7-2022

Comparison of Traditional and Eye-Gaze Preference and Reinforcer Assessments

Zhihao Ou

Comparison of Traditional and Eye-Gaze Preference and Reinforcer Assessments

by

Zhihao Ou

A thesis submitted to the College of Psychology and Liberal Arts of
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Applied Behavior Analysis and Organizational Behavior Management

Melbourne, Florida
July, 2022

We the undersigned committee hereby approve the attached thesis,
"Comparison of Traditional and Eye-Gaze Preference and Reinforcer Assessments."
by
Zhihao Ou

Kimberly Sloman, Ph.D., BCBA-D
Associate Professor
School of Behavior Analysis
Major Advisor

David Wilder, Ph.D., BCBA-D
Professor
School of Behavior Analysis

Vida Tyc, Ph.D.
Professor
School of Psychology

Robert A Taylor, Ph.D.
Professor and Dean
College of Psychology and Liberal Arts

Abstract

Title: Comparison of Traditional and Eye-Gaze Preference and Reinforcer Assessments

Author: Zhihao Ou

Advisor: Kimberly Sloman, Ph.D., BCBA-D

A variety of preference assessments have been used to identify potential reinforcers for individuals with autism spectrum disorder (ASD) and related disabilities. However, during traditional selection-based assessments, participants may fail to select items or selection may be under faulty stimulus control (e.g., side bias). Previous research has evaluated the utility of eye-gaze preference assessments with individuals with profound intellectual disabilities. The purpose of this study is to compare preference assessments by eye gaze and traditional preference assessment in three participants with ASD. The results of reinforcer assessment showed that among 2 out of 3 participants, the eye gaze method was more accurate compared to the traditional preference assessment. The breakpoint in the progressive ratio schedule matched with the result as well.

Keywords: preference assessment, reinforcer assessment, eye gaze.

Table of Contents

List of Tables

List of Figures

Acknowledgement

Thank you, Dr. Sloman. for giving me the thesis idea and patiently teaching me how to revise the paper.

Thank you, Dr. Wilder, for giving me suggestions for my thesis and teaching me valuable knowledge in the class.

Thank you, Dr. Tyc, for being my committee member and helping me to improve the deficiencies of my thesis.

Dedication

This thesis is dedicated to my parents who have supported me throughout my education. Thanks for making me see this adventure through to the end.

Chapter 1 Introduction: Comparison of Traditional and Eye-Gaze Preference and Reinforcer Assessments

For clients with an Autism Spectrum Diagnosis (ASD), especially those with severe intellectual disabilities, the goal of intervention is to teach functional communication skills, adaptive daily living skills, and social skills. Applied behavior analysis (ABA) achieves the purpose of promoting functional skills by identifying and arranging environmental variables that can affect behavior patterns (Cooper et al., 2019).

Applied behavior analytic teaching strategies often involve breaking down the target tasks into smaller, relatively independent steps. ABA therapists use reinforcement methods to gradually train each small step until the individual masters all the steps and finally completes the task independently in the same or different settings. The critical point is reinforcement. Reinforcement is defined as the presentation or removal of a stimulus, following a response, which results in the increase of that response. When the rate of response increases due to presenting a stimulus, it is positive reinforcement. For example, if parents praise their child for getting a good grade, and it increases the rate of getting good grades. On the other hand, when the rate of response increases due to removal of a stimulus, it is negative reinforcement. For example, a person scratches his back to reduce an itch. A reinforcer is defined as the stimuli that increase the rate of the target response. Using the above examples, praise would be a positive reinforcer, and scratching

that terminates the itch would be a negative reinforcer. In general, positive reinforcement is most often programmed in treatment.

Reinforcement is affected by different factors. One factor is the schedule of reinforcer delivery. It may be necessary to use different reinforcement delivery schedules under different situations. For example, continuous reinforcement schedules involve presenting a reinforcer on a fixed-ratio (FR) 1 schedule, or after every response. This schedule is typically used to establish a response. Other schedules involve intermittent delivery of a reinforcer based upon the number of responses (ratio schedule), the first response after a certain time period elapses (interval schedule), or on the passage of time in general (time-based or noncontingent schedule). A second factor is immediacy. Generally, reinforcement is most effective when the reinforcer is delivered immediately after the response, as it pairs the target response to the delivery of the reinforcer. The third factor is quantity or magnitude of the reinforcer. According to Lasserre et al. (2008), magnitude represents the quantity, intensity, and duration of reinforcement. Generally speaking, the greater the magnitude of reinforcement is, the more effective the reinforcer will be, meaning it will lead to more significant behavior change (Lasserre et al., 2008). Magnitude refers to the amount of reinforcement given and can be indicated by size of the reinforcer (e.g., a large candy bar vs. a small candy bar) or duration of access (e.g., 30 s of iPad access vs. 2 min of iPad access). A fourth factor is quality of the reinforcer. Quality is often assessed by

observing an individual's selections between one or more items, typically during preference assessments. Items which are selected most frequently are considered higher quality than items selected less frequently or never.

When programming for skill acquisition, new and more difficult tasks may receive greater quantities and higher qualities of reinforcers than easier or prompted tasks. For example, if a participant completed a mastered task, they are likely to get a small size reinforcer. However, if the participant completed a new task, they will be given a larger reinforcer. Research has shown that these arrangements may result in more effective and efficient acquisition of new skills than non-differential reinforcement procedures (e.g., Johnson et al., 2017; Paden & Kodak, 2015).

The last factor affecting reinforcer efficacy is motivating operations. Motivating operations are events that alter the effectiveness of reinforcers. For example, if a person has not consumed water for a long time, the value of the water may be increased as a reinforcer. On the other hand, if a person just had dinner, the value of food may decrease as a reinforcer. Practitioners must be aware of how reinforcers are arranged in the environment to optimize their effects for skill acquisition and avoid reinforcer competition or satiation.

Reinforcers should be identified via direct observation when possible. Some research has shown that reinforcers identified via indirect measures (e.g., asking the client or caregivers) may not be accurate. For example, Green et al. (1988)

compared the systematic assessment of preferences to staff opinion. The result indicated that staff opinion could not accurately predict the clients' preferences.

Preference Assessment

The first step in identifying reinforcers is to conduct preference assessments. The history of formal preference assessments can be traced back to 1985. The single stimulus (SS) preference assessment procedure was first created by Pace et al. In their study, participants were 3 years to 18 years old and diagnosed with a severe intellectual disability. Sixteen stimuli were used. Each item was presented one at a time in front of the participants. The participants had five seconds to engage in an approach response. If the participant approached the item, they were allowed to engage with it for an additional five seconds. If no approach was observed, the item was represented an additional time. Following no response, the item was removed, and the next item was presented. Each of the stimuli was presented a total of ten times across the assessment. Items that were approached on at least 80% of the trials were considered highly preferred and items that were approached less than 50% of the assessment were considered non-preferred.

In the second experiment, Pace et al. conducted a reinforcer assessment. There were three conditions: baseline, preferred item, and nonpreferred item. The dependent variable was the percentage of correct responses of individually determined arbitrary responses. In the baseline condition, researchers gave vocal instructions with gestures, but no consequence was provided for engaging in the

response. In the preferred item condition, if participants complied, preferred items were provided contingent on correct responses. In the nonpreferred condition, nonpreferred items were provided contingent on correct responses. Results showed that using preferred stimuli increased the percentage of correct target responses relative to the baseline and nonpreferred stimuli conditions.

The advantages of the SS preference assessment procedure are it empirically identifies preferred stimuli, and it is easy to conduct. However, there are some disadvantages. First, it may over-identify stimuli as reinforcers because the participant may approach and engage with all stimuli. Second, it does not show preference hierarchy or relative preferences for stimuli. Third, it may be time-consuming to conduct compared to other methods.

In 1992, Fisher et al. created the paired stimulus (PS) preference assessment (also known as paired choice or forced-choice preference assessment). This study involved 4 children between 2 to 10 years old and diagnosed with severe intellectual disabilities. The researchers first replicated the experimental procedures of Pace using the SS arrangement and then conducted the paired stimulus procedure. Each item was paired with another in a quasi-random format during the paired stimulus procedure until all items were paired at least one time. Both items were placed equidistant in front of the participants, and the participant was instructed to choose one. If neither of the items were chosen, the participant was prompted to sample each item, and then they were represented again. If no

approach was made, both items were removed, and the subsequent trial began. The dependent variable was the percentage of opportunities that the items were selected. Results showed that the PS preference assessment resulted in more significant differentiation among stimuli.

In the second experiment, the researchers conducted reinforcer assessments to compare the efficacy of items that were selected as highly preferred in both preference assessments to the items that were only highly preferred in the SS preference assessment. Reinforcer efficacy was evaluated using a concurrent operant arrangement. Both stimuli were available and placed in front of seats or squares in the room. The dependent variable was the percentage of trials that the participant engaged in the target response (i.e., “in seat" or "in square"). After comparison, the experimenters found that stimuli which were defined as a high preferred item in both preference assessments were more effective at increasing the target response than the stimuli identified by the SS preference assessment.

In 1999, Roscoe et al. replicated Fisher's experiment. Eight adult individuals with intellectual disabilities participated in this experiment. In the first study, researchers used the SS preference assessment method to assess preference for ten edibles. Researchers placed each item in front of the participants and observed if they would try to grab or touch the item. If participants touched the item, they were allowed to eat it. If they did not respond within 5s, researchers would prompt them to eat the item. If the participants still did not make a choice, the trial was

terminated, and a new trial started. Next, researchers conducted a PS preference assessment. Two edibles were randomly paired and presented to the participants. Once participants selected an item from the pair, they were allowed to eat. If there was no response within 5 s, researchers would let participants eat both items and restart the trial. If participants did not select either item again, the trial was terminated. The results of this study showed that under SS preference assessment conditions, six of the eight participants touched all stimuli. One of the remaining two participants touched 9 out of 10 and the other touched 7 out of 10. However, clearer preferences were observed under the PS preference assessment condition.

In experiment 2, researchers conducted two kinds of reinforcer assessment. Two stimuli were selected for each participant. One stimulus (High P) was the edible that was shown to have greatest correspondence between preference assessments. The other stimulus (Low P) was the edible that was shown to have the least correspondence between assessments. In the first phase of the reinforcer assessment, the researchers presented two color pads concurrently that the participant could press. In baseline, no consequence was provided for pressing either color pad. In the reinforcer assessment condition, each pad corresponded to a different food. When participants pressed the High P pad, they received the High P edible. When they pressed the Low P pads, they received the Low P edible. The results showed responding was low for 4 participants in the concurrent baseline. During the concurrent reinforcer assessment, responding to the High P edible pad

was higher than the low P edible pad. In the second phase of the reinforcer assessment, only one pad was presented at a time. In baseline, the previous Low P edible pad was placed in front of the participants and no consequences were provided for presses. Next, if participants pushed the pad, a Low P edible was provided. During baseline, response rates decreased. During the reinforcement condition, when only the Low P pad was available, participants engaged in high rates of responding for the Low P edible. The results indicate that items selected less during preference assessment still may function as reinforcers in single operant arrangements.

Compared to SS procedure, PS preference assessments yield a preference hierarchy among stimuli, may also yield fewer false positives, and can be completed more efficiently than the single stimulus procedure (Fisher et al., 1992). One limitation of this method is that it is time-consuming.

In 1996, Iwata and DeLeon compared three different preference assessments. Multiple stimuli without replacement preference assessment (MSWO), paired stimulus preference assessment, and multiple stimuli with replacement assessment (MSW). The experiment involved seven participants who were diagnosed with a severe intellectual disability. Before the experiment, participants were allowed to sample all of the items that were used. The PS assessment was conducted using the procedures described in Fisher et al (1992). In the MSWO assessment, all items were arranged in front of the participant

simultaneously. The participants were instructed to pick one of them. Once an item was selected, the participant was allowed to engage with the item, and then it was removed from the array. This process continued until all the items were selected or the participant did not select any items within 30 seconds. The MSW was identical to the MSWO; however, the only difference was that the selected items were placed back in the array. Results showed that the MSWO and paired stimulus procedure resulted in a similar preference hierarchy.

In the second experiment, the researchers conducted a reinforcer assessment. They used stimuli that were never chosen in the MSW but were chosen in the MSWO and PS assessments. The items were provided contingently on target listener response tasks. Results indicated that those items increased the response rate among three out of four participants. It means that items labelled as unidentified in MSW may still work as reinforcers. MSWO and paired choice procedures can identify these stimuli more effectively.

The strength of MSWO and MSW is that they can be done quickly, and MSWO produces a preference hierarchy. However, these methods may result in difficulties in determining what is preferred and nonpreferred. That is, all the stimuli available and selected may function as reinforcers.

Roane et al. (1998) conducted the free operant (FO) preference assessment. In the experiment, 20 participants were provided access to all stimuli during 5 min sessions and researchers measured engagement (defined duration that the

participant was touching the item) with the stimuli. Then, researchers conducted two reinforcer assessments. In the first one, there were two squares. One had a highly preferred item that was identified from the FO assessment, and another square had nothing. Participants were able to access highly preferred items contingent on staying in the square. Five out of six participants picked the square that had preferred items. The second reinforcer assessment had two workstations. One had a highly preferred item, and the other one had a nonpreferred item. Participants could complete a task in either station in order to get the corresponding item. Three of four participants picked the station with their high preferred item. In experiment 2, researchers compared the FO method to the PS method. Results showed that the FO method was completed more efficiently compared to the PS procedure. However, it may not yield a preference hierarchy among all stimuli evaluated. It may only yield to one or a few preferred stimuli.

To address limitations of the FO assessment, Hanley et al. (2003) conducted a response restriction (RR) analysis. Three individuals with intellectual disabilities participated in this experiment. In the first study, each session was 5 min and participants were able to freely choose between engaging in seven activities. After a clear preference was observed, restrictions were added in subsequent assessments. That is, the highest preferred activity was removed from the array and the participant selected from the remaining activities. This process continued until all seven activities were removed or participants showed little or no interaction with

the remaining activities. In the second experiment, the researchers compared the data from the FO analysis to the RR analysis. The results showed in FO sessions, participants focused only on one activity, which masked preference for other activities. The response restriction analysis resulted in a hierarchy of preference. However, the disadvantage of this method is that is it time consuming.

Further Comparisons of Preference Assessment Procedures

Kang et al. (2013) compared 14 studies on preference assessment, including SS, PS, MSWO, and FO, from 1985 to 2013. The results show that, compared to other methods, MSWO and PS assessments produced more accurate prediction for which stimuli would function as reinforcers. However, just looking at the accuracy rate is not enough. Efficiency is also important. The average time of PS is 31.5 minutes, while the average time of MSWO test is 15.9 minutes. Not only that, MSWO was also the most consistent procedure. On the other hand, although the effects of SS and FO are relatively inaccurate, the main reason is the relatively small number of studies. Although MSWO looks comprehensive, it also has shortcomings. For example, it may evoke problem behavior that maintained by access to tangible items. Another disadvantage is that when using picture cards in place of actual items to perform a MSWO, the results are less accurate.

Verriden and Roscoe (2016) conducted a study to compare the efficacy and stability of four preference assessments (i.e., PS, MSWO, FO, and RR). Six individuals with autism participated. The researchers first found seven preferred

items/ activities for participants and then conducted the four preference assessments six times each. Following each assessment, researchers converted item engagement to rank and calculated Spearman rank-order correlation coefficients and Kendall rank coefficient of concordance to evaluate stability of preference over time. They also measured problem behavior exhibited across the different preference assessments.

In PS, the researchers presented stimuli in pairs, and the therapist verbally instructed the participants to choose one of the presented items. If the participants selected an item, the therapist would provide 20 seconds of access to that item, and the other item would be removed. This process continued until all items were matched with every other item at least once. In the MSWO, seven items were presented at the same time. After participants chose one item, they were allowed to access the item for 20 s and the item was removed. This process was completed with the remaining 6 items. In the FO preference assessment, all items were presented simultaneously, and the participants could freely engage with any item during a five-minute session. In the RR preference assessment, there were several 2-min trials and initially, all items were made available. Participants chose between all seven items. Items that were highly engaged in 2 or 3 continuous sessions were removed or restricted in subsequent trials. The results showed that PS and MSWO methods were more stable (i.e., higher Spearman and Kendall coefficients) than FO and RR methods. It also showed that in FO, the frequency of problematic behaviors

was the lowest among all participants. The lower rates of problem behavior were mostly likely due to the fact that they had free access to all tangibles during FO in comparison to other methods in which some of the tangibles were removed.

In Experiment 2, the researchers tested the effect of preference stability on reinforcer efficacy. The dependent variable was the rate that the participants completed mastered tasks that required minimal effort to complete. Preference assessments from Experiment 1 that resulted in less stability over time were conducted before the reinforcer assessment to determine each participant's high preference item (HP initial). Immediately before each reinforcer assessment, researchers again conducted the preference assessment to determine HP immediate items. Reinforcer efficacy was assessed in a concurrent-operants arrangement for two participants and a single-operant arrangement for one participant. During baseline, tasks were presented, and no consequence was provided for engaging in the tasks. In the Concurrent-operants reinforcement session, three tasks were presented and each corresponded to the HP initial item, HP immediate item, and no item. In the Single-operant reinforcer assessment, only one task and one reinforcer were presented at a time. The results showed that overall, the participant engaged in tasks to access the HP immediate over the HP initial during the concurrent arrangement. However, items with less preference stability still functioned as reinforcers when compared to baseline or no reinforcement conditions. Therefore, preference stability may not have a significant impact on reinforcement efficacy.

Use of Eye Gaze to Assess Preference

While a number of methods to identify preferred items and reinforcers exist, there are limitations to reinforcer identification for some clients with severe disabilities. Ivancic and Bailey (1996) found that individuals with profound disabilities might show no preference (i.e., do not approach) items during preference assessments. This makes identifying potential stimuli for use in reinforcement-based intervention difficult. The potential reasons could be that the response effort for selection is too high, or that the participant does not have the fluent motor skills necessary to respond. Ivancic and Bailey suggested that alternative methods to assess preference such as eye gaze can be used for clients who failed to demonstrate preferences in selection-based assessments.

In order to address the limitation mentioned above, Fleming et al. (2010) used eye gaze as an alternative way to assess preference. In the experiment, they had four participants with severe to profound disabilities. All participants had inconsistent motor movements and difficulties with communication. In the initial experiment, researchers conducted a paired choice preference assessment using eye gaze toward items as the dependent variable. They first put one of the two items in front of the participants and operated each item to show its reinforcing properties. This item would slowly move back to the original location from participants' eyesight. When both items were displayed, the vision of the participants would be blocked by a board. When the participant's sight was in a natural state, the shield

was removed. Researchers then gave the SD: choose one. If participants looked at a specific item for a specific length of time (individually determined for each participant), they could engage with the selected item for 5 seconds. If participants did not select either item, the process would be repeated. If still no items were selected, the session would be terminated. Next, researchers conducted a reinforcer assessment. In baseline, the researchers gave the SD “look at me”, but no consequence was provided for responding. In the high preference condition, researchers first would let participants engage with the high preference stimulus for 5 seconds. After access, researchers said “if you look at me, you can have (item)”. Then the researchers gave the SD “look at me”. If participants responded within 5 s, they were allowed to engage with the item. The low preference condition was identical except that the researchers provided the low preferred item contingent upon responding. Results showed that among the 3 participants, eye gaze was an effective way to identify reinforcers as responding during the reinforcer assessment was differentially higher in the high preferred condition compared to the low preferred condition. However, this experiment had three limitations. First, although three participants’ data were clear, one participant’s data had high variability. The authors posited that the response patterns were due to his previous learning history outside of the experimental setting and lack of discrimination skills. Second, the discriminative stimulus in the reinforcer assessment "look at me" was likely to have a reinforcement history for participants, and the looking response might not be

solely a function of accessing the stimuli during the experiment. Third, one participant's assessment was interrupted due to personal reasons.

In 2015, Malone et al. replicated Fleming's study. They had three participants with severe disabilities, assessed eye gaze in a PS selection and followed up with a reinforcer assessment. Results showed a similar conclusion to Fleming's study. However, this study also had similar limitations. First, the instruction that was used in the reinforcer assessment was "look at me" and this response was already in the participant's repertoires. Future studies should use new target skills. Second, minimal reinforcement effects were observed for one participant, which may indicate that reinforcers were not fully identified and additional stimuli should be tested. Lastly, as with PS assessments in general, the time to complete the assessment may be excessive. Using the eye gaze response with alternative preference assessment methods may be more efficient and effective.

The purpose of the current study was to compare more traditional selection-based preference assessments to assessments using eye gaze. To date, no study has compared eye-gaze to selection-based assessments. In addition, although previous evaluations have used eye gaze preference assessments with individuals with severe intellectual disabilities, these assessments may also benefit very young children who exhibit limited skill sets. More specifically, we compared selection-based PS

and MSWO assessments with PS and MSWO using eye-gaze. We conducted reinforcer assessments to validate the preference assessment outcomes.

Chapter 2 Methods

Participants

Three individuals participated in the study, and all were receiving ABA services at a university-based treatment clinic. Luffy was 4 years old; Cody was 3 years old, and Mike was 30 years old. All participants had received an independent diagnosis of Autism Spectrum Disorder Level 3. Mike had also been diagnosed with Childhood Disintegrative Disorder. None of the participants could communicate vocally. Luffy and Cody primarily communicated by exchanging pictures of preferred items. Mike communicated via modified sign language and was also learning to mand for items using a simplified card touch response.

Setting and Materials

The study was conducted in a therapy room separate from the participant's typical classroom. The therapy room had a table, a chair, a timer, a camera to record the study session, and multiple items for preference assessments.

Dependent variable and data collection

The dependent variable in Experiment 1 was choice between two or more items. In the selection-based assessment, choice was defined as reaching toward and touching an item. In the eye-gaze assessment, choice was defined as orienting and looking toward an item for at least 2 seconds. Data were collected across trials on participants' choice of items. Data from the PS assessments were converted to rank order to felicitate comparisons across preference assessment types (e.g.,

Verriden & Roscoe, 2018). The measurement method was the percentage of an item selected. The formula was the number of times that an item was selected divided by the total number that the item was available and multiplied by 100%. The dependent variable in Experiment 2 was the rate of target response during sessions. The total count of the target response was recorded for each session and graphed.

Inter observer agreement (IOA)/ Treatment Integrity

A second observer independently collected data on at least 30% of sessions across Experiments 1 and 2. In Experiment 1, we used a trial-by-trial method to calculate IOA. The formula is # of trials items agreement / # of trials X 100. The average agreement for the selection-based preference assessments was 100% across all participants. The average agreement for eye gaze preference assessment was 80% for Luffy, 82.5% for Cody, and 82.5% for Mike. In Experiment 2, we used total agreement across sessions. The smaller number of responses was divided by the larger number of responses and multiplied by 100. For example, If Observer A scored 3 responses, and Observer B scored 2 responses, IOA would be 67% for this session. The IOA formula is the sum of all sessions' IOA/ # of sessions x100. The agreement was 100%.

Treatment integrity, or the degree to which the study's procedures were implemented as described, was assessed on at least 30% of sessions. The correct steps for each preference and reinforcer assessment are provided in the Treatment

Integrity checklist (Appendix A and B). Treatment integrity was calculated by dividing the number of correct steps by the opportunities to response and multiplying by 100%. Preference assessment treatment integrity averaged 88.2% and reinforcer assessment treatment integrity averaged 100% across all evaluations. For each participant, preference assessment treatment integrity was 88.2% for Luffy, 88.2% for Codey, and 94.1% for Mike.

Chapter 3 Procedure

Experiment 1. The purpose of experiment 1 is to compare the difference and similarity between traditional selection-based PS and MSWO preference assessments and these assessments using eye-gaze. Five items were selected for inclusion by interviewing the participants' case managers and staff members. Prior to each assessment, participants were able to briefly view and manipulate the items. In the selection-based MSWO, researchers put the items in front of participants and participants were told to “pick one.” After participants selected (i.e., touches) an item, they were allowed to engage with it for 20s and the item was removed from the array. This process continued until there were no items in the array or until the participant refused to make any choices. In the eye-gaze MSWO, the same items were presented in an array but out of reach to the participants. The distance between items and the participant was 60 cm (~2 feet). Before the start of the trial, items were blocked from view via a large three-fold poster board (91.44cm x 121.92cm) and then the participants were allowed to view the array. Looking at an item for at least 2 s resulted in delivery of that item to the participant to engage with for 20s. In selection-based PS assessment, we presented items in pairs and provided the verbal instruction “pick one”. When participants picked (i.e., touches) any item, they were allowed to engage with it for 20 s. This process continued until each item was presented with every other item at least once. In PS eye gaze assessment, two items were placed in front of the participants but out of reach. The

distance between items and the patient was 60cm (~2 feet) as well. Items were blocked from view and then participants could view the pair and select one by looking at it for at least 2 s. When they looked at a specific item, the item was presented for them to engage with for 20 s. Each preference assessment method was repeated two to three times to ensure the stability of results.

Experiment 2. The purpose of experiment 2 was to examine the accuracy of the preference assessments by conducting a reinforcer assessment. Sessions were 3 minutes in length and the dependent variable in this experiment was the rate of the target response. The study used a reversal design in which A is baseline, B is low preference condition, and C is high preference condition. The order of conditions was ACBABC for Luffy, ACBABC for Cody, and ACBABCA for Mike. The target response was to touch a red card on the wall. We chose this response as it was a relatively novel response without a long reinforcement history for Luffy and Cody. Although Mike had experience manding for food by touching cards on his desk, touching a card placed on the wall was also a relatively novel response. The position of the card was individualized for different participants. For Luffy and Cody, the position of the card was at knee level. They had to bend over or sit down to reach the card. For Mike, the card was positioned at the same level as his chest. The independent variables were different items that ranged from highly preferred to not preferred. High preference items were defined as items that had an average rank of 1 or 2 across at least 3 of the assessments. Low preference items were defined as

items that had an average rank of 3 or lower across at least 3 of the assessments. Three conditions were used: baseline, high preference, and low preference condition. In baseline, no consequence was provided regardless of participant response. In the high preference condition, highly preferred item was provided contingent with the correct response. In the low preference condition, the process was identical to the high preference condition but with the delivery of the low preference item. Before the start of each session, all participants received full physical prompts to conduct the target response and immediately received the specific condition consequence for 5-10s. That is, in baseline, no items were provided following each response. In the high and low preference conditions, the high and low preference items were delivered according to the condition. The process was repeated 3 times.

At the end of the single reinforcer assessment, progressive ratio (PR) schedule reinforcer assessments were conducted to compare the strength of the high-p and low-p items. In the progressive ratio schedule, participants received the item when they emitted the target response on a progressively increasing schedule starting with PR-1. The schedule then doubled after each delivery of the item as follows: PR-2, PR-4, PR-8, and PR-16. We chose to end the analysis following 16 responses as this was substantially more than the maximum reinforcement schedule used during typical therapy sessions and to mitigate the aversiveness of high reinforcement schedules (Poling, 2010). The session was terminated when the

participant emitted 16 responses or when the participant did not emit the target response for one minute.

Chapter 4 Results

Table 1 provides a summary of highest and lowest preferred items across the preference assessments for all participants. Figure 1 displays a comparison between selection-based and eye-gaze preference assessments for Luffy. In these graphs, rank for eye-gaze preference assessment is plotted on the x-axis and rank for selection-based preference assessment is plotted on the y-axis. If both assessments matched, all data points would fall on the diagonal line. The top panel shows the results from the MSWO assessment and bottom graph shows the results for the PS preference assessment. For Luffy, iPad was picked as the first choice in PS-S and PS-E. Other items generally matched across both assessments. However, car was a low preferred item during the PS-E and it was the second high preferred in the PS-S. The results of the MSWO-S MSWO-E matched in terms of the highest and lowest ranked item with some slight variation on the other items. We chose iPad and car for the reinforcer assessment.

Figure 2 displays the results for the reinforcer assessments for Luffy. In these graphs, sessions are plotted on the x-axis, and total count of the target response are plotted on the y-axis. In the single operant reinforcer assessment (top panel), Luffy had 0 target responses during baseline conditions. He emitted low or no target response in the low preferred conditions. Both high preferred conditions resulted in the highest response which was around 8 times. In the progressive ratio

schedule (bottom panel), Luffy emitted 16 responses during 3 out of 4 sessions. He did not engage in any responding for car.

Figure 3 displays a comparison between selection-based and eye-gaze preference assessments for Cody. The top panel shows the results from the MSWO assessment and bottom graph shows the results for the PS preference assessment. For Cody, among all four different preference assessments, iPad was always selected first. In general, the results from MSWO-S and MSWO-E matched across items. Results of the PS-S and PS-E varied across items. Additionally, the choice of low preferred item was different across assessments. In PS-E, bubble, puzzle toy, and magnet all ranked 2, because the total number Cody picked each of them was exactly the same across 3 repeated PS-E assessments. Due to that, there was no rank 3 and 4 in PS-E. Cups ranked 3 in both MSWO-S and PS-S, but it was the least preferred item in PS-E. Therefore, iPad and Cups were used in the reinforcer assessment.

Figure 4 displays the results of the reinforcer assessment for Cody. In the single operant reinforcement assessment (top panel), the target response in baseline conditions was around 1. In high preferred condition, responding was the highest, which was around 7 per session. In the low preferred condition, the target response initially maintained at 2 per session and then decreased to 0 responses per session during the second implementation of the low-p condition. In the progressive ratio schedule (bottom panel), Cody emitted 16 responses to access iPad twice. Cody

engaged in one response to access cups in the first session and responding decreased to zero upon subsequent low-p sessions.

Figure 5 displays a comparison between selection-based and eye-gaze preference assessments for Mike. The top panel shows the results from the MSWO assessment and bottom graph shows the results for the PS preference assessment. The results of the MSWO-E and MSWO-S matched on the top ranked item (snake toy) but varied across all other items. The results of the PS-S and PS-E generally matched across all items. Although the airplane toy was considered a low preferred item in MSWO-E, PS-S, and PS-E, it was the second-highest preferred item in MSWO-S. Therefore, snake and airplane were used in the reinforcer assessment.

Figure 6 displays the results of the reinforcer assessment for Mike. In the single operant reinforcement assessment (top panel), Mike showed high variability among all conditions. During the initial three phases, therapists noted that Mike's demeanor had changed and he was not eating and interacting with items as he normally did. He was out sick for several days. Upon return, Mike engaged in low and decreasing levels of responding in baseline. He engaged in high levels of responding in both the high-p and low-p conditions. In the progressive ratio schedule (bottom panel), the variability was high as well. Mike engaged in the maximum of 16 responses for both snake and airplane, but he also emitted 0 or 1 response at some point. While Mike continued to emit high and variable responding for the snake toy, his responding decreased to lower levels for airplane.

Chapter 5 Discussion

Preference assessments are an important component of behavioral intervention and needed to identify potential reinforcers. Although a number of variations of preference assessments exist, each has its own advantages and disadvantages. For example, the single stimulus (SS) preference assessment empirically identifies preferred stimuli, and it is easy to conduct. However, it is not only time consuming, but it may also over-identify stimuli as reinforcers, and it does not show a preference hierarchy. Paired stimulus (PS) preference assessment shows a preference hierarchy, but it is time-consuming. Free operant (FO) preference assessment may have a high efficiency compared to PS but it doesn't have a preference hierarchy. Multiple stimuli with and without replacement preference assessment (MSW/ MSWO) can be done quickly and MSWO specifically shows a preference hierarchy, but they may not tell what item is nonpreferred.

Relatedly, selection-based preference assessments may fail to identify potential reinforcers for clients who cannot physically engage in a selection response due to severe to profound disabilities. To overcome these issues, eye-gaze preference assessments have been utilized with individuals with such conditions. Other factors may also prevent selection-based assessments from accurately identifying reinforcers for other individuals receiving ABA therapy. That is, selection of items may be under the control of variables other than motivation for

the item such as location (side bias) or similarity to target materials (history of reinforcement). For example, Bourret et al. (2012) showed that five individuals with ASD and other developmental disabilities engaged in side-bias responding during preference assessments. These biases were only corrected when less preferred items were included in the assessment or repeated error correction trials were conducted.

The purpose of the current experiment was to compare traditional selection-based preference assessments to eye gaze preference assessments. We included a participant with profound disabilities as well as early learners who may lack prerequisite skills for traditional preference assessments. Our results were consistent with previous research on eye-gaze preference assessments. That is, preference assessment by eye gaze method generally matched the results from selection-based assessments. In all participants, items that were defined as high preferred in the eye gaze preference assessment did function as reinforcers in the subsequent reinforcer assessment. In the progressive ratio schedule, the highest preferred item also resulted in the highest breakpoint. However, low preferred items have some differences. For Luffy, the low preferred item car was the second-highest preferred item in the PS. However, car did not function as a reinforcer in the single operant or progressive ratio assessment. Therefore, PS eye gaze was more accurate in his case. For Cody, low preferred item cups were moderately preferred in traditional MSWO. However, contrary MSWO-S showed, cups did not

function as a reinforcer. The most accurate preference assessment for Cody was PS-E. Interestingly for Mike, the second-highest preferred item airplane in the MSWO-S was the lowest preferred item in the MSWO-E and it was a moderately preferred item in PS and PS eye gaze. Airplane did function as a reinforcer during the single operant reinforcer assessment when only one response was required to access the item. During the progressive ratio assessment, the breakpoint also reached the maximum criteria which was 16. However, responding decreased to low levels during the last three sessions of this condition, indicating that airplane did not maintain responding under high response requirements.

In conclusion, for 2 out of 3 participants, the data of eye gaze preference assessments are more accurate than the traditional preference assessment methods. Not only that, but the current study also overcame some shortcomings of previous studies. For example, the target response was brand new for Luffy and Cody. That is, they did not have experience nor reinforcement history on the target response. Next, the progressive ratio was introduced in the current study, which compared the strength difference between the high preferred item and the low preferred item. Last, 2 participants in this experiment were under the age of four, which illustrated that eye gaze method is also applicable to young children with limited skills.

However, there are several disadvantages of using an eye gaze method to assess preference. First, identifying the participants' eye gaze direction and focus point can be difficult to measure. In the current experiment, IOA was relatively

lower in the eye gaze preference assessment, particularly regarding the length of time the client was focusing on an item. Moreover, the eye gaze method may not necessarily be applicable to all types of clients. Although this preference assessment method is suitable for people with severe intellectual disabilities or people with limited motor skills, it may not be suitable for clients with fleeting eye gaze, vision issues, or side biases.

At the same time, this experiment has at least three main limitations. First, in general, Mike's responding was highly variable across conditions. Based on observation, the reason might be due to the ability to differentiate and the environmental factors. For example, Mike would not actively engage with toys when he was sleepy, fatigued, sick, or had a full stomach. There also appeared to be a lack of discrimination between conditions and carry over. For example, it took Mike several exposures to baseline conditions during the single-operant reinforcer assessment to stop responding. It is possible that some conditions needed to be conducted for longer periods of time to ensure accurate results. Second, Luffy used edible in the preference assessments. Edible preference may vary depending on the time of the day and previous research has shown that the inclusion of edibles may mask preference for other items (DeLeon et al., 1997). Third, we only conducted reinforcer assessments for one highly preferred item and one low preferred item. Future studies may use all items in the reinforcer assessment and add more

participants. This allows a complete comparison of selection-based method and eye gaze method.

The current study demonstrates preliminary evidence of the utility of eye gaze preference assessments in young children with an ASD diagnosis. Future research should continue to evaluate these assessments across varied populations to determine if eye gaze preference assessments are a useful tool.

References

Bourret, J. C., Iwata, B. A., Harper, J. M., & North, S. T. (2012). Elimination of position-biased responding in individuals with autism and intellectual disabilities. *Journal of Applied Behavior Analysis*, *45*(2), 241-250. doi: https://doi.org/10.1901/jaba.2012.45-241

Cannella‐Malone, H. I., Sabielny, L. M., & Tullis, C. A. (2015). Using eye gaze to identify reinforcers for individuals with severe multiple disabilities. *Journal of Applied Behavior Analysis*, *48*(3), 680-684. https://doi.org/10.1002/jaba.231

Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). *Applied Behavior Analysis (3rd Edition)*. Hoboken, NJ: Pearson Education.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*, 519-532. https://doi.org/10.1901/jaba.1996.29-519

DeLeon, I. G., Iwata, B. A., & Roscoe, E. M. (1997). Displacement of leisure reinforcers by food during preference assessments. *Journal of Applied Behavior Analysis*, *30*(3), 475-484. https://doi.org/10.1901/jaba.1997.30-475

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (*f*). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*, 491-498. https://doi.org/10.1901/jaba.1992.25-491

Fleming, C. V., Wheeler, G. M., Cannella-Malone, H., Basbagill, A. R., Chung, Y., & Day, K. G. (2010). An evaluation of the use of eye gaze to measure preference of individuals with severe physical and developmental disabilities. *Developmental Neurorehabilitation, 13*(4), 266-275. https://doi.org/10.3109/17518421003705706

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21*(1), 31-43. https://doi.org/10.1901/jaba.1988.21-31

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*(1), 47-58. https://doi.org/10.1901/jaba.2003.36-47

Ivancic, M. T., & Bailey, J. S. (1996). Current limits to reinforcer identification for some persons with profound multiple disabilities. *Research in developmental disabilities, 17*(1), 77–92. https://doi.org/10.1016/0891-4222(95)00038-0

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

Kang, S., O'Reilly, M., Lancioni, G., Falcomata, T. S., Sigafoos, J., & Xu, Z. (2013). Comparison of the predictive validity and consistency among preference assessment procedures: A review of the literature. *Research in Developmental Disabilities, 34*(4), 1125-1133. https://doi.org/10.1016/j.ridd.2012.12.021

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. https://doi.org/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. https://doi.org/10.1002/jaba.239

Poling, A. (2010). Progressive-ratio schedules and applied behavior analysis. *Journal of Applied Behavior Analysis, 43*(2), 347-349. https://doi.org/10.1901/jaba.2010.43-347

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

Roscoe, E. M., Iwata, B. A., & Kahng, S. (1999). Relative versus absolute reinforcement effects: Implications for preference assessments. *Journal of Applied Behavior Analysis, 32*(4), 479-493. https://doi.org/10.1901/jaba.1999.32-479

Trosclair-Lasserre, N., Lerman, D. C., Call, N. A., Addison, L. R., & Kodak, T. (2008). Reinforcement magnitude: An evaluation of preference and reinforcer efficacy. *Journal of Applied Behavior Analysis, 41*(2), 203-220. https://doi.org/10.1901/jaba.2008.41-203

Verriden, A. L., & Roscoe, E. M. (2016). A comparison of preference‐assessment methods. *Journal of Applied Behavior Analysis, 49*(2), 265-28. https://doi.org/10.1002/jaba.302

Table 1:

Summary table for all participants (highest/ lowest preferred item)

		MSWO-S	MSWO-E	PS-S	PS-E
Luffy	**Highest**	**Beans**	**Beans**	**iPad**	**iPad**
	Lowest	Goldfish	Stack ring	Stack ring	Stack ring
Cody	**Highest**	**iPad**	**iPad**	**iPad**	**iPad**
	Lowest	Magnet	Cups	Magnet	Cups
Mike	**Highest**	**Snake**	**Snake**	**Snake**	**Snake**
	Lowest	Ball	Airplane	Airplane+ Ball	Ball

Figure 1

Preference Assessment Comparison for Luffy

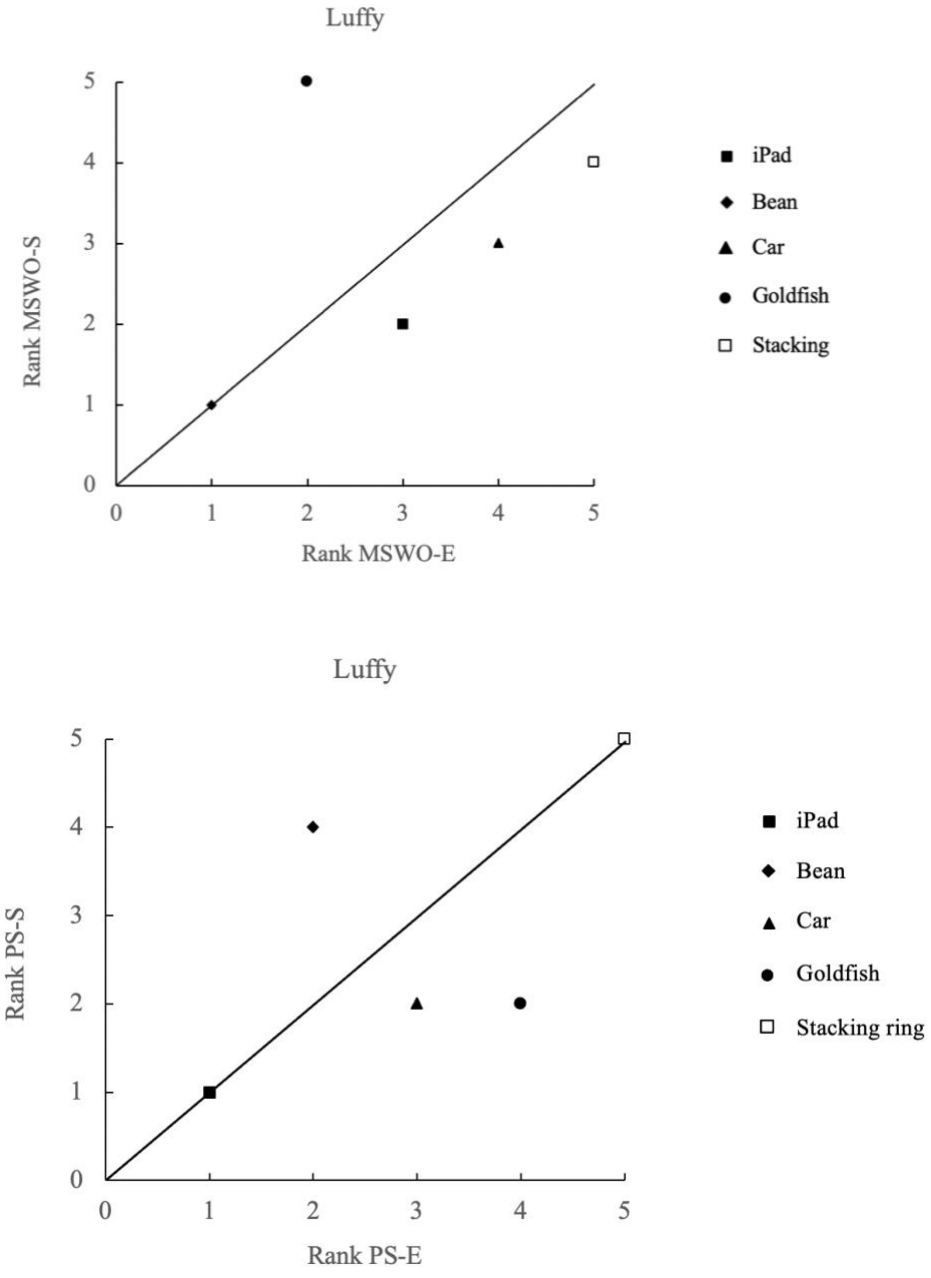


Figure 2

Reinforcer Assessment Results for Luffy

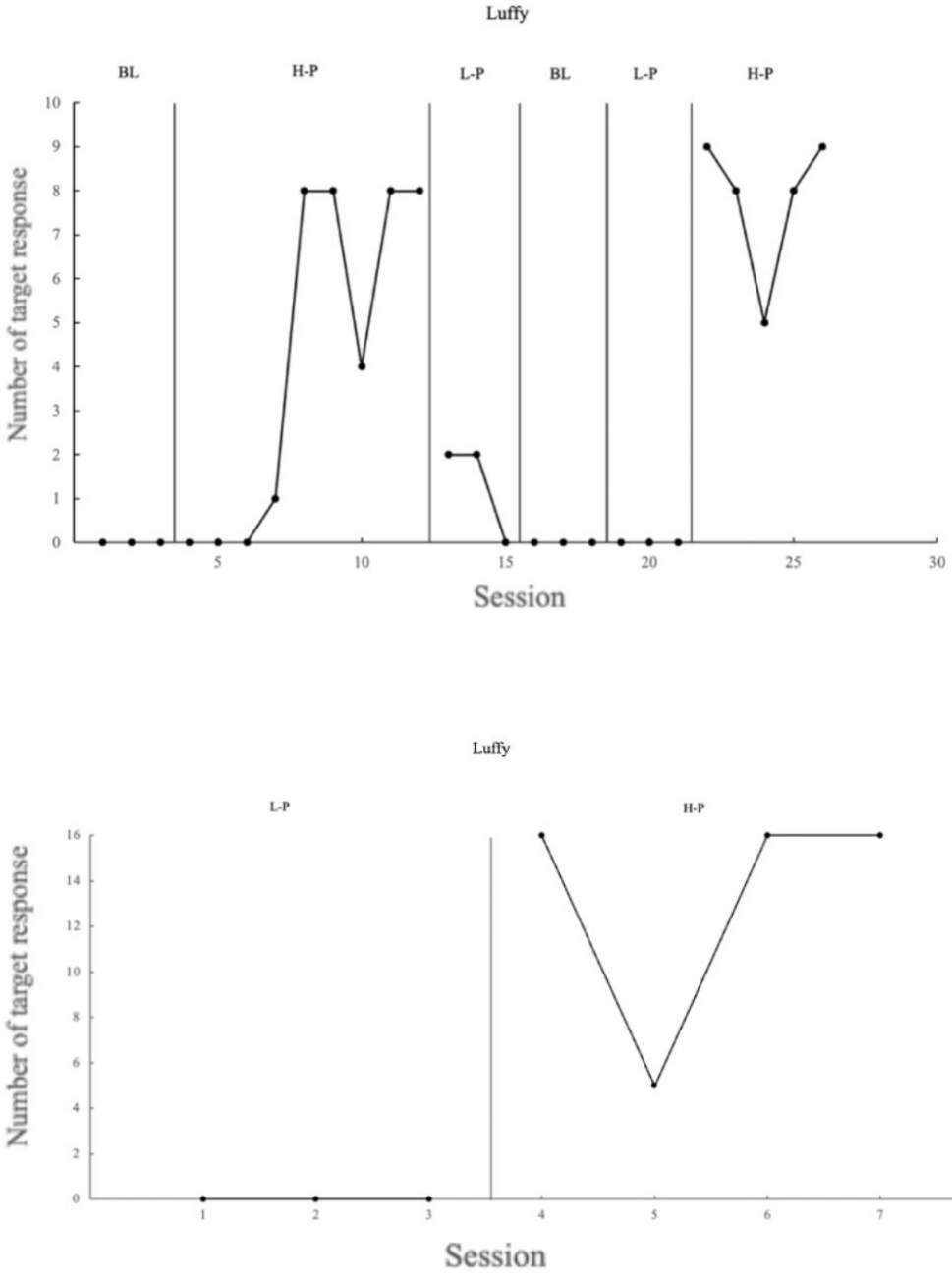


Figure 3

Preference Assessment Comparison for Cody

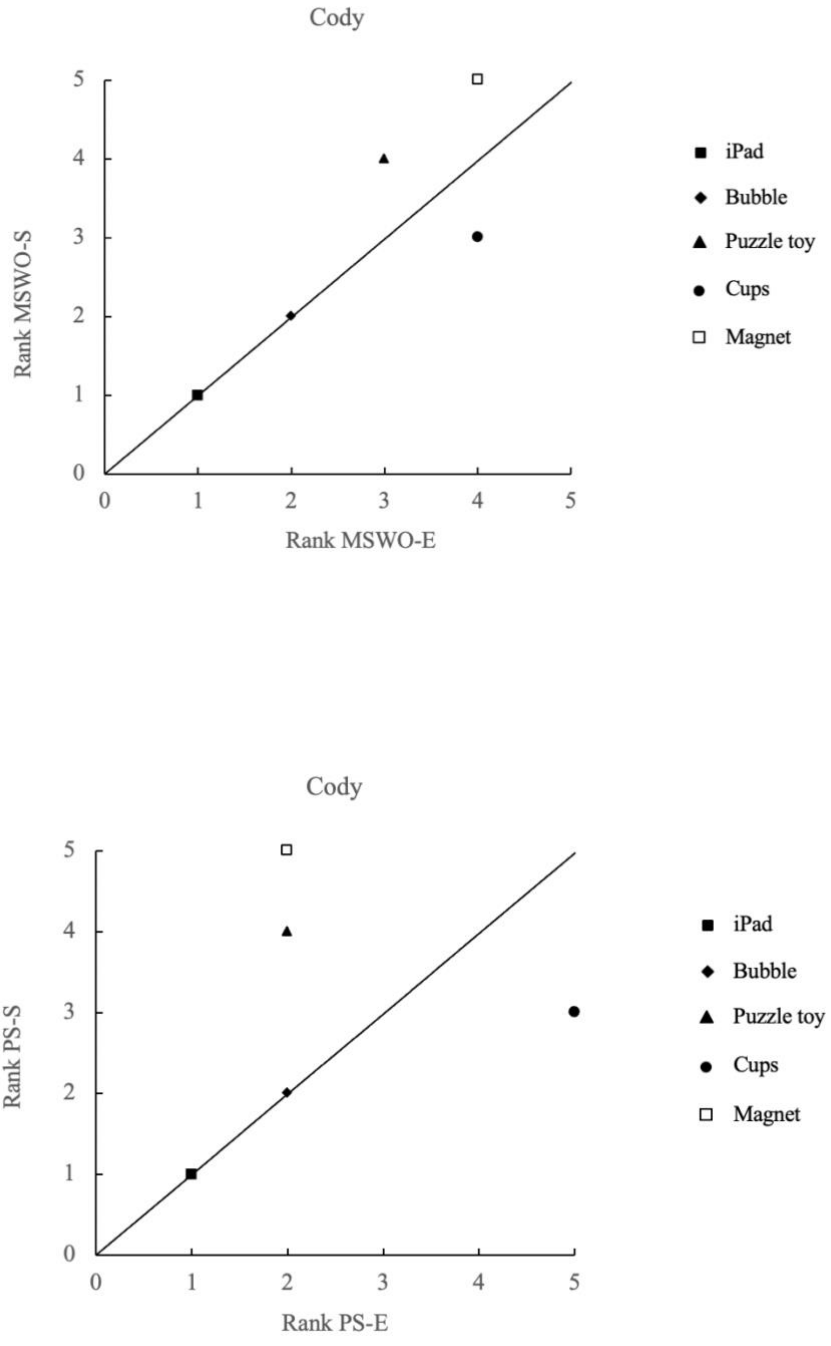


Figure 4

Reinforcer Assessment Results for Cody

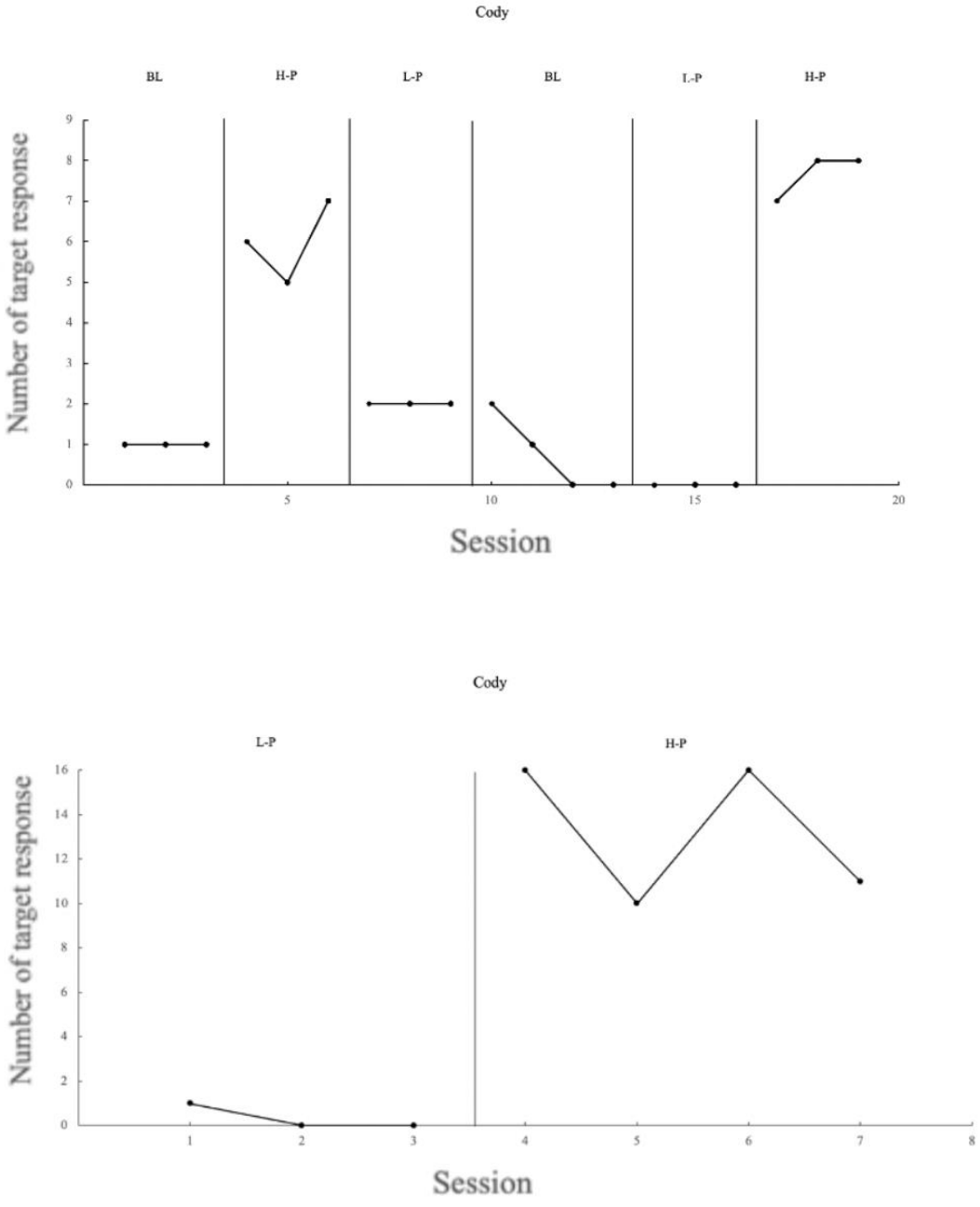


Figure 5

Preference Assessment Comparison for Mike

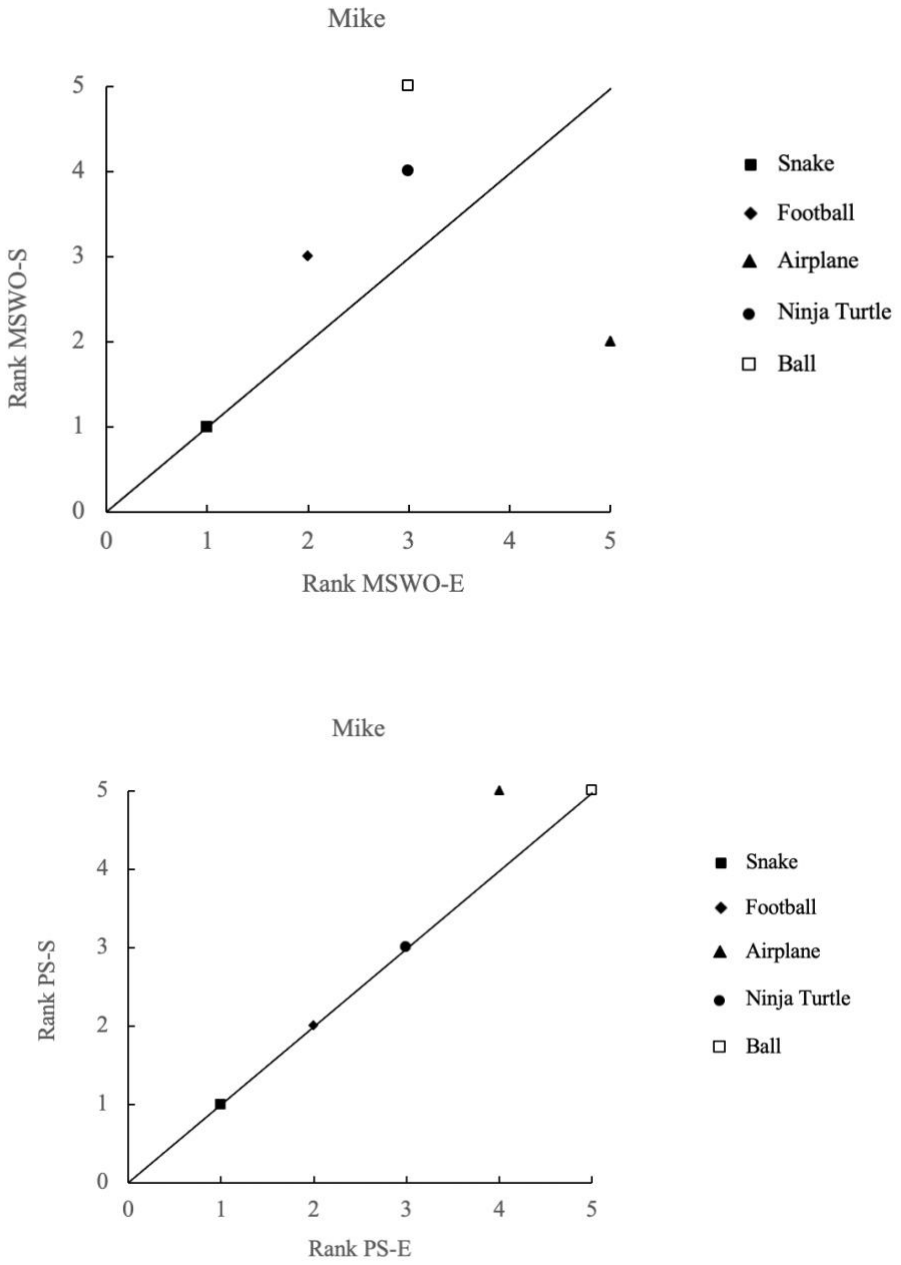
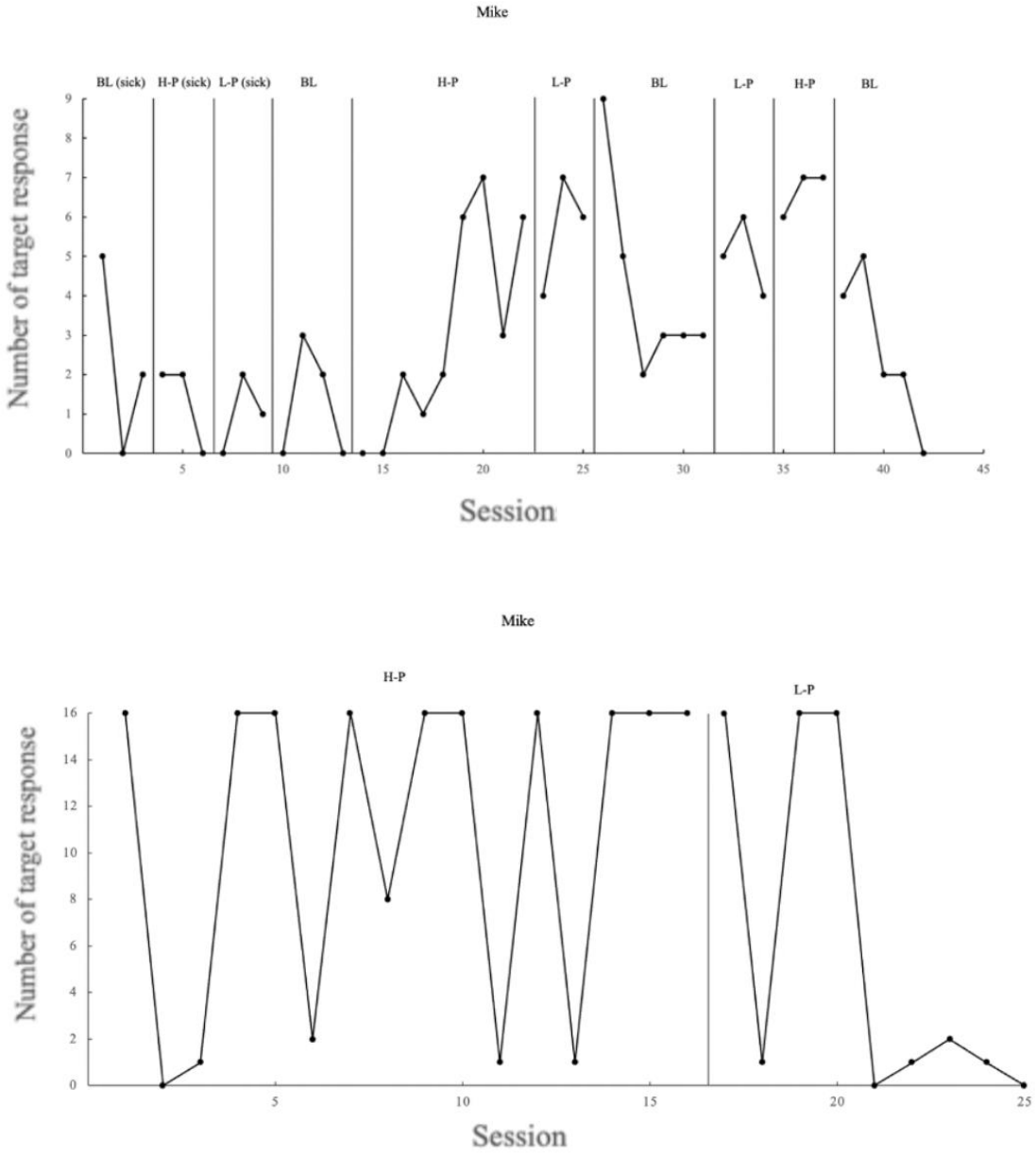


Figure 6

Reinforcer Assessment Results for Mike



Appendix A

Preference Assessments
Treatment Integrity Checklist

Date: 5th of July	**Observer Meagan**

General Procedures **must be completed for each observation*

Prepares for assessment with a variety of toys, activities and/or edibles	Y/ N
Data sheet ready and is completed correctly	Y/ N
Seat client at the table/therapy area	Y/N
Use a board to block sight during eye gaze preference assessments	Y/N
Score	**4/4**

**only evaluate staff performance on the assigned preference assessment*

Paired Stimulus	Assign a code to each stimuli/ prepare data sheet – ensure random pairing/ selection	Y/ N	If no response for 60s, removes both stimuli present again	Y/ N
	Present two stimuli to the child approx.	Y/ N	If reaches for both, re-set the trial	Y/ N
	If approach response, allows engagement for up to 25-seconds	Y/ N	Reaps until all combinations have been presented	Y/ N
			Score	**6/6**
Paired Stimulus (eye gaze)	Assign a code to each stimuli/ prepare data sheet – ensure random pairing/ selection	Y/ N	If no response for 60s, removes both stimuli and present again	Y/ N
	Present two stimuli to the child approx. 2ft apart	Y/ N	If client physically touches stimuli, re-set the trial	Y/ N
	If approach response, allows engagement for up to 25-seconds	Y/ N	Reaps until all combinations have been presented	Y/ N
	Client only gets the item when gazing for at least 2 s	Y/ N	*record not selected for stimuli not chosen	Y/ N
			Score	**6/8**
MSWO replacement Preference Assessment	Place stimuli in a randomly sequences straight line	Y/ N	MSWO *remove the selected item from the array	Y/ N
	Instruct the child to select 1 and allow access for up to 25-seconds	Y/ N	Rotate stimuli placement	Y/ N
	Block access to the other stimuli during the period of engagement	Y/ N	Continue until all items are selected OR until the child has no response to the last item twice	Y/ N
	If client reaches for multiple stimuli, re-set the trial	Y/ N	*record not selected for stimuli not chosen	Y/ N
			Score	**8/8**
MSWO replacement Preference Assessment (eye gaze)	Place stimuli in a randomly sequences straight line, 2 ft apart	Y/ N	MSWO remove the selected item from the array	Y/ N
	Instruct the child to select 1 and allow access for up to 25-seconds	Y/ N	Rotate stimuli placement	Y/ N
	Client only gets the item when gazing for at least 2 s	Y/ N	Continue until all items are selected OR until the child has no response to the last item twice	Y/ N
	If client physically touches stimuli, re-set the trial	Y/ N	*record not selected for stimuli not chosen	Y/ N
			Score	**6/8**

Assessment Scoring

Percentage Score 88.2% (30/34)	**(Score PF / Total Steps) x 100**

Appendix B

Reinforcer Assessments
Treatment Integrity Checklist

Date: June 26	Observer: Meagan

General Procedures **must be completed for each observation*

Prepares for assessment with a variety of toys, activities and/or edibles	Y/ N
Data sheet ready and is completed correctly	Y/ N
Place client at the therapy area	Y/N
The red card is placed at the right place	Y/N
Score	**4/4**

Reinforcer Assessment	Pre-teach before the beginning of each session	Y/ N	Deliver the item immediately when the client emits the target response	Y/ N
	If emit response, allows engagement for up to 20-seconds	Y/ N	Per session lasts 3 min	Y/ N
			Score	**4/4**

Progressive Ratio schedule	Pre-teach before the beginning of each session	Y/ N	Deliver the item immediately when the client emits the correct number of target response (eg. 1, 2, 4, 8, 16)	Y/ N
	If emit response, allows engagement for up to 20-seconds	Y/ N	The session ends when the client reaches 16 breakpoint or does not emit target response for 60s	Y/ N
			Score	**4/4**

**only evaluate staff performance on the assigned preference assessment*

Assessment Scoring

Percentage Score **100% 12/12**	**(Score PF / Total Steps) x 100**