5-11-2023

Using free-operant preference assessments and reinforcer evaluations to measure the stability of preferences over time in individuals with developmental disabilities

Lindsey Histand
James Madison University

Using Free-Operant Preference Assessments and Reinforcer Evaluations to Measure the Stability of Preferences Over Time in Individuals with Developmental Disabilities

Lindsey Histand

A thesis submitted to the Graduate Faculty of

James Madison University

In

Partial Fulfillment of Requirements

For the Degree of

Master of Arts

Psychological Sciences, Behavior Analysis Concentration

James Madison University

May 2023

FACULTY COMMITTEE

Committee Chair: Trevor F. Stokes, Ph.D., BCBA-D

Krisztina V. Jakobsen, Ph.D.

Tracy E. Zinn, Ph.D.

Acknowledgements

I would like to express my appreciation to Dr. Trevor Stokes for being a knowledgable supervisor and his willingness to share his clinical knowledge with me during my time at James Madison University. His strong support of my success both professionally and personally has provided comfort during difficult and hectic times while working together. I also would like to thank Dr. Krisztina Jakobsen and Dr. Tracy Zinn for being apart of my thesis committee and taking the time to provide feedback and guidance related to this study. Lastly, I am immensely appreciative of Sydney Setchel, Haley Gardner, and Morgan Mabery for all the time and attention they spent as research assistants for this study. They all showed great attention to detail and knowledge of the ABA principles and concepts as they were applied in this study. You will do amazing things in your careers and I am excited to watch you grow as practitioners in the field of ABA!

Table of Contents

List of Figures

Abstract

The present study investigated the stability of preferences over time using an alternating treatment multielement design across participants. The participants were children between the ages of eight and twelve with developmental disabilities, who also displayed interfering behaviors. The study had a baseline condition and two experimental conditions. During the first condition, the therapist provided contingent reinforcement for correct responses during reinforcer evaluations using highest preferred items. During the second condition, the therapist provided contingent reinforcement for correct responses during reinforcer evaluations using least preferred items. Repeated measures of free-operant preference assessments produced participants' preference rankings, thus identifying highest and least preferred items to be used during the reinforcer evaluations. This study identified that preferences remained stable for four and a half weeks on average, across both participants. It also identified that as preferences became unstable, the participants' performance decreased during the highest preferred reinforcer evaluation condition and/or increased during the least preferred reinforcer evaluation condition.

Keywords: Free-operant, preference assessments, reinforcer ranking, reinforcer evaluation, preference stability, interfering behavior

Introduction & Literature Review

The systematic use of reinforcers is a vitally important aspect of behavior interventions implemented by Applied Behavior Analysis professionals (Leaf, et al., 2015). Once a clinician has conducted a preference assessment, and has identified the client's most desirable items, the items can be used as reinforcement for quality performance and desirable behaviors (Johnson, et al., 2017). This is most commonly used with individuals with Autism Spectrum Disorder (Leaf, et al., 2015), but could also be used with neurotypical individuals or individuals with neurodegenerative disease or other developmental disabilities.

Throughout our lifetime, our preferences change depending on our life experiences and motivating operations (Hanley, et al., 2006). For some, preferences change rather quickly and for others preferences stay generally the same over longer periods of time. In order for clinicians to always offer the highest preferred reinforcers to their clients, clinicians must have a *current* list of highly reinforcing items. This brings the question: How often should clinicians conduct preference assessments to maintain an updated list of highly preferred reinforcers? Having an updated reinforcer ranking is essential to set up individuals for success because using reinforcers that are less preferred may have a negative effect on individuals' performance of tasks (Carr, 2000). This could prevent the individual from progressing in skill acquisition or behavioral goals. Especially with individuals with developmental disabilities, it is vital that individuals progress in their goals so that they can increase their level of independence and improve their quality of life.

Previous research has been conducted studying variables related to preference assessments, such as different methods of preference assessments, reinforcer stimuli selection,

and interfering behavior during preference assessments. Within this review, these variables will be discussed in more detail, and current research will be presented about the current knowledge of preference change over time. As stated earlier in this section, preference assessments can be used to identify reinforcers for individuals with neurodegenerative disease, other developmental disabilities, and neurotypical individuals; however, the focus of the current research is with children who have Autism Spectrum Disorder, therefore research related to alternative populations will not be discussed in length in this review.

Preference Assessment Methods

There are three main preference assessment methods that can be implemented. The three main methods of preference assessment are free-operant, paired-stimulus, and multiple-stimulus assessments. Each method has advantages and limitations that need to be considered carefully when selecting which type of assessment to use with an individual. The advantages and limitations must also be compared to the individual's current repertoire of skills, as certain skills are required for each type of assessment.

Paired-Stimulus Preference Assessment

The first type of assessment is paired-stimulus, in which the participant is presented with an array of two items. After the participant chooses and interacts with an item, the participant is presented with a new array of items from which to choose (Kang, et al., 2010). An advantage of this method is that each item is presented to the participant in multiple combinations. This increases the accuracy of the results because the participant comes into contact with each item in multiple pairings (Verriden & Roscoe, 2016). A limitation of this

method is that it takes longer to complete the assessment using this method in comparison to other preference assessment methods (Roane, et al., 1998).

Multiple-Stimulus Preference Assessment

The next assessment type is multiple-stimulus, which is similar to paired-stimulus, except the participant is presented with an array of three or more items instead of just two. Similarly to paired-stimulus, after the participant chooses and interacts with an item, the participant is presented with a new array of items from which to choose (Curiel & Poling, 2019). There are two subtypes of multiple-stimulus preference assessment. The first type, multiple-stimulus *without* replacement, requires that after an item is chosen from the array, it is not put back into successive arrays. Meanwhile, the second type, multiple stimulus *with* replacement, allows for all items chosen to be placed back into successive arrays.

Free-Operant Preference Assessment

Lastly, is free-operant preference assessments. For this type of assessment, the participant is given free-rein to interact with any of the items present in the environment for as long or as short as they would like over a predetermined amount of time (Ortiz & Carr, 2000). An advantage of free-operant preference assessments is that items are never withdrawn from the participant. Because of this, free-operant preference assessments are associated with fewer instances of interfering behavior (Ortiz & Carr, 2000), which suggests this method is best for individuals who exhibit self-injurious behavior or behavior that results in injuring others. Additionally, practitioners can set a shorter amount of time to conduct free-operant assessments to decrease the amount of time spent on identifying preference rankings (Roane, et al., 1998). A limitation of this method is that there is a higher chance of false negative

results because the participant may never interact with all the items presented (Verriden & Roscoe, 2016).

Roane and colleagues (1998) conducted assessments with individuals with developmental disabilities to analyze if brief (five minute) free-operant preference assessment methods were accurate in identifying preference rankings. The stimuli identified as preferred reinforcers during the study were differential effective reinforcers in comparison to non-preferred stimuli. They also found that less interfering behavior was exhibited during the free-operant assessments, and the free-operant assessments took less time, in comparison to paired-stimulus assessments.

Verriden and Roscoe (2016) compared correlation coefficients and Kendall rank coefficients of concordance across paired-stimulus, multiple-stimulus, and free-operant preference assessments. The highest correspondence was found for the paired-stimulus and multiple-stimulus methods, while the lower correspondence was found for the free-operant method. These findings suggest that there was more variability across administrators during the free-operant assessments. This was likely caused by multiple items tying for the least preferred stimuli ranking because those items were not engaged with at all during the assessments.

A study from 2000 by Oriz and Carr compared free-operant and multiple-stimulus preference assessments. Their study found that both methods resulted in stimulus rankings that were virtually identical; however, the free-operant method was only able to identify the most highly desirable items because the participants did not forcibly interact with each item, unlike with multiple-stimulus preference assessments. These results, in addition to the results

found in Verriden and Roscoes' (2016) study, exemplify the limitation previously mentioned in this section of false negative results resulting from the opportunity of no engagement with items included in free-operant assessments.

The present study used this type of assessment to determine preference rankings of the participants. This type of assessment was chosen because they have a decreased risk of evoking interfering behavior in comparison with paired-stimulus and multiple-stimulus methods (Kang, et al., 2010). This benefit mitigated the risk of participants engaging in an increase of interfering behaviors during the assessments. They also take less time to complete in comparison to other assessment methods (Roane, et al., 1998). This benefit allowed researchers to maximize their time since preference assessment is only one element of the study. Free-operant assessments have also been able to identify similar preference rankings as other methods, which suggests that free-operant assessments are comparable in preference identification (Roane, et al., 1998).

Reinforcer Evaluation

The purpose of reinforcer evaluations are to measure reinforcer effectiveness. This is often used following a preference assessment to identify the accuracy of the preference ranking identified by a preference assessment. Two studies by Piazza and colleagues (1996) and Roane and colleagues (1998) used similar reinforcer evaluations to assess the accuracy of preference rankings discovered by various preference assessment methods. The results of these two studies found that the preference rankings discovered during the preference assessments were able to relatively accurately rank the reinforcing value of various stimuli, based on the findings of the reinforcer evaluations.

Interfering Behavior

Interfering behavior, such as inattentiveness, may lead to interruption of the assessment, thus increasing the likelihood of skewed results. Interfering behavior occurs when a child or adult finds that exhibiting interfering behavior provides them access to items or activities, or when the interfering behavior is reinforced by socially mediated reinforcement (Kang, et al., 2010). Certain types of preference assessments can cause more interfering behavior than others (Kang, et al., 2010).

According to a study published by Kang and colleagues in 2010, tangible positive reinforcement is often the cause of interfering behavior during preference assessments. Their research found that most interfering behavior occurred during paired-stimulus and multiple-stimulus preference assessments. They also found that the least interfering behavior occurred during free-operant preference assessments. Another study conducted by Roane and colleagues in 1998 found similar effects of decreased interfering behavior during free-operant assessments in comparison to paired-stimulus assessments. A third study by Verridan and Roscoe (2016) also found that less interfering behavior occurred during free-operant assessments.

The most likely reason for the aforementioned findings is that during paired-stimulus and multiple-stimulus preference assessments, the item chosen by the participant from an array must be taken away after a predetermined amount of time in order to continue the assessment (Kang, et al., 2010). This can be an aversive effect of preference assessments for the participants, especially for individuals who are non-verbal, because they often want to continue interacting with the highly desirable item they have just chosen. This is a concern for

many professionals because some displayed interfering behaviors can physically hurt the individual engaging in them or those around them. Meanwhile, during free-operant preference assessments, the individual has free rein to interact with any and all of the items presented for as long as desired during the duration of the assessment. This is why free-operant assessments may be a better option for individuals who engage in severe interfering behavior (Ortiz & Carr, 2000).

Preference Stability

Changes in preference over time are often variable depending on the individual as well as many other aspects (Hanley, et al., 2006). Since preference can be affected by so many variables, it has been difficult for researchers to identify the variables that causes preference change. However, Hanley and colleagues' study (2006) was able to identify some potential causes of preference change: satiation, paired access to low-preference reinforcers, changes in establishing operations, and conditioning histories. Satiation occurs when an individual no longer finds an item as desirable because of continuous, free-access to the item. Paired access to low-preference reinforcers is when a low-preference item is presented with another more desirable item, thus the low-preference item becomes more desirable. Changes in establishing operations occur in the natural environment, which causes momentary alterations in the reinforcing value of an item.

A study by Carr and colleagues (2000) found that preference rankings of two out of three participants remained relatively stable over the course of four weeks when preference rankings were evaluated every two to five days using multiple-stimulus without replacement preference assessments. These findings suggest that many of the participants' preferences

remained stable over the course of four weeks; whereas, one of the participants showed variability in preference over the course of four weeks.

Verriden and Roscoes' study (2016) compared initially identified and recently identified high-preference stimuli using reinforcer assessments. They found that the identified and recently identified stimuli were equally effective reinforcers during the single-operant schedule of the reinforcer assessment. This suggests that changes in preference does not necessarily affect reinforcer efficacy, thus does not affect individuals' performance. However, higher levels of responding occurred with the recently identified stimuli than with the initially identified stimuli during the concurrent-schedule of the reinforcer assessment. The concurrent-schedule findings of Verriden and Roscoe's (2016) study was a replication of the findings by DeLeon and colleagues' original study (2001).

Preference Assessment Frequency

Part of the purpose of this study was to identify how often preference assessments should be conducted to maintain an updated reinforcer ranking list, while not conducting them too frequently or too infrequently. A study by Butler and Graff (2021) conducted monthly paired-stimulus preference assessments over a one year period. Their study suggests that preferences were more stable when assessed over shorter periods of time (monthly), rather than longer periods of time (yearly). Based on these findings, the researchers suggested that practitioners should conduct preference assessments more frequently with new clients to identify how quickly their preferences change. Then, increase the amount of time between assessments, depending on their findings from the initial assessments.

Purpose of Present Study

After completing her undergraduate thesis research (Histand, 2021), the primary researcher identified preference change over time as an area needing future research. Up until this point, research has been able to identify that preferences remain generally stable over the course of four weeks (Carr, et al., 2000); however, little research has been conducted to identify how long that stability persists beyond four weeks. More research on preference changes and individuals with developmental disabilities needs to be conducted. It is important to further this research to better understand how often professionals should conduct preference assessments with their clients to have the most accurate preference rankings, without wasting service time on conducting preference assessments more often than necessary. If practitioners are not using highest preferred stimuli, their clients might not progress in skill acquisition or behavioral goals. Especially with individuals with developmental disabilities, it is vital that individuals progress in their goals so that they can increase their level of independence and improve their quality of life.

Considering that a contraindication of using certain preference assessment methods includes an increased risk of evoking interfering behavior, it is important that preference assessments are not performed too frequently to mitigate this risk. Conversely, clinicians should not wait too long to conduct a new preference assessment because offering a client a lesser preferred item could cause a decrease in performance; thus, leading to a slower progression in reaching their goals and increasing their independence and quality of life.

The present study used free-operant preference assessments and reinforcer evaluations to track preference stability over time. Interfering behavior was also tracked to consider the

affect preference change has on the frequency of interfering behavior. Free-operant assessments were chosen to be used in this study because they have a decreased risk of evoking interfering behavior (Kang, et al., 2010) and take less time to complete in comparison to other assessment methods (Roane, et al., 1998). Free-operant assessments have also been able to identify similar preference rankings as other methods, which suggests that free-operant assessments are comparable in preference identification (Roane, et al., 1998).

The current study hypothesized that minor changes in reinforcer rankings would occur over the course of the study, with moderately preferred rankings being more variable. Interfering behavior was hypothesized to fluctuate throughout the study depending on which type of session was conducted (preference assessment, highest preferred reinforcer evaluation, and least preferred reinforcer evaluation) and the function of the behaviors. The current study also hypothesized that stability in preferences would persist for six to eight weeks. At that point, performance during the reinforcer evaluations were hypothesized to decrease during highest preferred sessions and/or increase during least preferred sessions. This hypothesis was made based on trends identified during the primary researcher's undergraduate thesis research. Preference stability was not a focus of that study, however it appeared that the participants' highest and least preferred items remained consistent for six to seven weeks on average.

Method

Participants

As stated in Leaf and colleagues' study (2015), preference assessments are most commonly used with individuals with Autism Spectrum Disorder. The current study included

two male participants, ages eight and twelve, with Autism Spectrum Disorder, who displayed behavior challenges. Assent/consent was obtained by each participant/the participants' parents prior to starting the study in accordance with James Madison University's Institutional Review Board's approved protocol.

Setting

Two clinic rooms were used for this study, which were connected by a one-way mirror. The observation room is where researchers looked through the one-way mirror to observe the participant-therapist interactions and collect data for the study. The observation room was 2.75 meters wide by 2.75 meters long. The treatment room displayed the mirror side of the one-way mirror and is where the participant and therapist interacted for the purpose of the study. The treatment room was also 2.75 meters wide by 2.75 meters long. The treatment room included a couch, a child-sized, round table with three child-sized chairs, and a rollable set of drawers. All furniture items were positioned around the perimeter of the treatment room to create more open space in the middle of the room for the participants to engage with the stimuli used during the preference assessments.

Independent Variable

The stimuli used to provide contingent reinforcement during the reinforcer evaluations changed in each condition. During baseline, non-contingent reinforcement was provided for both correct and incorrect responses. An example of this is saying, "Thanks for answering" with a neutral tone after each response, regardless of whether the individual provided a correct answer. In the first intervention condition, highest preferred stimuli identified by the free-operant preference assessments was used to provide contingent reinforcement for correct

responses. In the second intervention condition, least preferred stimuli identified by the free-operant preference assessments were used to provide contingent reinforcement for correct responses.

Dependent Variable

The dependent variable of the present study was the stability of preferences over time as demonstrated by shifts in preference rankings produced by free-operant preference assessments and performance during reinforcer evaluations. Interfering behavior was also assessed as part of the preference assessments and reinforcer evaluations to discover whether preference changes affected the frequency of interfering behavior.

Data Collection

During each assessment session, data was collected on the participants' interactions with the items included in the free-operant preference assessment to identify preference rankings. The preference rankings were identified based on the frequency of the participant approaching each item and the amount of time the participant engaged with each item. Approached was defined as the participant coming within one meter of an item and engaged was defined as the participant touching an item. Highest preferred items were approached most frequently and engaged with for the longest duration. Moderately preferred items were approached but engaged with for the shortest duration. Least preferred items were not approached or engage with at all.

Data was also collected on the reinforcer effectiveness of the highest and least preferred items identified by the free-operant preference assessments during the reinforcer evaluations. Data collectors recorded frequency data of interfering behaviors exhibited during each portion

of the sessions. The interfering behaviors tracked during the study were individualized to the participant's most commonly exhibited interfering behaviors, which were reported during the parent interviews.

Data sheets were provided to the data collectors during each session. One data sheet was used to record data during, and to score, the baseline reinforcer evaluations (Appendix A). A second data sheet was used to record data during, and to score, the free-operant preference assessments (Appendix B). A third data sheet was used to record data during, and to score, the reinforcer evaluations (Appendix C). A third data sheet was used to record the frequency of interfering behavior emitted during each session (Appendix D). There was a primary and secondary data collector to assess Interobserver Agreement. Data was displayed and analyzed graphically using Apple Numbers.

Research assistants participated in training from the primary researcher on how to code data for the current study. The primary researcher reviewed the Procedural Checklists with the assistants and provided training about how to complete the form. The primary researcher defined and modeled instances of approached and engaged behaviors included in free-operant preference assessments. The primary researcher also defined and model correct and incorrect responses included in the reinforcer evaluation. Lastly, the primary researcher provided functional definitions of the interfering behaviors being tracked for each participant during the duration of the study. Each research assistant watched videos depicting a free-operant preference assessment, a reinforcer evaluation, and interfering behaviors and were required to score the mock data to affirm that the research assistants were trained to criteria. The research

assistants were required to meet 90% Interobserver Agreement with the primary researcher using the total count method in order to meet mastery criteria.

The sessions were recorded using Video Audio Learning Tool (VALT), which is downloaded onto a secure computer that has no access to the internet. The computers, VALT software, and recordings were encrypted and password protected. The treatment room had two cameras and two microphones that transmitted visual and audio to a computer that contains the VALT software, where the sessions were recorded. The recordings were used by the primary researcher and research assistants for data collection purposes.

Interobserver Reliability

Interobserver Agreement was assessed across 100% of sessions during each phase of the current study. Interobserver Agreement was calculated for each of the dependent variables using the total count method. This was calculated by the primary researcher, who took the smaller number of reported responses and divided it by the larger number of reported responses, then multiplied that number by 100, which is outlined in Johnston and colleagues' book (2020). For example, if one of the data collectors reported 15 instances of interfering behavior, while the other data collector reported 16, the primary researcher would divide 15 by 16 to get 94% IOA. For the current study, sessions that receive 80% or higher Interobserver Agreement were considered acceptable.

With Paxton, percentage of agreement for the frequency of interfering behaviors was 82% during the baseline condition, 81% during preference assessment sessions, and 91% during reinforcer evaluation sessions. Percentage of agreement for preference assessment

rankings was 94%. Percentage of agreement for the rate of correct responses per minute was 97% during the baseline condition and 96% during reinforcer evaluation sessions.

With Rowan, percentage of agreement for the frequency of interfering behaviors was 83% during the baseline condition, 94% during preference assessment sessions, and 88% during reinforcer evaluation sessions. Percentage of agreement for preference assessment rankings was 94%. Percentage of agreement for the rate of correct responses per minute was 86% during the baseline condition and 95% during reinforcer evaluation sessions.

Experimental Design

The present study assessed stability of preferences over time using an alternating treatment, multielement design, across participants. Each participant experienced a reinforcer evaluation in baseline and subsequent intervention conditions. Using this design demonstrated experimental control by exhibiting that responding changed when there was a change in conditions between non-contingent reinforcement (baseline) and contingent reinforcement provided using highest (intervention condition A) and least (intervention condition B) preferred items (Johnston et al., 2020). As an additional source of experimental control, the primary researcher ensured that the stimuli used during the intervention conditions were presented an equal amount of times during the reinforcer evaluations. This experimental design is similar to the design used in Francisco and colleagues' study from 2008.

When analyzing the graphs, the primary researcher examined changes in level, trend, variability, and stability. A free-operant preference assessment was automatically repeated four weeks after the initial assessment to demonstrate preference stability over the course of four weeks, in accordance with the findings of previous research (Carr, et al., 2000).

Response guided decision-making was used to decide when the participants participated in subsequent free-operant preference assessments. The following is criteria was used to decide when to repeat the free-operant preference assessments: 1) A decrease in performance during the highest preferred reinforcer evaluation for two consecutive highest preferred sessions; 2) An increased performance during the least preferred reinforcer evaluations when compared to the performance during the most recent highest preferred reinforcer evaluations for two consecutive least preferred sessions; and 3) An increase in performance during the least preferred reinforcer evaluation for two consecutive least preferred sessions.

Pre-Experimental Procedures

During the intake appointment with the parent(s) of the participant, the primary researcher conducted a parent interview to identify the participant's preferred tangible items and activities by using a reinforcer survey (Appendix E), which was adapted from two commonly used reinforcer surveys: 1) The Reinforcement Assessment for Individuals with Severe Disabilities (Fisher, et al., 1996); and 2) The Reinforcement Inventories for Children and Adults (Institute for Applied Behavior, 1993). The parent-indicated preferred items and activities better informed the primary researcher of which items to choose during item selection for the free-operant preference assessments.

Additionally, the primary researcher enquired about mastered and unmastered tasks (i.e., motor imitations, tacting, listener-responding, etc.) of the participant during the parent interview. This information was used to identify mastered skills to be targeted during the alternative activities, and unlearned skills to be targeted during reinforcer evaluations. Lastly, the primary researcher included questions in the parent interview regarding the participant's

interfering behaviors. This informed the researcher about interfering behaviors the participant displayed that could be tracked throughout the study, and provide information to functionally define those behaviors.

Prior to starting the current study, the primary researcher collected baseline probe data on the mastered and unlearned skills identified during the parent interview. Skills that receive a correct percentage of 21-79% were not used during the present study. Skills that receive a correct percentage of 80% or higher during the baseline probe were considered mastered skills, and skills that receive a correct percentage of 20% or less during the baseline probe were considered unlearned skills. The skills identified as mastered were used during the alternative activity portion of the free-operant assessment sessions and the unlearned skills were used during the reinforcer evaluation sessions.

Materials

The materials used throughout the study depended on the information provided during the parent interview. The primary researcher identified four different items (i.e., toy cars, balls, musical toys, etc.) per participant to be used throughout the course of the study during the free-operant preference assessments. During the preference assessments, the items were spaced out around the treatment room for the participant to have direct access to all of the items. The items were spaced 1 to 2 meters apart. Additionally, the items were placed in different locations around the room to prevent side bias from occurring.

Baseline Reinforcer Evaluation

The researcher conducted three consecutive 10-minute baseline reinforcer evaluations of the chosen target behaviors across participants during the first session, separated by

engagement in an alternative activity for three to five minutes. The reinforcer evaluations during the baseline phase were shorter in time in comparison to the intervention reinforcer evaluations to mitigate the risk of potential increases in interfering behaviors due to the absence of reinforcement during the baseline phase. The target behaviors used during the baseline condition were the unlearned tasks or activities identified during the parent interview.

Free-Operant Assessment Sessions

After the baseline condition, the participants underwent two consecutive 10-minute free-operant preference assessments, which were scored together. A second free-operant preference assessment was repeated four weeks after the initial assessment. Subsequent free-operant preference assessments were decided using response guided decision-making. All free-operant preference assessments were conducted using the following method.

The primary researcher set up the treatment room with four of the participant's preferred items and activities, as identified by the parent interview during the intake appointment. The items were spread around the room so that the child has easy access to each item. The primary researcher brought the participant to the treatment room and said, "You have ten minutes to play with whichever toys you want." During the assessment, the primary researcher stayed in the room but did not interact with the participant or the items in the room.

After the first 10-minute preference assessment concluded, the primary researcher prompted the participant to sit in the child's chair located at the child-sized table in the treatment room. The primary researcher conducted an alternative activity with the participant, which included mastered tasks as identified during the parent interviews. After three to five minutes of the alternative activity, the participant was prompted to do a second free-operant

preference assessment. The researcher conducted the second free-operant assessment in the same way as the first assessment. After both 10-minute free-operant assessments were completed, the data collectors scored both of the 10-minute assessments together.

Intervention Reinforcer Evaluation Sessions

The target behaviors used during the intervention conditions were the same unlearned tasks or activities used during the baseline reinforcer evaluation. The researcher provided contingent reinforcement while conducting two consecutive 10-minute reinforcer evaluations, separated by three to five minutes of an alternative activity, which included mastered tasks as identified during the parent interviews. The two intervention conditions alternated every week. The first intervention condition provided contingent reinforcement using the highest preferred reinforcer identified during the free-operant preference assessment. The second intervention condition provided contingent reinforcement using the least preferred reinforcer identified during the free-operant preference assessment.

For each correct response, the participant was provided 10-second access to the stimuli assigned to that condition on a fixed-ratio 1 schedule. For each incorrect response, the participant was not provided access to the stimuli assigned to that session. Instead, the therapist implemented an error correction procedure: 1) Re-present targeted skill with an immediate prompt providing the correct response, and 2) Participant emits correct response. If the participant did not emit the correct response, the therapist prompted the participant to engage in a motor imitation and then represented the error correction. Correct answers emitted during the error correction procedure were not counted as correct responses during

data collection. The methods described above are similar to the reinforcer evaluation methods used in Carr and colleagues' study (2000) and Francisco and colleagues' (2008) study.

Procedural Fidelity

Procedural fidelity data was taken throughout the duration of the present study. This was an essential aspect of this study to ensure that all procedures of the study were being implemented as described in the session protocols. A procedural checklist (Appendix F), similar to the one described in Roscoe and Fisher's study (2008), was used to collect procedural fidelity data during each session. Procedural checklist items used during the current study pertained to procedures such as placement of the items in the treatment room, the researcher's behavior during the preference assessment, and conducing the preference assessments and reinforcer evaluations for the appropriate duration.

With Paxton, the therapist conducted the baseline condition with 100% fidelity, the preference assessment sessions with 95% fidelity, and the reinforcer evaluation sessions with 100% fidelity. With Rowan, the therapist conducted the baseline condition with 100% fidelity, the preference assessment sessions with 100% fidelity, and the reinforcer evaluation sessions with 100% fidelity.

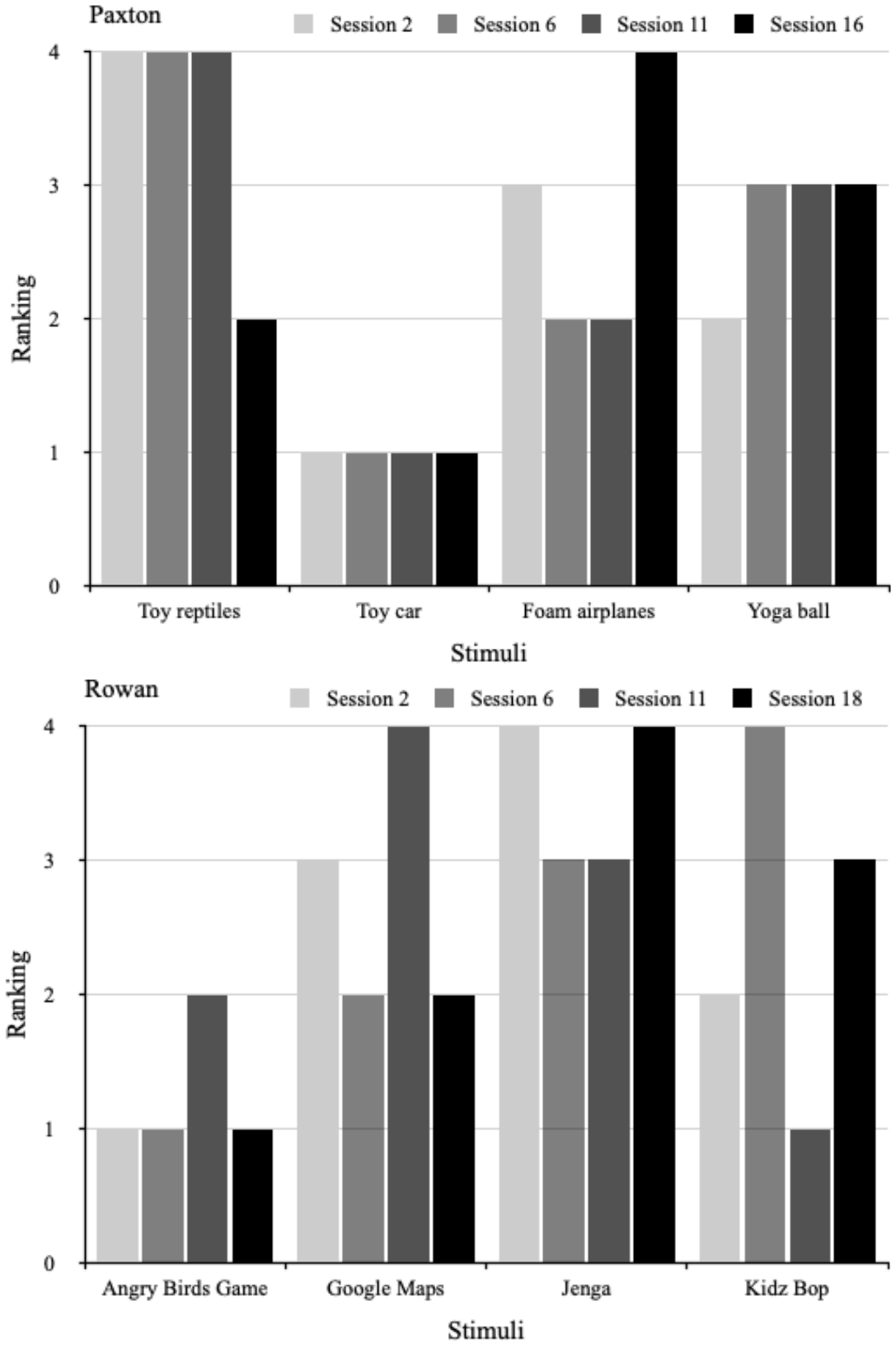
Results

Reinforcer Rankings

The study hypothesized that reinforcer rankings would show minor changes in preferences over the course of the study. The stimuli used during the preference assessments were ranked one through four, with one being the highest preferred, two and three being moderately preferred, and four being least preferred. Figure 1 shows reinforcer rankings of

each participant over the course of the study. When looking at the bar graphs, the highest preferred stimuli are displayed with the shortest bars, the moderately preferred stimuli with the middle-sized bars, and the least preferred stimuli with the tallest bars. With Paxton's rankings, his least preferred and highest preferred stimuli remained consistent until session 16 where his least preferred stimuli became a moderately preferred stimuli. Rowan's rankings were more variable than Paxton's. His highest and least preferred stimuli changed during each preference assessment, in addition to his moderately preferred stimuli changing during most of the assessments. This variability is consistent with trends identified in the primary researcher's undergraduate research (Histand, 2021).

Figure 1.
Reinforcer Rankings



Interfering Behavior

Paxton's frequency of interfering behaviors was variable during the preference assessment sessions. Rowan's frequency of interfering behaviors was decreased during the preference assessment sessions, until he broke a light switch cover during the last assessment, which led to an increased frequency of misconduct conversations surrounding breaking the light switch cover. Both Paxton and Rowan had an increased frequency of interfering behaviors during the highest preferred evaluation sessions and a decreased frequency during the least preferred evaluation sessions.

The differences in frequency of interfering behaviors between the highest and least preferred evaluations for both participants is likely due to the function of their behaviors. Although function of behaviors were not assessed during this study, it is possible that the function is at least partially maintained by denied access. If this is the case, the participants were more motivated to gain access to the highest preferred reinforcer. Therefore, they emitted more interfering behaviors when the reinforcer was removed at the end of the reinforcement period. There was not a correlation between the frequency of interfering behaviors and preference stability with either participant.

Figure 2 shows Paxton's frequency of interfering behaviors and Figure 3 shows Rowan's frequency of interfering behavior across the different sessions types. The solid phase change line on the evaluation graphs designates when the baseline conditions ended and intervention reinforcer evaluations began. Shaded and unshaded data points were used to differentiate data points that overlap, to make the graphs more readable. Some of the unshaded data points may appear shaded if a shaded data point overlaps an unshaded data point.

Figure 2.
Paxton's Frequency of Interfering Behavior

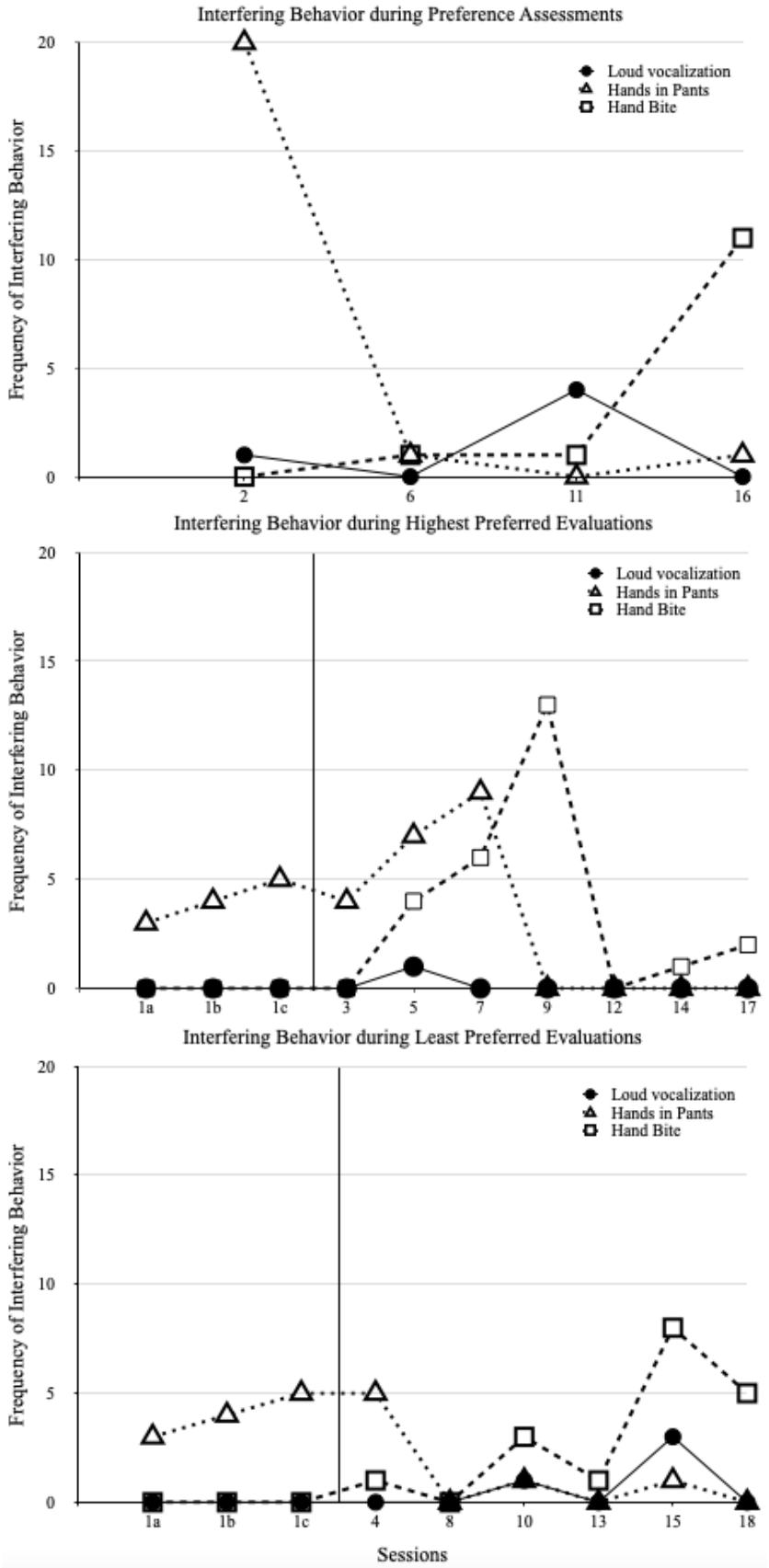
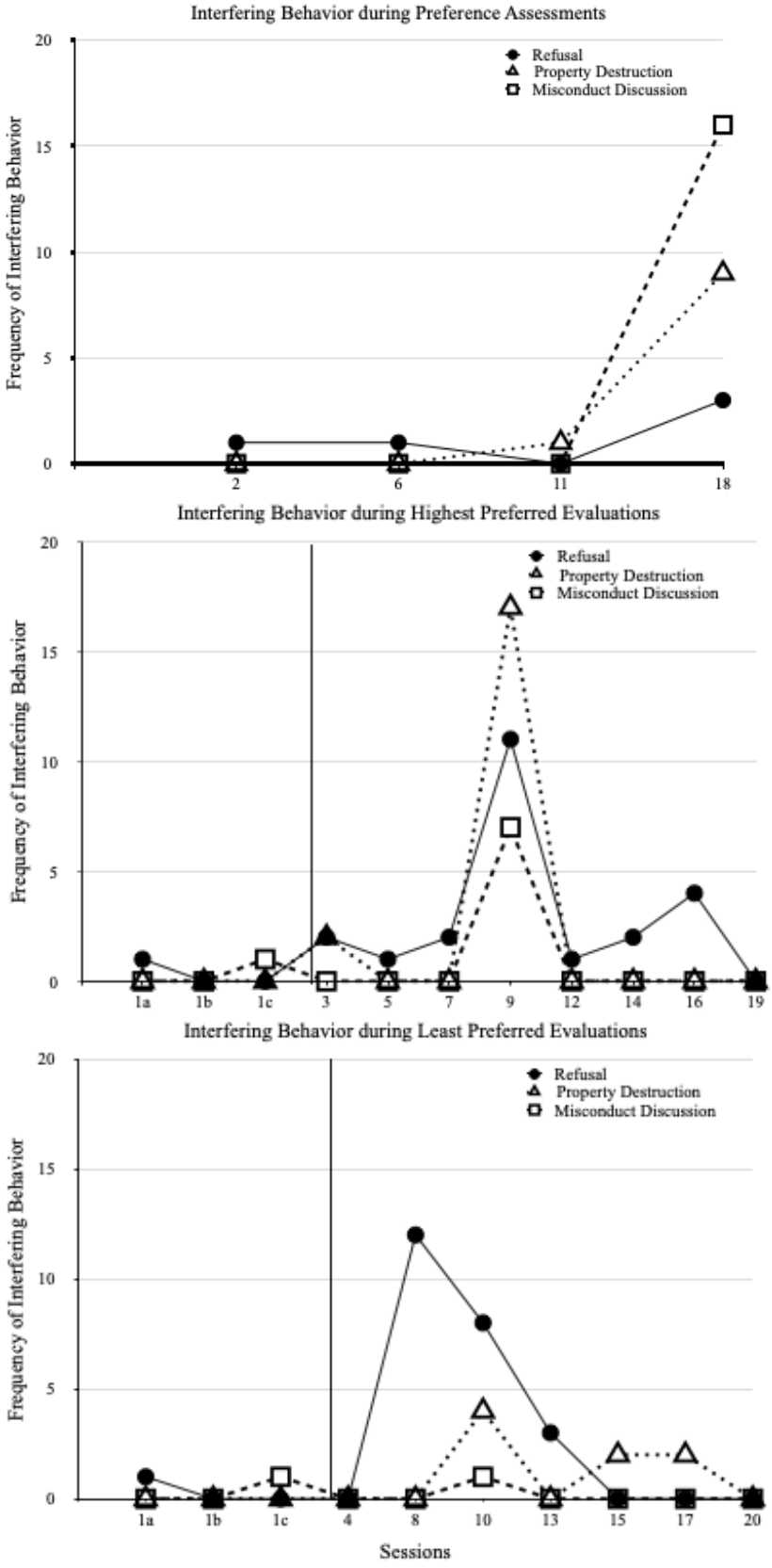


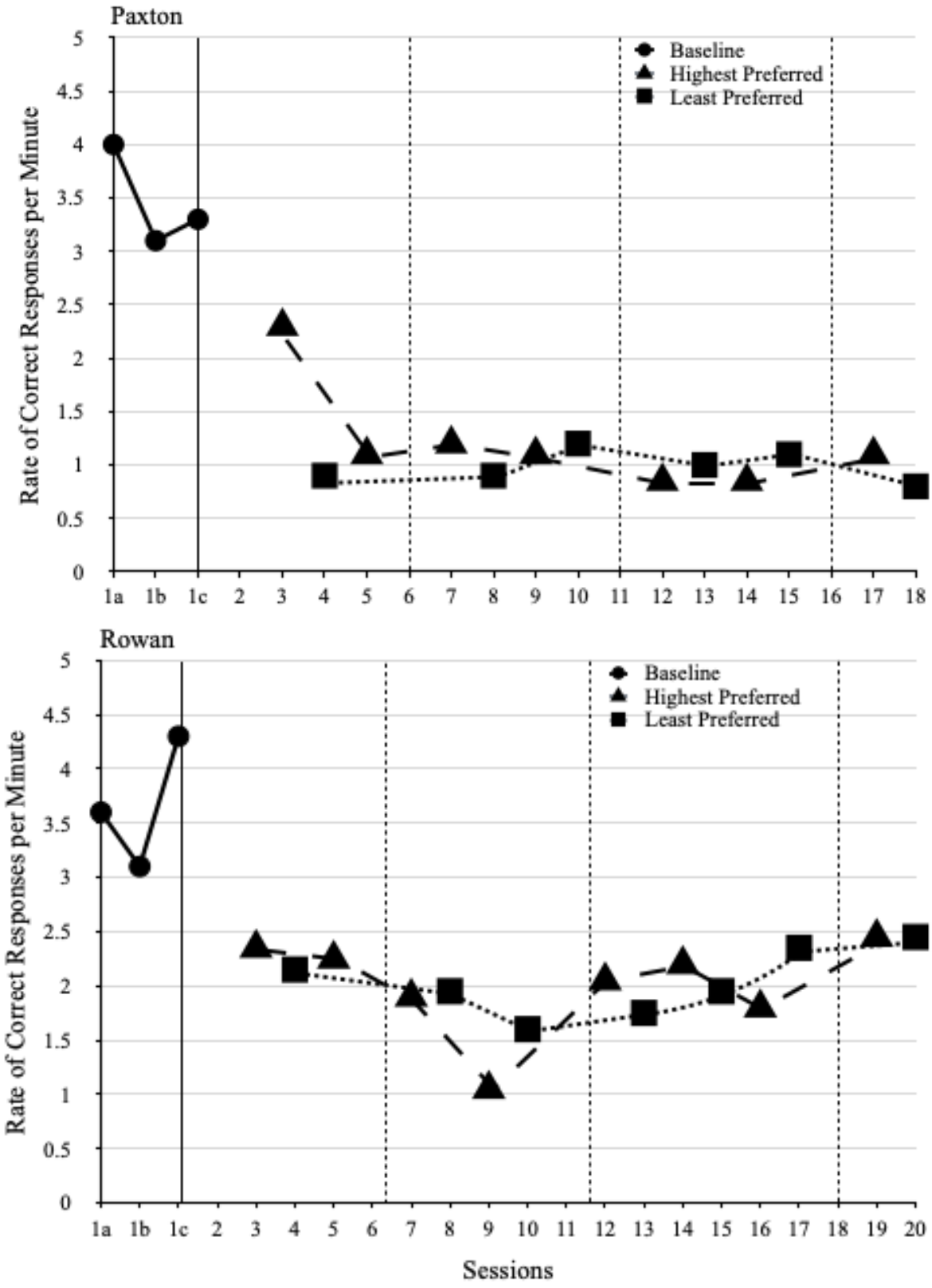
Figure 3.
Rowan's Frequency of Interfering Behavior



Preference Stability

The study hypothesized that stability in preferences would persist for six to eight weeks. At that point, performance during the reinforcer evaluations were hypothesized to decrease during highest preferred sessions and/or increase during least preferred sessions. Both participants' preferences became unstable during session 10, which was four weeks after the most recent preference assessment. Additionally, Paxton's preferences become unstable during session 15, which was also four weeks after the most recent preference assessment. This is consistent with previous research, which found that preferences remained stable over the course of four weeks (Carr, et al., 2000). However, Rowan's preferences became unstable during session 17, which was six weeks after the most recent preference assessment. This timeframe is more consistent with the present study's hypothesis. Figure 4 shows the participants' rate of correct responses per minute, which represents their performance during the reinforcer evaluation sessions. The solid phase change line designates when the baseline condition ended and intervention reinforcer evaluations began. The dashed phase change lines designate when preference assessments were repeated.

Figure 4.
Rate of Correct Responses per Minute



Discussion

This study sought to determine how long the stability of preferences persist using free-operant preference assessments and reinforcer evaluations, with children who have developmental disabilities. The study used an alternating treatment multielement design across participants, where each participant experienced a baseline probe condition and two alternating reinforcer evaluation conditions: 1) highest preferred stimulus for contingent reinforcement for correct responses; and 2) least preferred stimulus for contingent reinforcement for correct responses.

The results showed that preferences remained stable for an average of four and a half weeks across both participants. Paxton's preferences persisted for four weeks and Rowan's preferences persisted for five weeks on average. This demonstrates that each individual's preferences change over the course of different timeframes. This study replicate those found in Carr and colleagues' (2000) study, which identified that preferences remained stable for four weeks. Additionally, the present study extended this research and discovered that preferences may remain stable for longer periods of time for some individuals.

This study demonstrated the importance of having updated reinforcer rankings when working with individuals because, as preferences changed during the study, the participants' performance while using the highest preferred reinforcer decreased. Using reinforcers that no longer function as highest preferred reinforcers is a potential barrier to success when working with a client who is learning new skills.

Future Directions

Further research into preference stability using other types of preference assessments would contribute to the field whether the results observed in the present study are consistent across other types of preference assessments. It is possible that preferences could appear more or less stable when evaluating preference rankings using other methods and modes of assessment. Examples of other types of assessment that could be explored are paired-stimulus and multiple-stimulus assessment methods and web-based modes of assessment.

Web-based preference assessments are taken on an electronic device, such as a computer or tablet (Curiel, et al., 2020). This form of assessment administration would not require an assessment administrator to be physically present to perform the assessment. This method might be a more desirable method for individuals who spend much of their time using tablets or computers for communication or leisure activities (Curiel, et al., 2020).

In 2019, Curiel and Poling conducted an experiment using a web-based, multiple-stimulus without replacement preference assessment method. Their study found that web-based multiple-stimulus preference assessments were effective in identifying preference rankings based on reinforcer assessments. Further research into the stability of preferences using web-based, paired-stimulus, and multiple-stimulus preference assessments would contribute greatly to the field of Applied Behavior Analysis.

Limitations

Biased Data Collection

Another limitation of the present study is that data collectors were not blind to which intervention condition was assigned to each reinforcer evaluation. This could have led to biased data collection because the assumption is that highest preferred items should evoke

higher rates of correct responding in comparison to least preferred items. While this is a potential risk, it is important to note the high percentage of agreement across both intervention conditions, which suggests a low threat of bias in the present study.

Pre-existing Therapeutic Relationships

A final limitation of this study is that the primary researcher had previously established relationships with both participants in the study. These relationships were established in a clinic setting. Because of this, it is possible that the primary researcher already held stimulus control over the participants, which could have affected the participants' performance during reinforcer evaluations. Stimulus control is when a behavior is triggered by the presence or absence of a particular stimulus (Cooper, 2020). In this case, higher rates of correct responding would be under stimulus control of the therapist being present. This risk was mitigated by only having the primary researcher conduct research sessions to maintain consistent stimulus control across all sessions.

Clinical Implications

The current study demonstrated the importance of having updated reinforcer rankings when working with clients because decreasing reinforcer effectiveness led to decreased rates of correct responding among both participants. If new preference assessments were not conducted with the participants, it is likely that their rates of correct responding would either continue to decrease or plateau at lower rates of responding.

During this study, the researchers identified that preferences remained stable for four and a half weeks on average across both participants. This provides a rough timeframe when preference assessments should be repeated with clients; however, Paxton's preferences

remained stable for four weeks and Rowan's preferences remained stable for five weeks on average. This suggests that each individual's preferences persist for different amounts of time.

A recommendation based on this study's results is that clinicians should initially repeat preference assessments with new clients every five weeks. After the clinician gathers performance data for a minimum of 10 weeks, they should look for decreasing trends in responding leading up to the repeated preference assessments. If there is a decreasing trend present, the clinician should start conducting preference assessments every four weeks. If there is not a decreasing trend, the clinician should start conducting preference assessments every six weeks. These suggestions maximize service time by not conducting preference assessments more often than necessary, while still maintaining updated reinforcer rankings.

Appendix A

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Baseline Reinforcer Evaluation Data Sheet

Probe 1:

Minute	+ Responses (Talley)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Total + responses: _____	

Probe 2:

Minute	+ Responses (Talley)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Total + responses: _____	

Probe 3:

Minute	+ Responses (Talley)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Total + responses: _____	

Appendix B

Participant: ________ **Date:** ________ **Session #:** ________ **Data collector initials:** ________

Free-Operant Assessment 1 Data Collection Sheet

Item:	**Approached:** (within one meter)	**Engaged:** (touches)	**Duration of Engagement:**
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.

Total duration of assessment 1 (in minutes): ________

Participant: ________ **Date:** ________ **Session #:** ________ **Data collector initials:** ________

Free-Operant Assessment 2 Data Collection Sheet

Item:	**Approached:** (within one meter)	**Engaged:** (touches)	**Duration of Engagement:**
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.
			____ min., ___ s.

Total duration of assessment 2 (in minutes): ________

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Free-Operant Assessment Scoring Sheet
(score both assessments together)

Items:	Total Frequency Approached:	Total Duration of Engagement:
		_____ min., _____ s.
		_____ min., _____ s.
		_____ min., _____ s.
		_____ min., _____ s.
		_____ min., _____ s.

Total duration of both assessments (in minutes): _____

Highest preferred item(s) (approached frequently, engaged with for longest durations):

Moderately preferred item(s) (approached, engaged with for shortest durations):

Lowest preferred item(s) (not approached or engaged with):

Appendix C

Intervention Reinforcer Evaluation Data Sheet

Stimuli (circle): Highest or Lowest

Minute	+ Responses (Talley)
Part One	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Alternative Activity	
Part Two	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Total + responses: _____	

Appendix D

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Baseline Session Problem Behavior Frequency Data Collection Sheet

Problem behaviors: **Session Sections:**	**Hand/Arm Bites** Successful or attempt of teeth making forceful contact with skin on hand/arm	**Hands in Pants** Enters 1+ fingers through any opening of his pants. 1 hand = 1 occurrence; 2 hands = 2 occurrences. New instance is when fingernails are no longer visible	**Loud Vocalization** Any vocalization that is above conversational level. New instance is separated by 3 seconds of silence or appropriate volume	**Aggression toward Clinician** **Hitting**: Making forceful contact with any part of another person's body **Scratching**: Making contact with another person by moving fingernails across skin/ clothing	**Total Frequency**
Baseline Probe 1					
Alternative Activity					
Baseline Probe 2					
Alternative Activity					
Baseline Probe 3					
Total Frequency					

Total frequency of problem behaviors: _____

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Assessment Session Problem Behavior Frequency Data Collection Sheet

Problem behaviors: **Session Sections:**	**Hand/Arm Bites** Successful or attempt of teeth making forceful contact with skin on hand/arm	**Hands in Pants** Enters 1+ fingers through any opening of his pants. 1 hand = 1 occurrence; 2 hands = 2 occurrences. New instance is when fingernails are no longer visible	**Loud Vocalization** Any vocalization that is above conversational level. New instance is separated by 3 seconds of silence or appropriate volume	**Aggression toward Clinician** **Hitting**: Making forceful contact with any part of another person's body **Scratching**: Making contact with another person by moving fingernails across skin/ clothing	**Total Frequency**
Assessment Prt I					
Alternative Activity					
Assessment Prt II					
Total Frequency					

Total frequency of problem behaviors: _____

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Reinforcer Evaluation Session Problem Behavior Frequency Data Collection Sheet

Stimuli (circle): Highest or Lowest

Problem behaviors: **Session Sections:**	**Hand/Arm Bites** Successful or attempt of teeth making forceful contact with skin on hand/arm	**Hands in Pants** Enters 1+ fingers through any opening of his pants. 1 hand = 1 occurrence; 2 hands = 2 occurrences. New instance is when fingernails are no longer visible	**Loud Vocalization** Any vocalization that is above conversational level. New instance is separated by 3 seconds of silence or appropriate volume	**Aggression toward Clinician** **Hitting**: Making forceful contact with any part of another person's body **Scratching**: Making contact with another person by moving fingernails across skin/ clothing	**Total Frequency**
Session Part I					
Alternative Activity					
Session Part II					
Total Frequency					

Total frequency of problem behaviors: _____

Appendix E

Participant: _____ **Date:** _____ **Relation to Participant:** _____

Parent Reinforcer Survey

Toys:			
Toy	**Never**	**Sometimes**	**Regularly**
Toy car			
Train set			
Doll			
Mr. Potato Head			
Make-up/Dress-up			
Musical toy			
Block			
Lincoln Logs			
Smaller toy ball			
Yoga ball			
Book			
Puzzel			
Kitchen set/Toy food			
Toy animal			
Board game			
Legos			
Balloon			
Bubbles			
Trampoline			
Play-Doh			
Stuffed animal			
Fidgets/Pop-it			
Slinky			
Other toys **regularly** used: _____ _____			

Participant: _____ **Date:** _____ **Relation to Participant:** _____

Activities:			
Activity	**Never**	**Sometimes**	**Regularly**
Listening to music			
Dancing			
Singing			
Vacuuming			
Coloring			
White board drawing			
Watching YouTube			
Playing tablet games			
Other activities **regularly** engaged in: _____ _____			

Appendix F

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Assessment Session Procedural Checklist Data Collection Sheet

Checklist Item	Observed	Not Observed	Notes
Furniture placed around treatment room perimeter			
Same stimuli included as previous assessments			
Stimuli spaced out around room perimeter			
1st assessment lasted within 15 sec. of 10 min.			
During 1st assessment therapist **did not** engaged with stimuli			
During 1st assessment therapist **did not** engaged with participant			
Mastered skills were targeted during alt. activity			
During 2nd assessment therapist **did not** engaged with stimuli			
During 2nd assessment therapist **did not** engaged with participant			
2nd assessment lasted within 15 sec. of 10 min.			
Total score: _____/10 observed			

Participant: _____ **Date:** _____ **Session #:** _____ **Data collector initials:** _____

Reinforcer Evaluation Procedural Checklist Data Collection Sheet

Checklist Item	Observed	Not Observed	Notes
Furniture placed around treatment room perimeter			
Stimuli pre-selected			
Only assigned stimuli present			
Target behavior materials prepared			
Unmastered skills were targeted			
Reinforcer **not** provided for **incorrect** response			*Measure 1st 3 opportunities of session 1
Appropriate error correction provided			*Measure 1st 3 opportunities of session 1
Reinforcer **not** provided for **incorrect** response			*Measure 1st 3 opportunities of session 2
Appropriate error correction provided			*Measure 1st 3 opportunities of session 2
Total score: _____/_____ (17 - total number of no opportunities)			

Slash through observed/not observed boxes if there was no opportunity to observe checklist item (i.e., if the participant did not emit 3 incorrect responses to perform error correction).

References

Butler, C. & Graff, R.B. (2021). Stability of preference and reinforcing efficacy of edible, leisure, and social attention stimuli. *Journal of Applied Behavior Analysis. 54*(2), 684-699. doi: 10.1002/jaba.807

Carr, J., Nicolson, A. & Higbee, T. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis. 33*(3), 353-357. doi: 10.1901/jaba.2000.33-353

Ciccone, F., Graff, R., Ahearn, W. (2015). Increasing the efficiency of paired-stimulus preference assessments by identifying categories of preference. *Journal of Applied Behavior Analysis. 48*(1), 221-226. doi: 10.1002/jaba.190

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied Behavior Analysis* (3rd ed.). Pearson Education, Inc.

Curiel, H., Curiel, E., Adame, A. & Li, A. (2020). Multiple-stimulus-without-replacement preference assessment tool. *Behavioral Interventions: Early View.* doi: 10.1002/bin.1732

Curiel, H. & Poling, A. (2019). Web-based stimulus preference assessment and reinforcer assessment for videos. *Journal of Applied Behavior Analysis, 52*(3). doi: 10.1002/jaba. 593

DeLeon, I., Fisher, W., Rodruiguez-Catter, V., Maglieri, M., Herman, K. & Marhefka, J. M. (2001). Examination of relative reinforcement effects of stimuli identified through pretreatment and daily brief preference assessments. *Journal of Applied Behavior Analysis. 34*(4), 463-473. doi: 10.1901/jaba.2001.34-463

Fisher, W, Piazza, C., Bowman, L., & Amari, A. (1996). Integrating caregiver report with a

systematic choice assessment. *American Journal on Mental Retardation, 101,* 15–25.

Francisco, M.T., Borrero, J.C. & Sy, J.R. (2008). Evaluation of absolute and relative reinforcer value using progressive-ratio schedules. *Journal of Applied Behavior Analysis. 41*(2), 189-202. doi: 10.1901/jaba.2008.41-189

Hanley, G., Iwata, B. & Roscoe, E. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis, 39*(2) 189-202. doi: 10.1901/jaba. 2006.163-04

Histand, L. (2021). *A comparison of preference assessment modes and methods: In-person versus electronic-based modes and paired-stimulus versus MSWO preference assessment methods* [unpublished undergraduate thesis]. Eastern Mennonite University.

Institute for Applied Behavior. (1993). Section 3 Data Sheets. *Reinforcement Inventory for Children and Adults* (32–49).

Johnston, J. M., Pennypacker, H. S., & Green, G. (2020). *Strategies and Tactics of Behavioral Research and Practice*. doi: 10.4324/9781315537085

Johnson, K., Vladescu, J., Kodak, T. & Sidener, T. (2017). An assessment of differential reinforcement procedures for learners with autism spectrum disorder. *Journal of Applied Behavior Analysis, 50*(2), 290-303. doi: 10.1002/jaba.372

Kang, S., Lang, R.B., O'Reilly, M.F., Davis, T., Machalicek, W., Rispoli, M., & Chan, J. (2010). Problem behavior during preference assessments: An empirical analysis and practical recommendations. *Journal of Applied Behavior Analysis, 43*(1), 137-141. doi: jaba. 2010.43-137

Leaf, J., Leaf, R., Alcalay, A., Leaf, J., Ravid, D., Dale, S., Kassardjian, A., Tsuji, K., Taubman, M., McEachin, J., & Oppenheim-Leaf, M. (2015). Utility of formal preference assessments for individuals diagnosed with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities, 50*(2), 199-212. doi: 205.132.43.136

Livingston, C. & Graff, R. (2018). Further evaluation of preference categories to identify novel reinforcers: A systematic replication. *Behavioral Interventions, 33*(2), 173-184. doi: 10.1002/bin.1519

Ortiz, K. & Carr, J. (2000). Multiple-stimulus preference assessments: A comparison of free-operant and restricted-operant formats. *Behavioral Interventions, 15*(4), 345-353. doi: 10.1002/1099-078X(200010/12)15:4<345::AID-BIN69>3.0.CO;2-K

Piazza, C., Fisher, W., Hagopian, L., Bowman, L. & Toole, L. (1996). Using a choice assessment to predict reinforcer effectiveness. *Journal of Applied Behavior Analysis, 29*(1), 1-9. doi: 10.1901/jaba.1996.29-1

Roane, H., Vollmer, T., Ringdahl, J. & Marcus, B. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis, 31*(4), 605-620. doi: 10.1901/jaba.1998.31-605

Roscoe, E.M. & Fisher, W.W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavior Analysis, 41*(2), 249-254. doi: 10.1901/jaba.2008.41-249

Verriden, A. & Roscoe, E. (2016). A comparison of preference-assessment methods. *Journal of Applied Behavioral Analysis, 49*(2), 265-285. doi: 10.1002/jaba.302