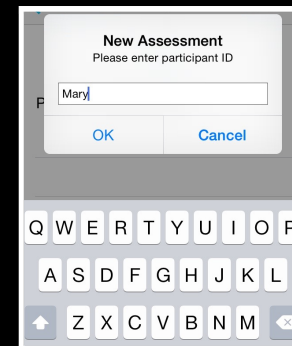
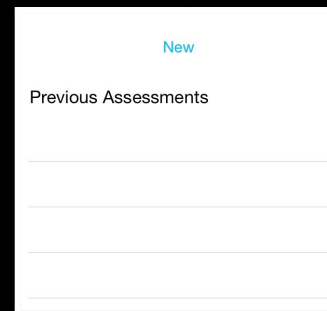
Don't know how to conduct Preference Assessments?

Touch "Tutorials" to launch Safari and see a YouTube playlist on how to conduct preference assessments. Once you are ready, open Prefer again, and choose a Protocol.



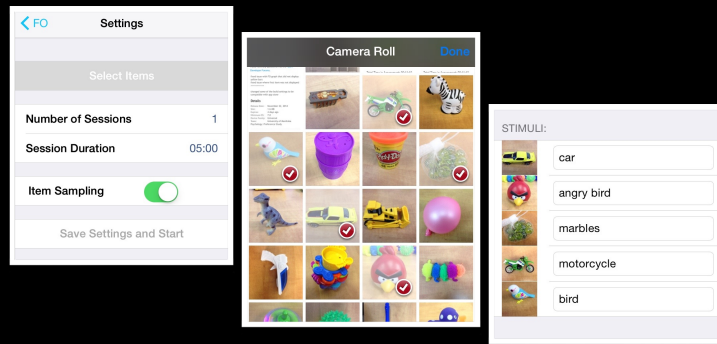
3 - After selecting a protocol, add a new assessment.

Touch "New", then add the client's name/code. And touch "OK."



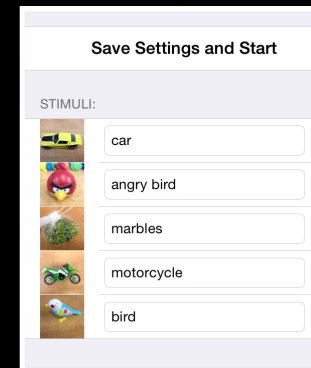
4 - Select the items from your Camera Roll (Step 1), and name it for the assessment.

Touch "Select items", select it, and name it.

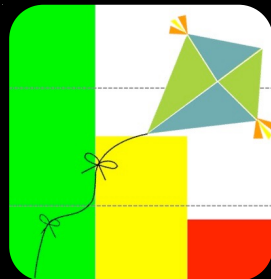


5 - Start the assessment.

Touch "Save Settings and Start."



HOW TO USE PREFER

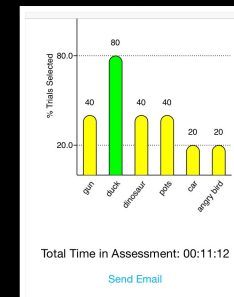


Results

RESULTS

Graphical

You will receive the graphical results of the assessment at the end of the assessment, or when you click on the Participant ID from a completed assessment.



New

Previous Assessments	
Noah May 18, 2015 - 21:15:10 - Complete	
Mary December 31, 1969 - 19:12:51 - Incomplete	

RESULTS

Graphical

Total Time in Assessment: 00:11:12

Send Email

You can also e-mail the assessment results in order to obtain detailed information or to back up the data.

RESULTS

E-mail: Detailed Table

The results will be e-mailed in a .csv format. Where you can obtain information about:

- Items location on each trial
- Item selection on each trial
- Access time
- Latency to select one item
- Intervals with engagement
- Total assessment time.


Example: Paired-Stimulus PA

Participant ID:	Jane 1			
Complete:	YES			
Date:	15-05-26			
Total Time in:	0:16:33			
PA Assessment:	PS			
Max Engagement	30.00s			
Initial Item Order:	Preference [%]			
1. bird	60			
2. bug	60			
3. ball	40			
4. caterpillar	20			
5. connect toy	0			
6. zebra	100			
Trial	Item Description	Selection Time	Engagement Time	Presented Items
1	bug	6.89	33.3	bird; BUG
2	bird	14.56	34.16	ball; BIRD
3	caterpillar	22.17	33.38	connect toy;
4	bird	11.23	34.29	BIRD; connect toy
5	zebra	13.34	33.84	bird; ZEBRA
6	bug	8.47	34.03	BUG; ball
7	bug	11.74	33.35	BUG; caterpillar
8	No item chosen	34.27	0	bug; connect toy
9	zebra	18.48	34.26	bug; ZEBRA
10	bird	11.4	33.43	caterpillar; BIRD
11	ball	9.44	33.64	BALL; caterpillar
12	ball	4.14	33.75	BALL; connect toy
13	zebra	9.87	33.46	ball; ZEBRA
14	zebra	10.03	33.7	caterpillar; ZEBRA
15	zebra	11.42	33.88	connect toy; ZEBRA

Example: Free-Operant PA

Participant ID:	Jane 5			
Complete:	YES			
Date:	15-02-17			
Total Time in Assessment:	0:09:00			
PA Assessment:	FO			
Default Presentation Time:	300.00s			
Initial Item Order:	Preference [%]			
1. pivot	62.62			
2. angry bird	66.25			
3. caterpillar	97.24			
4. motorcycle	0			
5. wand	0.31			
6. connect toy	0.13			
Session	Item Description	Engagement Intervals	Engagement Time	Presentation Time
1	motorcycle		0	300
1	caterpillar	00:07.23 - 00:17.33; 00:17.75 - 00:18.39; 00:19.01 - 05:00.01;	291.73	300
1	wand	02:36.37 - 02:37.32;	0.94	300
1	connect toy	04:47.84 - 04:48.23; 00:11.25 - 00:18.18; 00:18.74 - 00:32.40;	0.4	300
1	angry bird	00:32.83 - 01:08.03; 02:37.03 - 05:00.00;	198.76	300
1	pivot	01:52.13 - 05:00.00;	187.87	300

HOW TO USE PREFER

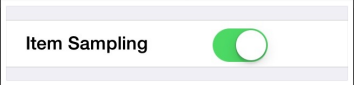


Protocol Variations.

SETTINGS

Change the default configuration of the assessments.

All procedures

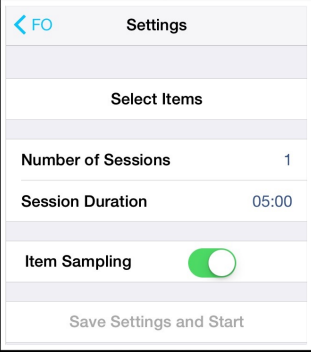


You can turn the prompt to sample the items at beginning of the session ON and OFF.

SETTINGS

Change the default configuration of the assessments.

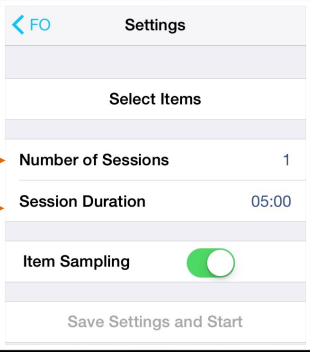
Free-Operant



SETTINGS

Change the default configuration of the assessments.

Free-Operant



Select the number of sessions repetitions.

Select the length of each session

SETTINGS
Change the default configuration of the assessments.

Multiple-Stimulus-with/without-Replacement

MSWO Settings

Select Items

Number of Sessions 5

Access to Selected Item 00:30

Item Sampling

Save Settings and Start

SETTINGS
Change the default configuration of the assessments.

Multiple-Stimulus-with/without-Replacement

Select the number of sessions repetitions.
TIP: A full-MSWO is 5 sessions, a brief-MSWO is 3 sessions.

MSWO Settings

Select Items

Number of Sessions 5

Access to Selected Item 00:30

Item Sampling

Save Settings and Start

SETTINGS
Change the default configuration of the assessments.

Multiple-Stimulus-with/without-Replacement

Select the count-down time for the access time.
TIP: If you are assessing edibles select "0", removing the prompt.

MSWO Settings

Select Items

Number of Sessions 5

Access to Selected Item 00:30

Item Sampling

Save Settings and Start

SETTINGS
Change the default configuration of the assessments.

Paired-Stimulus

PS Settings

Select Items

Permutate? No

Access to Selected Item 00:30

Item Sampling

Save Settings and Start

SETTINGS

Change the default configuration of the assessments.

Paired-Stimulus

Would you like to present the items in all possible positions?
(e.g., A - B and B - A)

< PS
Settings

Select Items

Permutate? No

Access to Selected Item 00:30

Item Sampling

Save Settings and Start

SETTINGS

Change the default configuration of the assessments.

Paired-Stimulus

Select the count-down time for the access time.
TIP: If you are assessing edibles select "0", removing the prompt.

< PS
Settings

Select Items

Permutate? No

Access to Selected Item 00:30

Item Sampling

Save Settings and Start

SETTINGS

Change the default configuration of the assessments.

Single-Stimulus-Engagement

< SS-E
Settings

Select Items

Number of Sessions 5

Access to Selected Item 02:00

Item Sampling

SETTINGS

Change the default configuration of the assessments.

Single-Stimulus-Engagement

Number of times all items will be presented.

< SS-E
Settings

Select Items

Number of Sessions 5

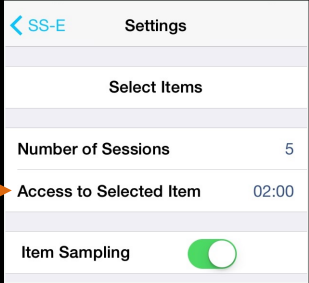
Access to Selected Item 02:00

Item Sampling

SETTINGS
Change the default configuration of the assessments.

Single-Stimulus-Engagement

Amount of time that each item will remain available during a session. →



Settings	
Select Items	
Number of Sessions	5
Access to Selected Item	02:00
Item Sampling	<input checked="" type="checkbox"/>

Appendix I
Clinical Utility Questionnaire
(Study 2)

**Clinical Utility Study of *Prefer*
Feedback Questionnaire**

Researcher: Flavia Julio
Phone: 256-4301 ext 5434
Email: juliof@myumanitoba.ca

The purpose of this questionnaire is to collect information about your experience using the stimulus preference assessment app *Prefer*. This questionnaire should take about 15 minutes to complete.

We ask you to provide your name in this questionnaire so that we can relate your experience to your educational and work experience that you provided at the beginning of this study. We will **not** include your name in any subsequent reports or presentations.

Answer each question by filling in the blanks or selecting the best options among the choices provided. Although it would be most helpful if you could answer all questions, you should feel free to decline answering any of the questions.

If you have any questions, please contact the researcher listed above.

Date: _____
Name: _____

HELP ME CHOOSE FEATURE

On the main screen of the application, there is a “Help Me Choose” icon that leads to a series of yes or no questions, and provides you with a suggestion of preference assessment procedure from the application, when applicable.

1. Rate the helpfulness of this feature by circling a number from 1 to 5. Circle n/a if you are unable to rate the feature (e.g., you did not use it).

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments: _____

VIDEO TUTORIALS

On the main screen of Prefer there is a “Tutorial” icon that connects you to a YouTube channel which contains video tutorials for three stimulus preference assessment procedures: paired-stimulus, multiple-stimulus-without-replacement, and free-operant.

Paired-Stimulus (PS) Stimulus Preference Assessment Video Tutorial

2. How many times did you watch the PS video tutorial? _____

3. Was the PS video tutorial difficult or easy to understand?

n/a	1	2	3	4	5
Unable to rate	Very difficult				Very easy

Comments (if any): _____

4. Do you believe that you conducted the PS preference assessment with your client(s) in the same way as you saw in the video tutorial?

n/a	1	2	3	4	5
Unable to rate	Very different				Very similar

Comments (if any): _____

Multiple-Stimulus-without-Replacement (MSWO) Preference Assessment Video Tutorial

5. How many times did you watch the MSWO video tutorial? _____

6. Was the MSWO video tutorial difficult or easy to understand?

n/a	1	2	3	4	5
Unable to rate	Very difficult				Very easy

Comments (if any): _____

7. Do you believe that you conducted the MSWO preference assessment with your client(s) in the same way as you saw in the video tutorial?

n/a	1	2	3	4	5
Unable to rate	Very different				Very similar

Comments (if any): _____

Free-Operant (FO) Preference Assessment Video Tutorial

8. How many times did you watch the FO video tutorial? _____

9. Was the FO video tutorial difficult or easy to understand?

n/a	1	2	3	4	5
Unable to rate	Very difficult				Very easy

Comments (if any): _____

10. Do you believe that you conducted the FO preference assessment with your client(s) in the same way as you saw in the video tutorial?

n/a	1	2	3	4	5
Unable to rate	Very different				Very similar

Comments (if any): _____

PROMPTS DURING ASSESSMENTS

11. **SAMPLE EACH ITEM:** The app prompts you to let the client sample the items before an assessment, did you find the reminder helpful?

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments (if any): _____

12. **PICTORIAL REPRESENTATION:** During the assessment, *Prefer* shows a picture to represent the items that should be presented on each trial. How helpful was it to have a pictorial representation of these items (i.e., having the picture of the item as opposed to have the name)?

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments (if any): _____

13. **ITEM LOCATION:** When applicable, during an assessment, *Prefer* will rotate the pictures after a trial, and show a number below the pictures representing the location that they should be presented. How helpful was it to be able to see the location of the items in *Prefer* (i.e., in addition to which item to present)?

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments (if any): _____

14. **NEXT TRIAL ITEM LOCATION:** For the PS procedure, a box on the lower half of the screen would show which items should be presented on the following trial, and its location. How helpful was it to be able to see the location of the items for the next trial?

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments (if any): _____

15. ACCESS TIME. For the PS and MSWO procedures, after a selection is made, a window pops-up prompting the user to allow the client to have access to the selected item, and offers a timer. How helpful was this prompt?

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments (if any): _____

16. RESULTS. After a preference assessment is concluded, *Prefer* shows a graph with the results. How helpful was this graph?

n/a	1	2	3	4	5
Unable to rate	Not helpful				Very helpful

Comments (if any): _____

Back-up data

You can e-mail the results of a preference assessment and have it as a backup, in addition to a more detailed record of preference assessment.

17. How often did you use this feature?

n/a	1	2	3	4	5
Unable to rate	Rarely				After every assessment

Comments (if any): _____

18. How important do you think it is to be able to back-up the results of the assessment (i.e., to be able to export the result to a different device/computer)?

n/a	1	2	3	4	5
Unable to rate	Not important				Very important

Comments (if any): _____

19. If you have used the backup feature and reviewed the data file for additional information, please rate the usefulness of the following information.

(a) Items presented on each trial or session with the location.

n/a	1	2	3	4	5
Unable to rate	Not useful				Very useful

(b) Item selected on each trial or interval (i.e., item with the name capitalized)

n/a	1	2	3	4	5
Unable to rate	Not useful				Very useful

(c) Engagement time

n/a	1	2	3	4	5
Unable to rate	Not useful				Very useful

(d) Latency to choose an item (i.e., how long it took tester to record the selection on each trials)

n/a	1	2	3	4	5
Unable to rate	Not useful				Very useful

(e) Duration of the assessment

n/a	1	2	3	4	5
Unable to rate	Not useful				Very useful

FUTURE USAGE

20. Do you plan to continue to use the app to conduct preference assessment in the future if an iOS device is available?

n/a 1 2 3 4 5
 Unable to rate Very unlikely Very Likely

Comments (if any): _____

21. Would you recommend *Prefer* to a colleague?

n/a 1 2 3 4 5
 Unable to rate Very unlikely Very Likely

Comments (if any): _____

22. Would you recommend the video tutorial to colleagues who have not received training yet?

n/a 1 2 3 4 5
 Unable to rate Very unlikely Very Likely

Comments (if any): _____

DID *PREFER* HELP YOU IDENTIFY REINFORCERS

23. Were the assessment results useful in helping you to identify reinforcer(s) for your interventions?

n/a 1 2 3 4 5
 Unable to rate Not useful Very useful

Comments (if any): _____

OTHER COMMENTS

Please provide any other comments you may have.
