The Dissertation committee for Heather Koch Gonzales certifies that this is the approved version of the following dissertation:

An Evaluation of Multiple Stimulus With Replacement Preference Assessment Variations: Effects on Motivation

Committee:

Mark O'Reilly, Supervisor

Terry Falcomata

Nina Zuna

Russell Lang

An Evaluation of Multiple Stimulus With Replacement Preference Assessment Variations: Effects on Motivation

by

Heather Koch Gonzales, B.A.; M.A.T.; M.Ed.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2016

Dedication

This dissertation is dedicated to my husband, Edward Gonzales, who first encouraged me to pursue a doctoral degree, and upheld me through every milestone and setback. He has challenged me to push myself intellectually and always perform to the best of my abilities. Secondly, I dedicate this dissertation to my parents, Robert and Candice Koch. Their unwavering support and encouragement made the completion of this project possible. They first taught me that the pursuit of learning and discovery is a noble cause, and served as my constant cheerleaders during this very long journey in academia.

Acknowledgements

The successful completion of a doctoral degree reflects not only the efforts of the candidate, but the instruction, feedback, encouragement, and support of the faculty of the conferring institution. I would like to acknowledge my advisor, Dr. Mark O'Reilly, for encouraging me to apply to this program, giving me a strong foundation of subject matter knowledge, and making my improvement as a practitioner and academician his priorities above all else in my development. I would also like to acknowledge my committee, comprised of Drs. Terry Falcomata, Russell Lang, and Nina Zuna for ensuring that my experiment had a solid rationale and methodology, and for their feedback throughout the process of researching and writing.

The staff at William's Community School deserves my gratitude for allowing me into their classrooms and to conduct research with their students. I would like to especially acknowledge Suzanne Byrne and Kara Huss for their input and time, and for always making me feel welcome.

In addition, I would like to acknowledge Katherine Hoffman for being a sounding board throughout both the formative stages and completion of this project. Our dissertation sessions and periodic walk breaks helped me push through end of this process. Along with Katherine, I would like to acknowledge Abby Hodges, Samantha Swinnea, and Nicolette Caldwell for conducting my Interobserver Agreement and fidelity evaluations efficiently and with great care.

Finally, I would like to thank the families of my participants, who trusted me to work with their children in an experimental context. The parents of Connor and Ethan deserve my special thanks for allowing me into their home, accommodating my schedule, and being especially welcoming and kind.

An Evaluation of Multiple Stimulus With Replacement Preference Assessment Variations: Effects on Motivation

Heather Koch Gonzales, PhD

The University of Texas at Austin, 2016

Supervisor: Mark O'Reilly

This experiment evaluated the differential effects of three Multiple Stimulus With Replacement preference assessment variations on the motivation of four participants with developmental disabilities to respond to a mastered task. No discernible difference was demonstrated between the Five Pre-Session Choices condition and the Five Within-Session Choices conditions, suggesting that the timing of choice opportunities did not affect motivation, demonstrated through lack of differentiation in a multi-element experimental design for the dependent variables of Latency to Task Initiation, Total Task Duration, Percentage Correct Responding, or No-Responses. Subsequently, a One Pre-Session Choice condition was initiated to evaluate the efficacy of a preference assessment method with lower practitioner response effort. The introduction of this condition also resulted in undifferentiated data. Implications and future research are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1: *Introduction*

It is now well-established that people with developmental disabilities are capable of making purposeful choices in a variety of contexts within their daily lives, including for food and drinks, activities, social interaction, and tactile or auditory stimulation (Hall, Morgan, & Salzberg, 2014; Lancioni, O'Reilly, & Emerson, 1996; Lancioni, Bellini, & Oliva, 1993; Mithaug & Hanawalt, 1978; Mithaug & Mar, 1998; Newton et al., 1993; Nurenberger, Smith, Czapar, & Klatt, 2012; Philips & Mudford, 2011; Reid & Parsons, 1991). The provision of choice-making opportunities can have a positive impact on the lives of individuals with developmental disabilities in a number of ways. First, it has been demonstrated that for some individuals, the provision of choice results in lower levels of challenging behavior when compared to conditions in which choice is not provided (Bambara, Koger, Katzer, & Davenport, 1995; Cole & Levinson, 2002; Dunlap et al., 1994; Dunlap, Kern-Dunlap, Clarke, & Robbins, 1991; Dyer et al., 1990; Seybert, Dunlap, & Ferro, 1996). Second, on-task behavior has been shown to increase when individuals are given the opportunity to choose their task or materials (Dunlap et al., 1991; Dunlap et al., 1994; Parsons & Reid, 1990; Seybert et al., 1996). Third, individuals may appear more capable when actively making choices, which may promote more positive attention from those around them, as well as a higher level of acceptance from persons within their environment (Lancioni et al., 1996). In general, the opportunity to choose tasks and reinforcers may provide individuals with a sense of control that increases motivation and neutralizes conflicts that lead to challenging behavior (Baer, Tishelman, Degler, Osnes, & Stokes, 1992; Bowers, Clement, Fantuzzo, & Sorensen, 1985; Lancioni et al., 1996; Stancliffe, 1995). An individual's choices are typically based on his or her preferences, so an awareness of preference is key to providing meaningful choice opportunities. As Cannella et al. (2004) pointed

out, when offered the choice of a ham or turkey sandwich, an individual would likely choose the meal he or she preferred at that moment. As that scenario implies, preferences may not always be static, so it is important for practitioners to be aware of stimuli that may potentially be reinforcing at any given moment, so as to provide choices that can promote significant behavior change. In this paper, preference refers to the "subjective liking or disliking of a particular item or person" (Kearney & McKnight, 1997, p. 219).

Over the past 30 years, a body of literature has expanded surrounding the technology of preference assessments (e.g., Ahearn, Clark, DeBar, & Florentino, 2005; Conyers et al., 2002; Davies, Chand, Yu, Martin, & Martin, 2013; Didden, Korzilius, Sturney, Lancioni, & Curfs, 2008; Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992; Lim et al., 2001; Snyder, Higbee, & Dayton, 2012; Zhou et al., 2001). Although there are many types of preference assessments, the two that will be relevant to this study are the Paired Choice (PC) preference assessment and the Multiple Stimulus With Replacement (MSW) preference assessment. In the PC preference assessment, two items are presented at a time to the participant and the selection of the participant is recorded. Each possible combination of items is presented in a randomized sequence in which position bias is controlled for by systematically varying on which side (i.e., left or right) each item is placed (Fisher et al., 1992). In the MSW preference assessment, all stimuli are presented initially. Upon selection of the first stimulus, that stimulus is presented for a predetermined period of time and then subsequently replaced in the array in a randomized position, such that all arrays contain identical items, randomly rotated to prevent position bias (DeLeon & Iwata, 1996).

Multiple studies have investigated how to reduce the duration of preference assessment methodologies while still producing a valid assessment of preference (e.g., DeLeon & Iwata,

1996; Roane, Vollmer, Ringdahl, & Marcus, 1998; Windsor et al., 1994). Reducing the time required to assess preferences may be significant for several practical reasons, making them both more appealing to practitioners and more effective in their designed purpose. First, although research demonstrates that provision of choice, in conjunction with the assessment of preference, has a positive impact on individuals' lives (Cannella et al., 2004; Lancioni et al., 1996), preference assessment methods can be considered effortful by practitioners (Matson, Bielecki, Mayville, Smalls, Bamburg, & Baglio, 1999). Brief preference assessment formats may be deemed more appealing, thereby increasing chances that practitioners will make the effort to administer them. Perhaps most importantly, brief preference assessments may be conducted with greater frequency, and may therefore capture momentary variations in preference. Studies have demonstrated that preferences for reinforcers are not fixed, but rather may change across time, such that highly preferred stimuli at one point in time may not be preferred at a later time, and vice versa (Dyer, 1987; Kennedy & Haring, 1994; Roane et al., 1998). Assessing preference frequently may ensure that the stimuli used in interventions are, in fact, the most highly preferred stimuli at that moment, thereby potentially leading to more effective use of reinforcement.

In order to assess the effect of frequent, brief preference assessments on reinforcer efficacy, DeLeon et al. (2001) compared response allocation of participants when reinforcers were chosen via a lengthy PC preference assessment prior to the experiment, to when reinforcers were chosen daily via a brief, MSWO preference assessment. After selecting seven to nine items on the Reinforcer Assessment for Individuals with Severe Disabilities (RAISD; Fisher, Piazza, Bowman, & Amari, 1996), interventionists conducted a complete PC preference assessment and ranked items according to the percentage of trials in which they were selected. The item with the highest percentage was designated as the lengthy PC item. A simple task that could be readily

completed was then selected for each participant. Participants were trained to complete the task that was set before the reinforcer with which they would prefer to engage. After training was completed, daily preference assessments were initiated. For daily assessment sessions, a brief, one-presentation MSWO preference assessment was conducted. The first item chosen during the preference assessment was designated as that day's daily-MSWO item. On days on which the daily-MSWO item differed from the lengthy-PC item, three identical tasks were placed on the table in front of the participant. Completion of each task was associated with one of three outcomes, depending on which task was completed: 30 s of access to the daily-MSWO item, 30 s access to the lengthy-PC item, or a 30 s interval with no access to either item (control task). Results indicated that on days when results of the two assessments differed, participants generally allocated more responses to the task associated with the daily top-ranked item than the item chosen from the lengthy-PC preference assessment. Interventionists concluded that, for many individuals, preferences can change from day to day, and that these changes in preference are usually associated with corresponding changes in response allocation when examined under a concurrent operant arrangement. Their results suggest that frequent, brief assessment of preference may reveal reinforcers that have greater efficacy, due to a participant's higher momentary preference.

Elliott and Dillenberger (2014) investigated the effects of response-contingent choice versus no choice among highly preferred reinforcers on motivation within the context of a discrete trial training (DTT) program. The authors drew on research by Fisher, Thompson, Piazza, Crosland, and Gotjen (1997), who concluded that providing choices "adds to the reinforcement value of the chosen stimuli and thus may have the potential to improve the effectiveness of the behavioural programme" (p. 433). The purpose of the study by Elliott and

Dillenberger (2014) was to determine if the introduction of reinforcer choice in DTT would have a motivating effect for learning. Prior to the study, parent and therapist interviews were used to construct a list of 12 potentially reinforcing stimuli for each child. A PC preference assessment was used to narrow the pool of reinforcing items to six. At the start of each session, a very brief preference assessment was conducted to identify the three stimuli that were most highly preferred on that day. An alternating treatments design compared choice and no-choice conditions. During the no-choice condition sessions, the interventionist chose one of the three highly preferred items and presented it to the child, contingent upon correct responses. During the choice sessions, the interventionist provided the participant with a choice of the three highly preferred reinforcers contingent upon correct responses. Results indicated that for two participants, skill acquisition, levels of non-responding, and time to return to table were comparable in choice and no-choice conditions. However, for the third participant, slight increases in skill acquisition were demonstrated in the choice condition, and considerable increases of non-responding were demonstrated in the no-choice condition. Furthermore, the third participant demonstrated longer response latency to the instruction, "Come sit" after engaging with the stimulus. Interventionists concluded that for some participants, response-contingent choice might increase motivation for completing tasks. Within a DTT format, the understanding that a choice of reinforcers will be provided contingent upon a specified number of responses may create an establishing operation (EO) that increases accurate responding and reduces challenging behavior. An EO is any event that increases the effectiveness of a given consequence (e.g., the delivery of a specific reinforcer; Laraway, Snycerski, Michael, & Poling, 2003).

This experiment extended the research conducted by DeLeon et al. (2001), while incorporating findings from Elliott and Dillenberger (2014). DeLeon et al. (2001) found that administering brief, but more frequent preference assessments led to greater response allocation, an indicator of how reinforcing the selected stimulus was at that moment. If daily preference assessments determined reinforcers that led to greater motivation, perhaps more frequent assessment of preference during a session could increase motivation to an even greater extent. Conversely, infrequent assessment of preference would result in the repeated presentation of the same stimulus, which researchers (Berlyne, 1960; Fowler, 1971; Glanzer, 1958; Myers and Miller, 1954) have demonstrated may result in satiation, defined as, "reinforcers losing their ability to increase, or maintain, behavior when they have been consumed" (DeMarse, Killeen, & Baker, 1999, p. 324). Such a phenomenon would indicate that an abolishing operation (AO), defined as any event that decreases the effectiveness of a given consequence, was in effect. Therefore, motivation, as established by the parameters of the experiment, would be lower (Laraway, Snycerski, Michael, & Poling, 2003). In previous experiments, Egel (1980; 1981) examined the effect of reinforcer variation on responding, and found that in the event of satiation, high rates of responding could be renewed by making changes in the consequence event (i.e., delivering different reinforcers). Therefore, providing the opportunity to access varied reinforcers through the mechanism of frequent preference assessments could limit satiation and maintain motivation to respond. Furthermore, Elliott and Dillenberger (2014) found that providing a choice of highly preferred reinforcers, rather than no choice, increased motivation in the form of greater response accuracy and shorter response latency (returning to the table after a break). They determined that response-contingent choice may increase motivation for task

completion, and suggested that future research continue to explore the effects of providing response-contingent choice within a DTT format.

Given previous findings, this experiment compared motivation, as measured by latency to initiation of the task, total task duration, percentage of correct responding, and frequency of "no-responses" between two conditions: one in which a daily, five-trial, MSW preference assessment was conducted prior to the session and the designated reinforcer held constant, and one in which a daily, five-trial, MSW preference assessment was spaced evenly throughout the session. The first condition, referred to as, "Five Pre-Session Choices" condition hereafter, was intended to emulate the format of the daily, brief, MSW preference assessment used in the DeLeon et al., (2001) experiment. The second condition, referred to as, "Five Within-Session Choices" condition from now on, was intended to incorporate the findings of Elliot and Dillenberger (2014), specifically, the inclusion of response-contingent choice. In addition, this condition was structured such that the participants had the opportunity to access multiple reinforcers throughout the session, in the event that satiation with the present reinforcer produced an abolishing effect.

A multi-element design was used to determine whether the combination of opportunities to select different stimuli throughout a session and response-contingent choice would produce greater motivation among participants than reinforcement with a stimulus chosen prior to the session and held constant. In addition to potentially increasing motivation by reducing the probability of an AO being in effect, the Five Within-Session Choices condition would provide practitioners with a method of assessing preference that is brief and occurs within the context of instruction, thereby not taking extensive time away from the instructional period for assessment. If effective at increasing motivation and determining momentary preferences, a brief, response-

contingent preference assessment that is evenly spaced throughout the instructional period could be an appealing option for practitioners.

RESEARCH QUESTIONS

This study will address the following research questions:

1. Will offering choices for reinforcers throughout a session lead to greater motivation in the form of lower latency to task initiation and total task duration, higher percentage of correct responding, and lower frequency of "no-responses"?
2. Can time-based dependent variables (latency to task initiation and total task duration) effectively evaluate motivation in a DTT context?
3. If differentiation between conditions in a multielement design does not occur, will an even briefer preference assessment format produce comparable results, so as to reduce response effort for practitioners?

Chapter 2: Literature Review

The purpose of this chapter is to review the literature concerning choice and preference assessment studies for individuals with developmental disabilities. Three previous literature reviews have addressed this topic (Cannella et al., 2004; Lancioni et al., 1996; Tullis et al., 2011), with the most recent published in 2011. This literature review is an update on the research done regarding choice and preference assessments with people who have developmental disabilities from 2011-2014. The review will be organized as follows. The first section of this chapter is the Method section, and describes inclusion and exclusion criteria, provides the search procedures, and explains how studies were coded. The next section of this chapter is the Results section, and categorizes the included studies, describes the nature of each category, and provides 1-2 examples in detail. The final Discussion section provides comments on the included studies, presents limitations of this review, and suggests areas for future research.

Method

Search Procedures and Selection Criteria

A literature search was conducted in the databases Academic Search Complete, ERIC, Medline, and PsychINFO from 2011-2014. The search terms "developmental disabilities" or "developmental disability" and "autism" or "autistic" were combined with the search terms "preference", "preference assessment(s)", and "choice". In addition, the search term "choice" was combined with the search terms "challenging behavior" and "problem behavior". The title and abstract of each study was read to determine whether it examined choice or preference assessments with people who have developmental disabilities. If the abstract suggested that the

study might meet the inclusion criteria, it was retained to be read more carefully. The reference sections of retained studies were also reviewed to ensure a more complete list of studies.

In order to be included in this review, a paper had to (a) be an experiment concerning either choice or preference; (b) include at least one participant with a developmental disability; and (c) be published between 2011-2014. Studies in which a preference assessment was used, but in which choice or preference were not the focus of the experiment were excluded.f Studies were coded by the first author using a coding sheet (see Figure 28 in the Appendix) that included number of participants, age of participants, diagnoses of participants, stimuli used in assessment, type of preference assessment, dependent variable(s), and experimental design.

Results

Overview of Studies

The initial search yielded 594 articles. Of these articles, 26 met all of the inclusion criteria. The 26 included studies have been grouped according to similar methodologies. See Table 1 for information regarding number of participants, diagnoses of participants, stimuli used, presentation method, dependent variables, experimental design, and results. Studies in the first category examined the effectiveness of preference assessments. In this category, positive findings indicate that the preference assessments identified items that functioned as reinforcers for all participants. Mixed findings indicate that the preference assessments identified items that functioned as reinforcers for some, but not all, participants. In order for a study to be grouped in this category, a reinforcer assessment must have been conducted in addition to preference assessments. Studies in the second category assessed various parameters of preference without testing reinforcer effectiveness. In this category, positive findings indicate that the preference assessments successfully identified a hierarchy of reinforcers for all participants. Mixed findings

indicate that the preference assessments identified a hierarchy of reinforcers for some, but not all, participants. Studies in the third category examined the effects of the provision of choice on behavior. In this category, positive findings indicate that the provision of choice led to positive changes in behavior (i.e., decreases in challenging behavior or increases in adaptive behavior) for all participants. Mixed findings indicate that the provision of choice led to positive changes in behavior in some, but not all, participants. Studies in the fourth category compared preference assessment types in terms of the effectiveness of the reinforcers they identified. In this category, positive findings indicate that when two different preference assessment types were compared, the most effective reinforcers were identified by the same method for all participants. Mixed findings indicate that when two different preference assessment types were compared, the most effective reinforcers were not identified by the same method. In order for a study to be included in this category, a reinforcer assessment must have been conducted. Studies in the fifth category examined protocols for training individuals to administer preference assessments. In this category, positive findings indicate that all staff members were trained to assess preference to mastery criterion. Mixed findings indicate that some, but not all, staff members were trained to assess preference to mastery criterion. Studies in the sixth category examined the effects of using preference in an intervention for challenging behavior. In this category, positive findings indicate that assessing preference had a positive effect on behavior for all participants. Mixed findings indicate that assessing preference had a positive effect on behavior for some, but not all, participants. Studies in the seventh category assessed various parameters of choice making. In this category, positive findings indicate that all participants demonstrated a preference for a choice-making procedure over a no-choice procedure. Mixed findings indicate that some, but not all, participants demonstrated a preference for a choice-making procedure over a no-choice

procedure. Studies in the final category examined the effects of assessment type on challenging behavior. The nature of this study was exploratory, rather than treatment-oriented, so the designation of findings as positive or mixed was not considered applicable.

The Effectiveness of Preference Assessments

Eleven studies examined the effectiveness of preference assessments (Call, Shillingsburg, Bowen, Reavis, & Findley, 2013; Clay, Samaha, Bloom, Bogoey, & Boyle, 2013; Graff & Larsen, 2011; Hall et al., 2014; Kenzer, Bishop, Wilke, & Tarbox, 2013; Keyl-Austin, Samaha, Bloom, & Boyle, 2012; Mangum, Frederick, Pabico, & Roane, 2012; Nuernberger, Smith, Czapar, & Klatt, 2012; Russo, Tincani, & Axelrod, 2014; Tullis, Cannella-Malone, & Fleming, 2012). Clay et al. (2013) examined a procedure to assess preference for different types of social interactions for participants with intellectual and developmental disabilities. For each participant, four forms of social attention that consisted of both physical contact and a vocal statement were identified through an open-ended informal interview. A total of nine forms of social attention were identified across participants and an individual interventionist was assigned to each form of attention exclusively. A brief exposure to each form of attention was followed by a PC preference assessment in which the interventionist assigned to each form of attention stood on opposite ends of a wall. The form of social attention selected most frequently was then evaluated in a reinforcer assessment in which a simple response was reinforced by the social attention on a Fixed-Ratio 1 (FR1) schedule of reinforcement. Results of the reinforcer assessment indicated that the form of social attention selected most frequently functioned as a reinforcer for all participants. Interventionists concluded that this procedure successfully identified forms of social attention that functioned as reinforcers.

Graff and Larsen (2011) used PC preference assessments and reinforcer assessments to assess the effectiveness of high-preferred, low-preferred, and novel stimuli for children with ASD. First, interventionists asked staff at the school the participants attended to fill out the RAISD to generate a list of eight preferred edible items. Next, interventionists conducted two paired-stimulus preference assessments (PS-1) with each participant using the eight items identified from the RAISD. Relative reinforcer efficacy was assessed for the most-preferred and least-preferred stimuli identified from PS-1 using a reinforcer assessment (RA-1). Following RA-1, a second preference assessment (PS-2) was conducted using the least-preferred stimuli from PS-1 and seven novel stimuli. Finally, a second reinforcer assessment (RA-2) was conducted using the most-preferred and least-preferred stimuli from PS-2. Results from RA-1 indicated that, for all five participants, both high-preferred and low-preferred stimuli functioned as reinforcers, producing nearly identical rates of responding. In RA-2, the items classified as most preferred (those classified as least-preferred in PA-1) functioned as reinforcers, but the novel stimuli selected by interventionists which were classified as least-preferred did not. Interventionists concluded that preference rankings are directly influenced by the composition of the array and that low-preference items may function as reinforcers.

Assessing Preference

Four studies assessed preference in people with developmental disabilities (Davies, Chand, Yu, Martin, & Martin, 2013; Kenzer & Bishop, 2011; Snyder, Higbee, & Dayton, 2012; Verschuur et al., 2011). Davies et al. (2013) compared the results of PC preference assessments and MSWO preference assessments to assess the stimulus preferences of adults with developmental disabilities. In the PC preference assessment, interventionists assessed six food items and ran the total assessment two times for a total of 60 trials. In the MSWO preference

assessment, interventionists assessed the same six items and ran the total assessment two times for a total of 60 trials. Results indicated both preference assessments identified the same stimuli as least preferred for all participants and identified the same stimuli as most preferred for five out of nine participants. Furthermore, the MSWO identified the stimuli that were ranked second in the paired choice preference assessment as most preferred for two participants, and identified the item ranked third in the paired choice preference assessment as most preferred for the remaining two participants. Correlations between items chosen on each assessment were moderate to high. Interventionists also conducted a post hoc analysis that suggested two patterns of responding. For the four participants who did not select the same stimuli as most preferred in both assessments, preferred stimulus selection seemed to increase on MSWO trials with fewer stimuli. For participants who selected the same stimuli as most preferred in both assessments, all participants chose the most preferred stimulus well above chance-level in the first two trials of a five- or six-choice MSWO. As the number of choices decreased, participants demonstrated a decline, relative to chance, in preferred stimulus selection.

Kenzer and Bishop (2011) examined participant preference for novel and familiar stimuli in children with ASD. Using an open-ended written interview, interventionists collected information about staff-reported high-preference stimuli, staff-reported low-preference stimuli, and experimenter-selected novel stimuli. Interventionists then conducted two PC preference assessments. The first consisted of high- and low-preference stimuli selected by the staff and the second consisted of high-preference stimuli selected by the staff and novel stimuli selected by the interventionists. Four outcomes of the two paired choice preference assessments were possible: new preferred stimuli were identified in (1) neither assessment; (2) the high/low assessment; (3) the high/novel assessment; or (4) both assessments. For two of the 31

participants, additional preferred stimuli were identified only during the high/low preference assessment. For 16 of the 31 participants, additional preferred stimuli were identified during the high/novel preference assessment. Finally, for nine out of 31 participants, additional preferred stimuli were identified in both assessments. Interventionists concluded that there was poor agreement between indirect and direct preference ratings and that restricting preference assessments to only familiar stimuli may limit the number of preferred stimuli identified.

The Effects of Choice on Behavior

Four studies examined the effects of providing choice on the behavior of people with developmental disabilities (Elliott & Dillenburger, 2014; Lough, Rice, & Lough, 2012; Phillips & Mudford, 2011; Rispoli et al., 2013). Lough et al., 2012 investigated the effect of providing choice in a coloring task with children ASD. A choice and a no-choice condition were presented to children at least one week apart. In the choice condition, participants were presented with three templates of a face and were asked to choose one. Whichever template they chose was placed in front of the participant, and he or she was instructed to take as much time as necessary to color the picture with the provided markers. In the no-choice condition, the interventionist provided the template without a choice, and again instructed the participant to take as much time as needed to color it. In both conditions, total time coloring was measured. In addition, the drawing was scanned into the computer in order to calculate the amount of pixels colored. Results indicated that there was no statistically significant difference between the amount of pixels colored in each condition, but participants used significantly more colors and spent significantly more time coloring in the choice condition. Interventionists concluded that the provision of choice increased occupational engagement in the participants in their study, and that

providing choices in a task may increase occupational engagement in children with ASD in general.

Rispoli et al. (2013) examined the effects of both across-activity and within-activity choices on the escape-maintained challenging behavior of children with ASD. First, interventionists assessed each participant using the QABF (Matson, Bambrug, Cherry, & Patlawskij, 1999) to determine the function of their challenging behavior. For each participant, results indicated that challenging behavior was maintained by negative reinforcement. Next, stimuli associated with six instructional activities were evaluated using an MSWO preference assessment. The four lowest ranked stimuli were selected for use in further sessions. This study used an alternating treatments design of choice within and across activities, embedded within a reversal design. The across-activity choice condition consisted of placing either two or four activities in front of the participant and instructing him or her to choose one. The session then proceeded using the selected activity. The within-activity choice condition consisted of the interventionist choosing the activity but providing the participant with choices about how to complete the activity. Results demonstrated that rates of challenging behavior were higher in baseline conditions for all participants. For three participants, lower rates of challenging behavior were exhibited when choice was provided across activities. For the remaining participant, rates of challenging behavior were undifferentiated between conditions. Interventionists concluded that providing across-activity choices might be more effective than providing within-activity choices at reducing escape-maintained challenging behavior in children with ASD.

Comparisons of Preference Assessment Types

Two studies compared the effectiveness of different preference assessment types (Call, Trosclair-Lasserre, Findley, Reavis, & Shillingsburg, 2012; Kelly, Roscoe, Hanley, &

Schlichenmeyer, 2014). Call et al. (2012) examined the effectiveness of reinforcers identified through a single PC preference assessment or a daily MSWO preference assessment using a progressive ratio reinforcement schedule. First, each participant was engaged in a PC preference assessment with either six or seven items. After the PC preference assessment, the target response was taught using a four-step, least-to-most intrusive prompt sequence. A progressive ratio baseline was conducted next to verify that participants would not emit the target response under extinction conditions. Next, daily MSWO preference assessments commenced, using the same items evaluated in the PC preference assessment. Concurrently, daily progressive ratio assessments were conducted, with one item from the preference assessments included per session in random order. The progressive ratio reinforcer assessment continued until stability was achieved based on the cumulative number of responses for each item. Data were examined in times of break points obtained for each item. The PC preference assessment revealed a hierarchy of reinforcers for all participants. Overall, the results of the progressive ratio reinforcer assessment were variable across participants. For three participants, the PC did a good job of identifying not only the most effective reinforcer, but also the efficacy of reinforcers across the continuum of reinforcers. For these three participants, the daily MSWO rankings generally corresponded well with average break points. For two participants, one preference assessment method corresponded to better average break points. One of the two participants showed good correspondence between the rankings of the MSWO assessment and average break points, while the other participant's results on the PC assessment corresponded better to average break points. The final two participants demonstrated low break points for all stimuli. The reinforcer identified as most preferred by both the PC and the MSWO preference assessments produced the highest average break points for both of these two participants.

Kelly et al. (2014) examined the effectiveness of seven types of social reinforcement, ranked through two types of preference assessment, and evaluated through a reinforcer assessment. Pre-experiment assessments included a social stimuli questionnaire, the collection of video samples for each participant across activities to determine frequently delivered forms of attention, and the manipulation of motivating operations to determine whether the participants would mand for any forms of attention if a therapist withheld attention. Two forms of preference assessment were evaluated: a pictorial PC preference assessment and a single stimulus (SS) preference assessment. After establishing reliable rankings of reinforcers from both types of preference assessment, a mand reinforcer assessment was conducted to assess the predictive validity of reinforcers deemed high- and low-preference from the more reliable assessment. The reinforcing effects of high-preference and low-preference stimuli were evaluated in both a reversal design and a concurrent operant arrangement with a control condition. For all participants, only the PC preference assessment provided reliable results, so the high-preference and low-preference items from that assessment were used in the mand reinforcer assessment. For four participants, both the HP and LP items functioned as reinforcers, but the HP items were more effective. For the final participant, the high-preference and low-preference items both functioned as reinforcers and were equally effective. Researchers concluded that the PC preference assessment was a reliable method of identifying preferred social stimuli, and had high predictive validity regarding the reinforcing efficacy of social stimuli.

Training Others to Assess Preference

Two studies examined protocols for training practitioners to assess the preferences of people with developmental disabilities (Graff & Karsten, 2012; Weldy, Rapp, & Capocasa, 2014). Graff and Karsten (2012) assessed whether an antecedent-only self-instruction package

regarding the implementation of preference assessments could produce accurate responding in a group of staff members at a school for students with ASD and related disorders. First, teachers received written instructions alone, comprised of information drawn from the Methods sections of previously published papers on preference assessments. Teachers were given 30 min to read the instructions and then were permitted to bring the written instructions into the testing environment. They were then engaged in a simulated preference assessment with a researcher playing the part of the student. In a second, enhanced phase, written instructions in which the technical jargon had been removed, and diagrams and data sheets were added, were provided to the teachers. They subsequently engaged in additional simulated preference assessments. A third phase included the written instructions from baseline with the data sheet from the enhanced condition. For the group of teachers who received written instructions, written instructions plus data sheet, and then enhanced instructions, none of the six participants achieved mastery in the first two phases, and five out of six of the participants achieved mastery in the final phase. In general, the addition of a data sheet to the original written instructions did not improve accuracy, but the introduction of the enhanced instructions produced highly accurate responding. Another finding was that, when provided with written instructions alone, teachers were more accurate in providing instructions on the MSWO assessment than on the PC assessment, but were more accurate with the placement of stimuli for the PC assessment.

Weldy et al. (2014) trained behavioral staff members to conduct two brief preference assessments using 30 min videos that contained instructions and modeling. Participants were assigned to one of two groups. Group 1 was trained to implement a brief MSWO assessment first and a free operant (FO) assessment second. Group 2 was trained to implement a FO assessment first and a brief MSWO second. Four group training sessions and two individual booster sessions

were conducted over four weeks. Video training was comprised of a PowerPoint© presentation with embedded video models of one to two steps of each type of preference assessment presented at a time. Either the same day as the presentation or the day after the presentation, each participant was engaged in an in vivo preference assessment with a designated client. All participants achieved mastery criterion on both assessments. Seven participants achieved mastery criterion for both types of preference assessments after the first viewing of the training video. Two participants required a second viewing of the MSWO training video. Researchers concluded that staff members can learn to implement preference assessments using videos in a group setting.

The Effects of Preference on Behavior

One study assessed the effect of using preference to guide an intervention for a specific challenging behavior (Lanovaz, Rapp, & Ferguson, 2012). In this study, researchers examined whether noncontingent access to music would have differential effects on vocal stereotypy when PC preference assessment with each participant using five different songs. Songs were ranked, with the most frequently selected song being designated as the high-preference song, and the least frequently selected song being designated as the low-preference song. Next, participants were observed during comparison conditions in which either their low-preference or high-preference song was played continuously. Participants were also observed in a condition in which toys that did not produce auditory stimulation were present but no music was played. For three participants, noncontingent access to high-preference music produced lower levels of vocal stereotypy than either the no-interaction condition or the low-preference condition. For the fourth participant, noncontingent access to high preference music produced higher levels of vocal stereotypy. Researchers concluded that it is important to assess for preference when contemplating introducing a noncontingent music intervention to individuals who engage in vocal stereotypy.

Assessment of Choice

One study assessed the preference of individuals with developmental disabilities for making choices or not making choices (Sellerset al., 2013). In this study, researchers assessed response allocation among a choice, no-choice, and control condition initially. First, researchers conducted a PC preference assessment with each participant and ranked items as either high-preferred or medium-preferre. Then, participants approached a table on which three identical tasks with identical materials (differing only by color) were placed in front of varied high-preferred items (choice condition), the highest preferred item from the assessment (no-choice condition) and nothing (control condition). Based on the response allocation for each participant, further assessments of choice, including using varied medium-preferred items and identical high-preferred items, were conducted. Results indicated that two participants most frequently completed the task associated with the varied high-preferred items (choice condition), than completed the task associated with the single high-preferred item (no-choice condition) or control condition. The remaining two participants more frequently completed the task associated with the no-choice condition than the choice or control conditions. Researchers concluded that, for two participants, preference for choice or no-choice conditions was driven by the alternative that was associated with the higher-preference items. For a third participant, choice was preferred, unless high-preferred items were associated with the no-choice condition. The final participant preferred the choice condition exclusively, regardless of whether he was accessing high-preferred or medium-preferred items.

The Effect of Preference Assessment Type on Challenging Behavior.

One study examined how different preference assessment types affect challenging behavior in participants with developmental disabilities (Kang et al., 2011). Researchers measured rates of challenging behavior maintained by different reinforcers across three different

preference assessment formats (i.e., PC, MSWO, and FO). First, researchers conducted a functional analysis for each participant. Data suggested that each of the seven participants exhibited challenging behavior maintained by tangibles, escape from demands, or attention. Next, each preference assessment format was administered five times in a random order, for a total of 15 assessments for each participant. The rates of challenging behavior per participant during each type of preference assessment were graphed. Results indicated that for participants who exhibited challenging behavior maintained by tangibles, more challenging behavior occurred during PC and MSWO assessments, with very little challenging behavior occurring during FO assessments. For participants who demonstrated challenging behavior maintained by attention, high rates of challenging behavior occurred during the FO assessment, with lower rates in the PC and MSWO assessments. Finally, for the participant who demonstrated challenging behavior maintained by escape from demands, near-zero levels of challenging behavior were exhibited in all assessment types. Researchers concluded that challenging behavior was more probable in assessments that involved the removal of items (PC, MSWO) for participants whose challenging behavior was maintained by access to tangibles, or reduced interaction (FO) for participants whose challenging behavior was maintained by attention.

Discussion

Overall Effectiveness

Thirteen of the twenty studies examining preference assessments reported positive results (65%; Clay et al., 2013; Graff & Karsten, 2012; Graff & Larsen, 2011; Hall et al., 2014; Kelly et al., 2014; Kenzer et al., 2011; Kenzer & Bishop, 2011; Keyl-Austin et al., 2012; Lanovaz et al., 2012; Mangum et al., 2012; Russo et al., 2014; Tullis et al., 2012; Weldy et al., 2014;). The seven remaining studies examining preference assessments reported mixed results (35%; Call et

al., 2012; Call et al., 2013; Davies et al., 2013; Karsten et al., 2011; Nuerenberger et al., 2012; Snyder et al., 2012; Verschuur et al., 2011). As in Cannella et al. (2004) and Tullis et al. (2011), the majority of studies reviewed reported positive results, supporting the use of preference assessment methodology in determining effective reinforcers for individuals with developmental disabilities. Similar to Cannella et al. (2004), a number of studies reported mixed results. Possible reasons for mixed results given by researchers included difficulties in discriminating between many stimuli in the MSWO preference assessment (Davies et al., 2013), a lack of necessary prerequisite skills, such as matching (Snyder et al., 2012), and a possible shift in preference from the time of administration of the indirect preference assessment to the time of administration of the direct preference assessment (Verschuur et al., 2011). Four studies (Call et al., 2012; Call et al., 2013; Karsten et al., 2011; Nuerenberger et al., 2012) did not provide hypotheses for their mixed results.

One of the five studies examining the provision of choice reported positive results (20%; Philips & Mudford, 2011). The four remaining studies examining the provision of choice reported mixed results (Elliott & Dillenberger, 2014; Lough et al., 2012; Rispoli et al., 2013; Sellers et al., 2013). The percentage of choice studies reporting positive results (20%) in this review is low compared to the percentages of choice studies reporting positive results in previous reviews (Cannella et al., 2004, 80%; Lancioni et al., 1996, 83%, Tullis et al., 2011, 77%). Possible reasons for mixed results given by researchers included a potential inability of participants to express choices (Elliott & Dillenberger, 2014), a lack of prerequisite skills (Lough et al., 2012), and perception that the task presented was not meaningful (Lough et al., 2012). Two studies (Rispoli et al., 2013; Sellers et al., 2013) did not provide hypotheses for their mixed results. Previous research has indicated that the provision of choice is an evidence-based practice

for individuals with severe to profound developmental disabilities (Tullis et al., 2011). The mostly mixed findings among these studies may indicate that choice interventions are becoming more complex, such that the mechanisms of choice making are being examined in more sophisticated ways. Because a wealth of research has demonstrated that choice-making opportunities can reduce challenging behavior and improve quality of life, it appears that recent research has shifted to focus on participant preferences for different choice arrangements and how to make choice-making opportunities optimally effective. When given varied choice-making opportunities, participants will respond according to their preferences, which means that not all participants will respond favorably to all variations. Such variability in responding, in addition to the hypotheses presented by researchers, may explain the mostly mixed results demonstrated by the reviewed choice interventions.

Response of Studies to Previous Literature Review Suggestions

The authors of previous literature reviews each presented suggestions for future research regarding choice interventions and preference assessment studies. While some of the suggestions have yet to be realized in the literature, a number of the studies in the current review have addressed issues identified by previous reviews. Lancioni et al. (1996) suggested that future research examine ways that choice can be incorporated into the lives of people with developmental disabilities in addition to choices among food, drinks, and leisure activities. The choice interventions in this review introduced choices to participants that included academic tasks (Rispoli et al., 2013), task materials (Lough et al., 2012; Rispoli et al., 2013), task locations (Rispoli et al., 2013), activities (Philips & Mudford, 2011), and between choice-making and no-choice conditions (Sellers et al., 2013) Although the number of choice interventions was limited, each study presented choice opportunities outside the realm of food, drink, and leisure choices

examined in the studies reviewed by Lancioni et al. (1996). Furthermore, the range of types of stimuli assessed in preference assessments was broad, and included toys (Elliott & Dillenberger, 2014; Graff & Karsten, 2012; Karsten et al., 2011; 2012; Kenzer & Bishop, 2011; Snyder et al., 2012; Tullis et al., 2011), edibles (Call et al., 2012; Davies et al., 2013; Graff & Larsen, 2011; Graff & Karsten, 2012; Keyl-Austin, 2012; Russo et al., 2014; Sellers et al., 2013; Verschuur et al., 2011; Weldy et al., 2014), activities (Call et al., 2012; Kenzer & Bishop, 2011; Lough et al., 2012; Philips & Mudford, 2011; Rispoli et al., 2013; Russo et al., 2014; Verschuur et al., 2011), social interactions (Call et al., 2013; Clay et al., 2013; Kenzer & Bishop, 2011; Kelly et al., 2014; Nuerenbeerger et al., 2012), stimulation (Mangum et al., 2012; Verschuur et al., 2011), occupations (Hall et al., 2014), and songs (Lanovaz et al., 2012). In general, it appears that both the choice and preference assessment literature has expanded considerably in the years since Lancioni et al. (1996) published their review, and now includes a wide variety of choices for individuals with developmental disabilities, as well as a wide variety of stimuli from which to choose.

Cannella et al. (2004) suggested that choice interventions and studies investigating preference assessments start taking place in the home and community of participants, rather than being solely clinic and school based. Although studies included in this review did take place in clinics (Call et al., 2012; Call et al., 2013; Clay et al., 2013; Davies et al., 2013; Mangum et al., 2012; Weldy et al., 2014) and schools (Clay et al., 2013; Graff & Larsen, 2011; Graff & Karsten, 2012; Kang et al., 2011; Karsten et al., 2011; Kelly et al., 2014; Keyl-Austin, 2012; Rispoli et al., 2013; Sellers et al., 2013; Snyder et al., 2012; Tullis et al., 2012; Verschuur et al., 2011), settings also included the home (Elliott & Dillenberger et al., 2012; Karsten et al., 2011; Kenzer et al., 2013; Lanovaz et al., 2012; Nuerenberger et al., 2012; Philips & Mudford, 2011; Russo et al.,

2014; Verschuur et al., 2011) and the community (Hall et al., 2014; Lanovaz et al., 2012). It appears that an effort has been made to incorporate choice and preference into home- and community-based interventions to a greater extent than in past years. Although more difficult to control conditions, home- and community-based interventions demonstrate the feasibility of incorporating individual choice and preference into daily life. Moving successful interventions to the home and community would be a simple way to replicate and extend previous research and provide families and practitioners with evidence that affording individuals with developmental disabilities with opportunities to choose and to exercise preference may be an effective way to improve quality of life.

Tullis et al. (2011) suggested that future assessments of preference should investigate the durability of reinforcers selected in preference assessments. One study (Call et al., 2012) in this review investigated whether reinforcers identified through PC or MSWO preference assessments were more effective when presented according to a progressive ratio schedule of reinforcement. Interventionists conducted a single PC preference assessment to determine the most-preferred reinforcer. Then, they conducted daily MSWO preference assessment to identify the most-preferred reinforcer on a specific day. Using either the item identified from the PC preference assessment or the item identified from the MSWO preference assessment, interventionists engaged participants in a task for which the schedule requirement doubled after each delivery of reinforcement (e.g., FR1, FR2, FR4, FR8, etc.). The correspondence between break points and preference, as assessed by both formats, was assessed. Results indicated that both preference assessment formats did equally well at predicting reinforcer efficacy, but that the PC format was able to identify the more durable reinforcer more consistently. The progressive ratio schedule of reinforcement presents an efficient way to investigate how durable the reinforcing efficacy of an

item is under increasing response effort. Hopefully, future research will continue to utilize this schedule of reinforcement to examine how different types of preference assessments can identify more or less effective reinforcers. Also important will be studies that examine the durability of reinforcers over time, and the types of preference assessments that identify items that best withstand satiation.

Tullis et al. (2011) also suggested that future research investigate how to make less-preferred stimuli more effective. Although no studies modified conditions surrounding low-preferred stimuli in order to make them more reinforcing, three studies (Graff & Larsen, 2011; Keyl-Austin et al., 2012; Mangum et al., 2012) investigated reinforcer effectiveness of high-, moderate-, and low-preference stimuli, or manipulated preference assessment presentations to include low-preference stimuli. For example, Mangum et al. (2012) conducted two separate preference assessments for participants before engaging them in a reinforcer assessment. The first preference assessment was titled the Full-Array preference assessment, and consisted of all 11 items identified through a structured interview. The second preference assessment was termed the Partial-Array preference assessment, and consisted of the five items that ranked lowest on the Full-Array preference assessment. Results indicated that for some participants, the stimulus chosen in the Partial-Array preference assessment functioned effectively as a reinforcer, although the high-preference stimulus chosen from the Full-Array was more effective in general. This study demonstrated that even low-preference stimuli can function as reinforcers when presented in the absence of higher-preference stimuli. Parents and practitioners could therefore use a wider array of reinforcers if they removed access to highly preferred stimuli temporarily. Future research could continue to investigate how both presentation of stimuli in preference assessments

and manipulation of context, including examinations of motivating operations, could improve the effectiveness of low-preference reinforcers.

The included studies did not address a number of suggestions made by previous literature reviews. Cannella et al. (2004) suggested that future research should investigate characteristics of participants that make them more or less likely to respond to choice interventions. In addition, Cannella et al. (2004) suggested that interventionists conduct further examination of the reasons for mixed results among choice interventions. Although Elliott and Dillenberger (2014) and Lough et al. (2012) did provide possible explanations for the mixed results demonstrated in their experiments (lack of prerequisite skills, or the perception that the task was not meaningful, respectively), no included studies conducted post hoc analyses to determine why the interventions were not successful with all participants. Tullis et al., (2011) suggested that future research determine whether there was a set of necessary prerequisite skills for successfully participating in various types of preference assessments. Although some studies required prerequisite skills for inclusion (e.g., the ability to make choices, Clay et al., 2013; the ability to scan items in an array; Kang et al., 2011), no studies systematically examined how prerequisite skills affected performance.

TRENDS IN THE CURRENT LITERATURE

Assessing Preference for Social Stimuli

In the Tullis et al. (2011) review, one study (Smaby, MacDonald, Ahearn, & Dube, 2007) examined a procedure for identifying preferred social consequences. In the current review, the number of studies examining preferences for a variety of social interactions has increased, with four studies (Call et al., 2013; Clay et al., 2013; Kelly et al., 2014; Nuerenberger et al., 2012) investigating how preference for social interaction can be determined and utilized effectively.

For example, Nuerenberger et al. (2012) conducted a MSWO preference assessment for a variety of forms of social interaction, with subsequent reinforcer assessments to determine if preferred forms functioned as reinforcers. First, informal interviews with teachers and caregivers identified multiple forms of social interaction that each participant seemed to enjoy. Photos of interventionists engaging in each identified form of social interaction were taken with each participant and printed in color. Next, a MSWO preference assessment using the photos was conducted with each participant. Selection of a photo resulted in 15 s of the type of social interaction depicted. After a ranking of preference for social interaction was determined, a reinforcer assessment was conducted using a sorting task previously mastered by each participant. Contingent upon a correct sorting response, 15 s of one kind of social interaction was provided per session. After delivery concluded, the sorting task was presented again. The reinforcer assessment continued until five minutes elapsed or challenging behavior occurred. Results indicated that the preference assessment was predictive of sorting responses in the reinforcer assessment for two participants, with unclear results for the third participant. Interventionists concluded that most of the forms of social interaction functioned as reinforcers for participants, and that even low-preference forms functioned as reinforcers. It appears that interest in both the use of social interaction as reinforcement, and the effective assessment of individual preference for social interactions is growing in recent years. Social interactions represent a no-cost, easily-delivered, and inexhaustible form of reinforcement. As such, establishing procedures for determining preference hierarchies for social interaction will likely be a topic of interest for many practitioners. Now that studies have established methods for assessing preference for social interactions that effectively function as reinforcers, one of the

next steps will be to examine whether increased response requirements diminish reinforcer potency, as compared to other forms of reinforcement.

Enhancement of Training Procedures

Although the training of others to assess for preferences or to establish choice-making opportunities was previously examined in studies reviewed by Tullis et al. (2011) (e.g., Machaliecek et al., 2009; Roscoe & Fisher, 2008;), more recent studies have focused on making training more effective and using fewer resources. For example, Graff & Karsten (2012) developed an antecedent-only training model that did not require in-person training with experts (see above). Results indicated that enhanced written instructions with reduced jargon and more visuals were sufficient to train all participants to mastery criterion. Weldy et al. (2014) developed a training model that relied on video examples to enhance training lessons (see above). Their results indicated that a video model was effective at training practitioners to conduct both MSWO and FO preference assessments. These studies address one of the major criticisms of preference assessment training models addressed by Tullis et al. (2011), who suggested that many practitioners would find them substantially time- and resource-intensive. Given that assessing preferences and providing opportunities to make choices has been established as an evidence-based practice for improving quality of life (Tullis et al., 2011), methods for making this process more efficient are important. As methodological enhancements for assessing preference are refined, hopefully practitioner models will continue to evolve into their most efficient and effective forms.

Indirect vs. Direct Preference Assessment Methods

Two studies in this review (Russo et al., 2014; Verschuur et al., 2011) compared indirect preference assessment methods, such as structured parent interviews, to established, direct

preference assessment methods. Verschuur et al. (2011) conducted a structured interview with the parents or teachers of participants that consisted of 10 questions that aimed to identify preferences for specific food, drinks, sensory objects, toys, and activities. Informants were asked to make rank orderings of their child/student's top six food and top six sensory/play items. Next, an MSWO preference assessment was conducted with the identified six items for each participant. Spearman rank correlations between the results of the structured interview protocol and the MSWO were calculated. Results indicated that the mean Spearman rank order correlation coefficient was large for four of the seven children, medium for one child, and small for one child. Interventionists concluded that parents were able to provide an accurate rank ordering of their children's preferred food and toys using a structured interview. Russo et al. (2014) compared the results of an open-ended interview for assessing preference and direct preference assessment methods for all participants, as well as a reinforcer assessment for two participants. First, interventionists conducted an open-ended interview with the parents of participants that asked them to rank five edible and five activity stimuli in order from most-to-least preferred. Next, an MSWO preference assessment was conducted for the edible and activity stimuli of each participant separately. For two participants, a reinforcer assessment was conducted using a simple sorting task. Spearman's rank order correlation coefficients were calculated for the results of the parent interview and the MSWO preference assessment. In general, there was little correspondence between the two assessment methods, with the exception of one participant, for whom strong correlations were demonstrated. The reinforcer assessment results indicated that both the highest-ranked item identified by parents and by researchers functioned as reinforcers for both assessed participants, but that for one participant, the item ranked highest by parents was more effective. Although the results of each assessment were not comparable per participant,

both formats identified effective reinforcers. Interventionists concluded that results tentatively supported the use of open-ended parent interviews for assessing preferences. The drive to assess the validity of parent interview preference assessments perhaps indicates that some practitioners may consider traditional preference assessment methods too time-consuming. Mixed results regarding the correlation between direct assessment methods indicate that only tentative conclusions can be drawn about the congruent validity of indirect methods. Future research should continue to refine structured interview procedures for assessing preferences for the purpose of establishing truly valid indirect methods. For practitioners who consider established preference assessment methodology to be too cumbersome, a valid structured interview could provide an accessible means of assessing preference, and ensure that individual preferences are being honored.

FUTURE RESEARCH

The reviewed studies represent expansion and refinement of an already robust body of literature regarding choice and preference assessment methodologies. New research questions emerge from both the impetus and findings of several studies. One priority of a number of studies was to assess preference in an efficient manner that did not consume substantial time. Russo et al. (2014) and Verchuur et al. (2011) examined the validity of indirect methods of assessing preference that consisted of brief, open-ended parent interviews. Tullis et al. (2012) compared the accuracy of the MSWO preference assessment format when comprised of different numbers of array presentations, with the finding that even one array presentation produced a stimulus that functioned as a reinforcer. Reducing the time commitment required to conduct direct preference assessment formats, or replacing them with valid indirect formats, can increase social validity, and make it more likely that teachers or caregivers will assess for individual

preference. Future research should continue to examine ways to assess preference and provide choice that do not take substantial time away from instruction or leisure. An ideal form of assessment would be a procedure that could be conducted concurrently with instruction.

Elliott and Dillenberger (2014) demonstrated that response-contingent choice improved response accuracy and reduced time to return to the table from a break for one participant. For that participant, clear differentiation was demonstrated between a condition in which he was given the opportunity to choose a reinforcer contingent upon a response, and a condition in which a highly preferred item was provided without a choice. The authors suggested that future research continue to investigate the impact of response-contingent choice and choice timing (i.e., when a choice is presented in relation to antecedents and behaviors) experimentally. Response-contingent choice presents a possible avenue for assessing preference at the same as conducting instruction. Providing a response-contingent choice could allow individuals to exercise their preferences through choice. Recording choices made by the individual across trials would provide a ranking of preference for the reinforcers presented similar to the ranking produced from a MSWO preference assessment. In both formats, a predetermined number of reinforcers is presented repeatedly without the removal of any items. In addition to being efficient, response-contingent choice could potentially improve motivation, as demonstrated by decreased time to return to work after a break. Future research could investigate whether response-contingent choice increases other variables of motivation and/or which types of individuals respond well to response-contingent choice, as compared to the provision of highly preferred reinforcers without choice.

Typical assessment of reinforcer efficacy involves either validation that more responses are allocated when reinforcers are provided contingent upon responses than when no

reinforcement is provided (e.g., Clay et al., 2013; Graff & Larsen, 2011) or an examination of the rate of response under increasing response requirements (e.g., Call et al., 2013). The former assessment determines whether a stimulus truly functions as a reinforcer, while the latter determines the durability of a reinforcer when response requirements increase. One study (Elliott & Dillenberger, 2014) measured an additional dependent variable that captured motivation to respond for access to a stimulus: time elapsed from the end of a break and the direction to return to the table to the participant actually sitting at the table. For the participant that demonstrated clear differentiation between choice and no-choice conditions, time to return to table was longer when choice was not provided. Temporal dependent variables provide an attractive option for measuring participant motivation. In addition to latency to return to table, other possible dependent variables include latency to the task initiation of a response or duration of total task. Future research could examine motivation for reinforcers selected by preference assessments through the dependent variables of latency to task initiation or total task duration.

Table 1
Study Characteristics

Category	Studies	*n*	Age(s)	Diagnoses	Stimuli	Choice Format	Dependent Variable(s)	Design	Findings
Effectiveness of Preference Assessments	Call, Shillingsburg, Bowen, Reavis, & Findley (2013)	6	2-11	ASD	Social interaction condition or being alone	Pairs	Percentage of time spent in side of room	Reversal	Mixed
	Clay, Samaha, Bloom, Bogoev, & Boyle (2013)	5	6-17	ASD, CP, ADHD, PDD-NOS, BP, ID, L-GD	Various forms of social interactions	PC	Percentage of selections	Reversal	Positive
	Graff & Larsen (2011)	5	12-15	ASD	Various edibles	PC	Approach responding; frequency of responses	Multi-element and reversal	Positive
	Hall, Morgan, & Salzberg, (2014)	4	19-20	DS, TBI, ASD, VI	Various occupations	Single	Percentage correct responding; Percentage of productivity	Alternating treatments	Positive
	Karsten, Carr, & Lepper (2011)	20	3-11	ASD	Various toys	MSWO, concurrent operant assessment, FO	Rank of stimulus, percentage of interval actively engaged	Concurrent operant assessment	Mixed
	Kenzer, Bishop, Wilke, & Tarbox (2013)	3	3-4	ASD	Not stated	PC	Percentage of selection; responses per minute	ABAB reversal, with embedded multi-element	Positive
	Keyl-Austin, Samaha, Bloom, & Boyle (2012)	1	4	ASD	Various edibles, leisure items	PC, MSWO	Mean number of responses; proportion of IRTs	PA, brief comparison session	Positive
	Mangum, Frederick, Pabico, & Roane (2012)	3	6-13	ASD, VI	Various toys, sources of auditory stimulation, visual stimulation, and	PC	Percentage of session spent in area; percentage of trials chosen	Concurrent operant assessment	Positive

Table 1 (continued)

					tactile stimulation				
	Nuernberger, Smith, Czapar, & Klatt (2012)	3	4	ASD	Various types of social interactions	MSWO	Percentage of trials chosen; percentage of intervals in which the behavior occurred	Multi-element within reversal	Mixed
	Russo, Tincani, & Axelrod (2014)	6	3-4	ASD	Various activities and edibles	Parent survey, MSWO	Correct responses	Alternating treatment within an ABAB reversal design	Positive
	Tullis, Canella-Malone, & Fleming (2012)	1	9	PDD-NOS	Various toys	MSWO	Number of responses	ABAB reversal	Positive
Assessing Preference	Davies, Chand, Yu, Martin, & Martin (2013)	9	26-52	ID	Various edibles	PC, MSWO	Percent of trials selected	PA only	Mixed
	Kenzer & Bishop (2011)	31	2-9	ASD	Toys, activities, edibles, social reinforcement	PC	Percentage of trials in which an item was selected	PA only	Positive
	Snyder, Higbee, & Dayton (2012)	6	3-5	ASD	Various toys	PC with and without video recording	Percentage of trials selected.	PA only	Mixed
	Verschuur et al. (2011)	7	4-10	ASD	Various edibles, activities, and sensory items	Structured parent interview, MSWO	Percentage of trials selected	PA only	Mixed
Effects of Choice on Behavior	Elliott & Dillenburger (2014)	3	4-5	ASD	Various toys	PC or no choice	Correct responding, non-responding, time to return to table	Alternating treatment	Mixed
	Lough, Rice, & Lough (2012)	26	8-15	ASD	Pictures to color	Choice from three or no choice	Amount of page colored, number of colors used, and time on task	Counter-balanced design with participants yoked into dyads	Mixed
	Phillips & Mudford (2011)	1	24	ID	Various activities	Choice from seven or no choice	Grabbing, task engagement, staff attention	ABCBCDBDABCD Reversal	Positive
	Rispoli et al. (2013)	4	5-11	Mixed ASD	Various activities, locations in which to complete an activity, activity materials	Pairs or Fours for choice conditions, MSWO for determining possible choices	Rate per minute of challenging behavior	Alternating treatment design imbedded within the 'B' phase of an ABAB design	
Preference Assessment	Call, Trosclair-	7	5-18	ASD, sensory	Edible items,	PC, MSWO	Breakpoints (the	Alternating treatment;	Mixed

Table 1 (continued)

Comparisons	Lasserre, Findley, Reavis, & Shillingsburg (2012)			disintegration disorder	leisure activities		highest schedule requirement completed during the PR reinforcer assessment for the item that day)	alternating treatment within reversal	
	Kelly, Roscoe, Hanley, & Schlichenmeyer (2014)	5	9-19	ASD	Various forms of social reinforcement	PC, single stimulus	Percentage selection, Number of mands, responses per minute	Reversal ABABACBC	Positive
Training Others to Assess Preferences	Graff & Karsten (2012)	11	Not stated	None; Teachers at a school for children with ASD and DD	Various toys or edibles	PC; MSWO	Percentage trials implemented correctly	Multiple baseline across assessment types	Positive
	Weldy, Rapp, & Capocasa (2014)	9	Not reported	None; participants were employees at a behavioral center for children with ASD	Various edibles and leisure items	FO, MSWO	Percent correct responses on task list	Nonconcurrent multiple probe across preference assessments	Positive
Effects of Preference on Behavior	Lanovaz, Rapp, & Ferguson (2012)	4	4-9	ASD	Various songs	PC	Percentage of trials selected, percentage of time engaged in vocal stereotypy	PA, brief comparison condition	Positive
Assessment of Choice	Sellers, Bloom, Samaha, Dayton, Lambert, & Keyl-Austin (2013)	4	4-5	ASD or PDD-NOS	Various edibles	PC	Responses per minute	Concurrent operant assessment	Mixed
The Effect of Assessment Type on Challenging Behavior	Kang, O'Reilly, Fragale, Aguilar, Rispoli, & Lang (2011)	7	4-8	ASD, DD	Not stated	PC; MSWO, FO	Responses per minute of problem behavior	Multi-element	N/A

Table 1: Studies listed by category, number of participants, ages, diagnoses, stimuli, choice format, dependent variables, design, and findings. ADHD: attention deficit hyperactivity disorder;; ASD: autism spectrum disorder; BP: bipolar disorder; CP: cerebral palsy; DD: developmental disability; DS: Down syndrome; FO: free operant; ID: intellectual disability; ; L-G: Lennox-Gastaut Syndrome; MSWO: multiple stimuli without replacement; PA: preference assessment; PDD-NOS pervasive developmental disorder, not otherwise specified; TBI: Traumatic brain injury; VI: visual impairment

Chapter 3: Method

Inclusion Criteria

In order to be included as participants in this experiment, individuals were required to have a diagnosis of a recognized developmental disability. In addition, participants needed to demonstrate preference (defined as the propensity to engage with an item for more than 30 seconds, or until consumed, if edible) for at least five different edible items, toys, or activities of the same stimulus class, of which three were utilized as reinforcers. Finally, participants needed to demonstrate the ability to perform reliably (i.e., five, concurrent, correct responses with stable Latencies to Initiation and Total Task Durations) on a simple task.

Participant Characteristics

There were a total of four participants in this experiment. A description of their characteristics follows.

James

James was a Caucasian male, aged 10, with a diagnosis of ASD. James attended a local school for children with ASD and developmental disabilities on a full-day schedule.

The Vineland-II was administered to James through interview with the Director of Operations at his school, who also typically conducted his one-to-one ABA therapy. James' Domain Standard Scores on the Vineland-II are as follows: Communication: 60 (0.4 percentile), Daily Living Skills: 60 (0.4 percentile), and Socialization: 64 (1 percentile). James' motor skills raw scores did not correspond to a Domain Standard Score in the Vineland-II scoring manual. See Table 2 for more detailed Vineland-II scoring information.

James periodically engaged in the following challenging behaviors: elopement (defined as moving his body away from the current task without permission), motor stereotypy (defined as intently focusing with his eyes on the motor movements of his hands), and vocal stereotypy (defined as emitting nonfunctional and/or repetitive vocalizations). During the experiment,

challenging behaviors were blocked and redirected (elopement and motor stereotypy), or ignored (vocal stereotypy).

Aimee

Aimee was a Caucasian female, aged 12, with a diagnosis of rare chromosomal deletion (on chromosome three). The behavioral phenotype of Aimee's diagnosis was similar to ASD. Aimee attended the same school as James, also on a full-day schedule.

The Vineland-II was administered to Aimee through interview with the Director of Academics at her school, who also typically provided Aimee's ABA services. Aimee's Domain Standard Scores on the Vineland-II are as follows: Communication: 54 (<1 percentile), Daily Living Skills: 47 (<1 percentile), Socialization: 61 (<1 percentile). Aimee's Motor Skills raw scores did not correspond to a Domain Standard Score in the Vineland-II scoring manual. See Table 2 for more detailed Vineland-II scoring information.

Aimee periodically engaged in the following challenging behaviors: whining (defined as vocally emitting an audible cry of more than 1 s that did not include shedding tears), crying (defined as visibly shedding tears), and head banging (defined as attempting to forcibly make contact between the forehead and the table). During the experiment, challenging behaviors were ignored (whining and crying), or blocked and ignored (head banging).

Ethan

Ethan was a Caucasian male, aged five, with a diagnosis of ASD. He attended a local public school in an Early Learning Environment (ELE) classroom.

The Vineland-II was administered to Ethan through interview with his mother. His Domain Standard Scores on the Vineland-II are as follows: Communication: 97 (42 percentile),

Daily Living Skills: 46 (20 percentile), Socialization: 72 (3 percentile), and Motor Skills: 75 (5 percentile). See Table 2 for more detailed Vineland-II scoring information.

Ethan periodically engaged in the following challenging behaviors: elopement (defined as moving his body away from the current task without permission), noncompliance (defined as not responding to the S^D within ten seconds), and vocal protest (defined as audibly protesting in a volume that exceeded conversational tones). During the experiment, challenging behaviors were blocked and ignored (elopement), ignored (vocal protest), or addressed with escape extinction (noncompliance).

Connor

Connor was a Caucasian male, aged five, with a diagnosis of Beckwith-Weidemann syndrome and comorbid diagnosis of ASD. He attended a local public school in an Early Structured Learning Environment (ESLE) classroom on a half-day schedule. Connor and Ethan were identical twins and lived in the same house with their mother and father.

The Vineland-II was administered to Connor through interview with his mother. His Domain Standard Scores on the Vineland-II are as follows: Communication: 85 (16 percentile), Daily Living Skills: 64 (1 percentile), Socialization: 59 (0.3 percentile), and Motor Skills: 64 (1 percentile). See Table 2 for more detailed Vineland-II scoring information.

Connor periodically engaged in the following challenging behaviors: elopement (defined as moving his body away from the current task without permission), noncompliance (defined as not responding to the S^D within ten seconds), and crying (defined as visibly shedding tears). During the experiment, challenging behaviors were blocked and redirected (elopement), ignored (crying), and addressed with escape extinction (noncompliance).

Table 2
Vineland-II Scores

Participant	Domain	Subdomain	Raw Score	V-Scale Score	%tile Rank	Adaptive Level	Age Equivalent
	Communication	Receptive	13	8		low	<3.0
		Expressive	28	8	0.4%	low	<3.0
		Written	22	9		low	4:10
	Daily Living Skills	Personal	26	9		low	4:9
		Domestic	23	8	0.4%	low	6:4
		Community	12	8		low	<3:0
James	Socialization	Interpersonal Relationships	7	7		low	<3:0
		Play and Leisure Time	15	8	1%	low	<3:0
		Coping Skills	20	10		Moderately low	3:4
	Motor Skills	Gross	28	N/A	N/A	N/A	3:3
		Fine	45	N/A	N/A	N/A	>6:11
	Communication	Receptive	13	8		Low	<3:0
		Expressive	17	7		Low	<3:0
		Written	10	7	<1	Low	5:0
	Daily Living Skills	Personal	11	5		Low	<3:0
		Academic	6	5		Low	3:6
		School Community	12	8	<1	Low	3:3
Aimee	Socialization	Interpersonal Relationships	7	7		Low	<3:0
		Play and Leisure Time	11	7	<1	Low	<3:0
		Coping Skills	19	10		Moderately low	3:2
	Motor Skills	Fine	19	N/A	N/A	N/A	<3:0
		Gross	14	N/A	N/A	N/A	<3:0
	Communication	Receptive	34	15		Adequate	5:6
		Expressive	70	11	42 %tile	Moderately low	3:2
		Written	21	18		Moderately high	6:1
	Daily Living Skills	Personal	37	8		Low	2:9
		Domestic	8	1	<0.1 %tile	Low	2:11
Ethan		Community		11		Moderately low	3:5
	Socialization	Interpersonal Relationships	35	10	3 %tile	Moderately low	2:3
		Play and Leisure Time	13	7		Low	1:2
		Coping Skills	23	13		Adequate	3:11
	Motor Skills	Gross	63	10	5 %tile	Moderately low	2:9
		Fine	49	12		Moderately low	4:5
	Communication	Receptive	28	12	16 %tile	Moderately low	2:11
		Expressive	61	9		Low	2:10
		Written	19	17		Adequate	5:11
	Daily Living Skills	Personal	22	7		Low	1:11
		Domestic	5	11	1 %tile	Moderately low	2:2
		Community	12	10		Moderately low	3:2
	Socialization	Interpersonal Relationships	28	8		Low	1:5
Connor		Play and Leisure Time	12	7	0.3 %tile	Low	1:1
		Coping Skills	7	8		Low	1:9
	Motor Skills	Gross	57	9		Low	2:3
		Fine	34	9	1 %tile	Low	3:0

Table 2: Vineland-II expanded score information

SETTING

The experiment took place in two locations. Aimee and James engaged in the experiment at a local school for children with ASD and other developmental disabilities. For Aimee, the specific setting was a classroom within the school. Aimee was positioned at a desk with the experimenter seated next to her in a second desk. Approximately three other students, a teacher, and an individual holding the camera were in the room on any given day. Sessions took place between 12:00 pm and 1:00 pm on weekdays, 2-4 times per week.

For James, the specific setting was a partitioned area in the front room of the school. James was positioned at a single desk with the experimenter seated in a second chair at the same desk. No other students were present in the area at the time of the experiment. An individual holding the camera stood next to the desk and filmed James engaging with the experiment materials. Sessions took place between 12:00 pm and 1:00 pm on weekdays, 2-4 times per week.

Connor and Ethan engaged in the experiment at their home. For both participants, the specific setting was a guest room on the second floor of the house. Connor and Ethan, individually, were positioned in a chair at a small table, with the experimenter seated next to the participant in a second chair. No other individuals were present, and the camera was placed on a tripod. Sessions took place between 4:00 pm and 5:30pm on weekdays, 2-4 times per week.

MATERIALS

Reinforcers

A structured interview (RAISD; Fisher et al., 1996) was conducted with a relevant adult (a teacher for James and Aimee, and their mother for Connor and Ethan) to determine a list of at least five toy, edible, activity, or stimulation-related reinforcers of the same stimulus class per participant. After engaging each participant in PC preference assessment with the five identified stimuli, the top and bottom ranked items were dropped, leaving three reinforcers available for use in the experiment. This protocol was followed to mitigate both the risk of a highly preferred stimulus being chosen exclusively in all conditions and resulting in undifferentiated data, and/or a low-preferred stimulus evoking challenging behavior.

Task Materials

A simple task of a reliable duration was selected based on input from the relevant adult about the current skill set of each participant. For James, the task materials included a folder with two printed sight words and four laminated, circular images. Each piece was adorned with Velcro, such that James could affix the laminated image to the sight word with which it rhymed. For Aimee, Connor, and Ethan, the task materials included a folder with four solid shapes and four laminated shape pieces. Similar to the task for James, each shape was adorned with Velcro, such that participants could affix the shape piece to the image to which it matched. See Table 3 for participant information.

Table 3
Participant Characteristics

Participant	Age	Gender	Diagnosis	Setting	Reinforcers	Task Materials
James	10	Male	ASD*	School	Edibles (popcorn, orange Starburst©, powdered donut)	Laminated folder, small circular images of rhyming words
Aimee	12	Female	Unnamed chromosomal disorder (series of genes missing on 3rd chromosome)	School	Video clips (Dora the Explorer©, Bubble Guppies© "Line Up" song, Bubble Guppies© "Lunch" song)	Laminated folder, large shapes
Ethan	5	Male	ASD	Home	Video clips (school bus, subway, Connor riding an elevator)	Laminated folder, large shapes
Connor	5	Male	ASD Beckwith Weidemann Syndrome	Home	Video clips (Baby Einstein©, Pete the Cat©, Kinder Surprise© eggs)	Laminated folder, large shapes

Table 3: Participant Characteristics. ASD: autism spectrum disorder

DEPENDENT VARIABLES

Latency to Task Initiation

Latency to Task Initiation was defined as the time elapsed between the termination of the discriminative stimulus (S^D) and the initiation of the requested task, as

expressed in seconds. For each participant, initiation of the task was defined as contact between the participant's digits and the laminated pieces. Latency to Task Initiation was determined from video recording using a timing device, and graphed as a mean across responses within a single session.

Total Task Duration

Total Task Duration was defined as the time elapsed between the initiation of the task and the termination of the task, as expressed in seconds. Termination of the task was defined as the placement of the final laminated piece on the folder and the removal of the digits. This dependent variable was determined from video recording using a timing device, and graphed as a mean across responses within a single session.

Percentage Correct Responding

Percentage Correct Responding was defined as the number of observed correct responses, divided by the number of total possible correct responses (i.e., 20), multiplied by 100. This dependent variable was recorded as correct or incorrect per response in vivo, and was calculated once all responses were recorded for a single session. Percentage Correct Responding was verified via video recording.

No-Responses

No-Responses were defined as trials in which the Latency to Task Initiation exceeded 5 s. If the participant was in the process of reaching for the laminated piece at 5 s, a No-Response was not scored. This dependent variable was determined in vivo using a timing device, and verified from video recording using the same timing device.

PROCEDURE

Pre-Experiment

Simple Task Assessment

In order to establish that the simple task used in the experiment was a mastered skill, and therefore, that fluctuations in the dependent variables were due to motivating operations rather than a skill deficit, performance on a simple task was assessed prior to beginning data collection. A relevant adult identified a task that she was confident that the participant in question had mastered. Subsequently, performance on the task was assessed with each participant prior to initiating the experiment.

Reinforcement occurred on a FR1 schedule of reinforcement, using whichever stimulus was used to train the response initially (popcorn for James, and video clips for Aimee, Ethan, and Connor). The experimenter delivered the relevant S^D ("Do your work" for James, and "Match" for Aimee, Ethan, and Connor). Contingent upon a correct response, the selected reinforcer was delivered for 30 s, or until consumed (James). Following an incorrect response, the correct response was modeled, the laminated pieces were removed, and the S^D was delivered again. The assessment was concluded when the participant correctly responded five times consecutively. From film, Latency to Task Initiation and Total Task Duration were calculated using a timing device, and graphed in Microsoft Excel©. Visual analysis determined whether or not these two dependent variables were stable or variable. If five consecutive correct responses, with stable Latency to Task Initiation and Total Task Duration, were observed, the participant was

accepted into the study. See Figures 5-12 in Chapter 4 for data related to the Simple Task Assessment for each participant.

Preference Assessment.

A PC preference assessment was administered for each participant with the five stimuli identified from the RAISD. In a PC preference assessment, two items are presented at a time to the participant and the selection of the participant is recorded. Each possible combination of items is presented in a randomized sequence in which position bias is controlled for by systematically varying on which side (i.e., left or right) each item is placed (Fisher et al., 1992). The percentage of trials in which each stimulus was chosen was calculated, and a ranked order of reinforcers was established for each participant. See Figures 1-4 for a visual representation of each participant's selections. Once a ranked order was established, the top and bottom items were dropped, leaving ranked items 2-4 available for the experiment for each participant. For James, reinforcers included popcorn (one kernel per trial), orange Starbursts (approximately one-sixth of a single Starburst per trial), and powdered donuts (a piece measuring approximately 3cm by 4cm per trial). For Aimee, reinforcers included three, 30 s video clips from the children's television shows, Bubble Guppies© (two clips) and Dora the Explorer© (one clip). For Ethan, reinforcers included three, 30 s video clips of a school bus, his brother, Connor, riding in a hospital elevator, and a subway train. For Connor, reinforcers included three, 30 s video clips from the children's television show, Baby Einsteins©, a Youtube© video of an individual opening Kinder Surprise Eggs©, and a video reading of a Pete the Cat© book.

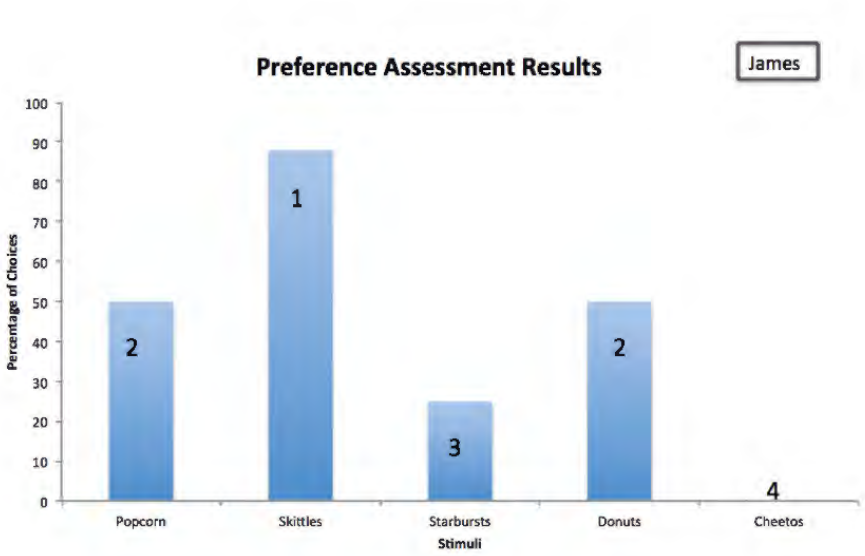

Figure 1: Preference Assessment Results for James

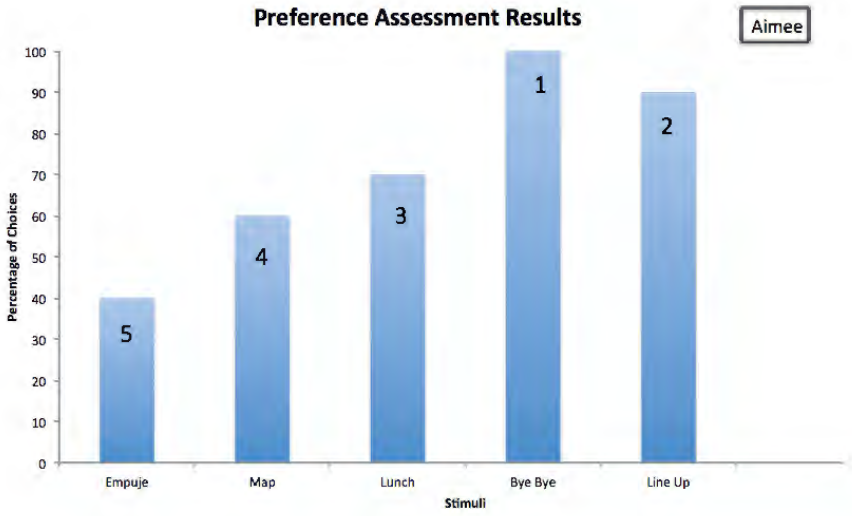

Figure 2: Preference Assessment Results for Aimee

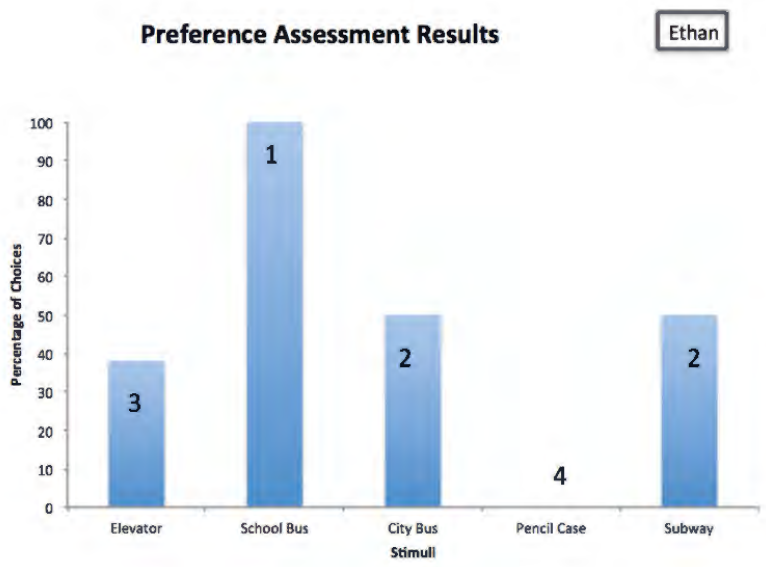

Figure 3: Preference Assessment Results for Ethan

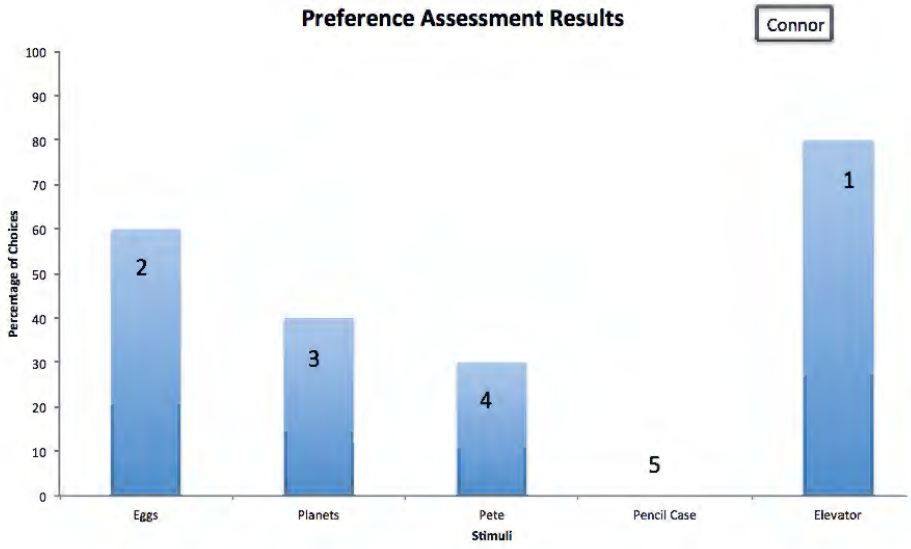

Figure 4: Preference Assessment Results for Connor

INDEPENDENT VARIABLES

Three individual conditions were randomized across sessions, with the added requirement that no more than two sessions utilized the same condition consecutively. Conditions included a Five Pre-Session Choices condition, a Five Within-Session Choices condition, and a One Pre-Session Choice condition. Duration of sessions for James ranged between 8 m and 12 m (consisting of 21 trials and the duration of consumption of the edible 26 times (Five Pre-Session Choices condition and Five Within-Session Choices condition) or 22 times (One Pre-Session Choices condition). Duration of sessions for Aimee ranged between 20 m and 40 m (consisting of 21 trials and 26 viewings of a 30 s video clip (Five Pre-Session Choices condition and Five Within-Session Choices condition) or 22 viewings of a 30 s video clip (One Pre-Session Choice condition)). Duration of sessions for Ethan and Connor ranged between 25 m and 35 m (consisting of 21 trials and 26 viewings of a 30 s video clip (Five Pre-Session Choices condition and Five Within-Session Choices condition) or 22 viewings of a 30 s video clip (One Pre-Session Choice condition)). Fluctuation in session duration was due to the slight variations in conditions and the occurrence of challenging behavior (see Participant Characteristics for operational definitions of each participants' challenging behaviors).

Five Pre-Session Choices Condition

On sessions designated as Five Pre-Session Choices conditions, a 5-trial MSW preference assessment was administered before the experiment. In the MSW preference assessment, all stimuli are initially presented in an array, in a randomly determined order. Upon selection of the first stimulus, that stimulus is presented for a predetermined period of time (e.g., 30 s, or until consumed, if edible) and then replaced in the array for the next presentation, such that all arrays contain identical items, randomly rotated for order of

presentation. The reinforcer selected for the highest percentage of trials was determined to be the Designated Reinforcer for that session. In the event that two items were chosen with the same frequency (i.e., twice), a coin was flipped, and the item selected by the coin toss was used for that session.

After selection of the Designated Reinforcer, the simple task was initiated. An initial trial occurred so that the participant could identify in which condition he/she was operating for the day, and data was not taken for said trial. Contingent upon a correct response, the Designated Reinforcer was presented for 30 s, or until consumed (James). Following the initial trial, data collection began. Trials consisted of the verbal delivery of the relevant S^D, followed by the participant's response, and the delivery of the reinforcer on a FR1 schedule of reinforcement. If the participant responded correctly to the S^D, the trial was scored as correct, the reinforcer was delivered for 30 s, or until consumed (James), and the next trial was subsequently initiated. If the participant responded incorrectly to the S^D, the correct response was immediately modeled with a relevant vocalization ("Oops. (Word) rhymes with (word)" for James, and "Oops. (Shape) matches with (shape) for Aimee, Ethan, and Connor). Next, the laminated pieces were removed, and the trial was initiated once again. Incorrect responses were scored on the data sheet in vivo. If a participant demonstrated five, consecutive, incorrect responses, it was concluded that satiation with the reinforcer had occurred, and that an abolishing operation was in effect. In that event, the session was terminated early due to low motivation. Latency to Task Initiation was calculated in vivo via a timing device, and, in the event that the participant was not actively reaching for the laminated pieces when 5 s

had elapsed, the experimenter stated, “Oops. Let’s try again” in a neutral voice, removed the laminated pieces, and initiated a new trial. If a participant demonstrated five, consecutive trials in which Latency to Initiation Task exceeded 5 s, it was also concluded that satiation with the reinforcer had occurred, and that an abolishing operation was in effect. In that event, the session was terminated due to low motivation. These termination criteria were established in order to determine an appropriate termination point in the event that an abolishing operation was in effect to the extent that responding would cease to occur. The experimenter did not prompt responses, other than to silently tap the folder approximately every ten seconds when a participant paused substantially during the task. Provided that the termination criteria expressed above were not met, sessions were terminated when twenty trials had been conducted (excluding the initial trial, for a total of 21 trials). For James, Aimee, and Ethan, eight, Five Pre-Session Choices conditions, eight, Five Within-Session Choices conditions, and six, One Pre-Session Choice conditions were conducted, for a total of 22 sessions. For Connor, 10, Five-Pre-Session Choices conditions, 10, Five Within-Session Choice conditions, and four, One Pre-Session Choices conditions were conducted, for a total of 24 sessions.

Five Within-Session Choices Condition

All procedures for this condition were identical to the Five Pre-Session Choices condition, with the exception of the timing of the MSW. Instead of offering five consecutive choices in the MSW format before the session, the choices were spaced evenly throughout the session, such that one choice was offered after the initial trial, and then subsequently after every four trials. The number of choices per session (five) was

held constant between these two conditions. The rationale for this condition was that evenly spacing the trials of the MSW preference assessment throughout the session would allow the participant to access a novel reinforcer if his/her preferences had shifted (i.e., satiation with the current reinforcer had occurred).

One Pre-Session Choice Condition

Due to lack of differentiation between conditions upon visual analysis of the data for all participants after ten sessions (five, Five Pre-Session Choices condition and five, Five Within-Session Choices condition), a third condition was initiated, and a randomized schedule of all three conditions was created for the remainder of sessions. In the One Pre-Session Choice condition, a single presentation of the three reinforcers was offered, and the item selected was determined to be the Designated Reinforcer for that session. After the initial choice, the One Pre-Session Choice condition was identical to the Five Pre-Session Choices condition. This condition was initiated to determine whether a single, pre-session choice could produce comparable results across the four dependent variables, thereby suggesting that a more practitioner-friendly model (due to lower response effort) would be acceptable in an applied setting.

DATA ANALYSIS

Responses were recorded in vivo as correct or incorrect. Percentage Correct Responding was calculated by dividing the observed number of correct responses by the total number of possible correct responses (i.e., 20) and the quotient was multiplied by 100. Percentage Correct Responses was graphed per session using Microsoft Excel©. Latency to Task Initiation and Total Task Duration of correct responses were determined from video recording using a timing device. Each fully completed (i.e., not terminated

early) session consisted of 20 correct responses, so the sum of the Latency to Initiation or Total Task Duration for each response was divided by 20 and multiplied by 100 to determine the mean per session. Means for these two dependent variables were graphed using Microsoft Excel©. No-Responses were determined in vivo, using a timing device, and verified using the video recording and the same timing device. The frequency of No-Responses was graphed per session in Microsoft Excel©. In the event that the session was terminated early according to the termination criteria, dependent variable values were divided by the total number of trials successfully completed (e.g., 13), and the quotient was multiplied by 100. Visual analysis determined whether or not experimental control was achieved.

EXPERIMENTAL DESIGN

This experiment utilized a multi-element experimental design (Kennedy, 2005). Two, and later three, experimental conditions were presented in a randomized order. Visual analysis determined whether or not experimental control was established, based on differentiation between/among the data paths.

INTEROBSERVER AGREEMENT

Two data collectors independently observed at least 30% of sessions per condition for each participant. For Percentage Correct Responses and No-Responses, Total Agreement Interobserver Agreement (IOA) was calculated. Data from each of the observers were compared for agreements and disagreements. Agreements were scored when both observers recorded an occurrence or nonoccurrence of the dependent variable on a specific trial. Discrepancies between the observers' scoring resulted in a

disagreement. IOA for Percentage Correct Responses and No-Responses was calculated per session using the formula:

$$\text{IOA} = \frac{\text{Agreement}}{\text{Agreements + Disagreements}} \times 100$$

For James, Percentage Correct Responses IOA was 98.8% [range: 95%-100%] and No-Responses IOA was 100%. For Aimee, Percentage Correct Responses IOA was 98.8% [range: 90%-100%] and No-Responses IOA was 100%. For Ethan, Percentage Correct Responses IOA was 99.0% [range: 96.0%-100%] and No-Responses IOA was 100%. Finally, For Connor, Percentage Correct Responses IOA was 98.8% [range: 90%-100%] and No-Responses IOA was 100%.

For Latency to Task Initiation and Total Task Duration, Total Duration IOA was calculated. Durations for each correct response were summed and divided by the total number of possible correct responses (i.e., 20) per session by each observer. Subsequently, the shorter duration recorded by an observer was divided by the longer duration recorded by the second observer, and multiplied by 100, using the formula:

$$\text{Dur IOA} = \frac{\text{Shorter duration recorded by observer}}{\text{Longer duration recorded by observer}} \times 100$$

Total Duration IOA per participant was determined by calculating the mean of all Duration IOA scores for that participant, using the formula:

$$\text{Total Dur IOA} = \frac{\text{Dur IOA } R_1 + \text{Dur IOA } R_2 + \text{Dur IOA } R_n}{n \text{ responses}} \times 100$$

For James, Latency to Task Initiation IOA was 68.4% [range: 38%-87.9%], and Total Task Duration IOA was 97.3% [range: 94.6%-99.5%]. For Aimee, Latency to Task Initiation IOA was 89.1% [range: 75.4%-99.0%], and Total Task Duration IOA was 94.5% [range: 85.3%-97.0%]. For Ethan, Latency to Task Initiation IOA was 81.9% [range: 70%-98.8%], and Total Task Duration IOA was 94.2% [range: 84%-100%]. Finally, for Connor, Latency to Task Initiation IOA was 80.5% [range: 42.4%-96.4%], and Total Task Duration IOA was 98.1% [range: 93.8%-99.7%]. See Table 5 in the Appendix for an example of the coding sheet on which IOA was calculated per session and Table 6 in the Appendix for expanded IOA information for all participants.

Table 4
Interobserver Agreement

Participant	Latency to Task Initiation	Total Task Duration	Percentage Correct Responding	No-Responses
James	68.4%	97.3%	98.8%	100%
Aimee	89.1%	94.5%	98.8%	100%
Ethan	81.9%	94.2%	99%	100%
Connor	80.5%	98.1%	98.8%	100%

Table 4: Interobserver Agreement across dependent variables

TREATMENT FIDELITY

Treatment Fidelity was calculated by a second observer, using one of three coding sheets (See Table 6 in the Appendix for an example), for 30% of each condition per participant. A specific task analysis was created for each condition. Raters were asked to watch the video recordings of randomly selected sessions and circle "Yes", "No", or "Not Applicable" for each step of the relevant task analysis. Treatment fidelity was calculated per session by dividing the number of "Yes" responses, minus the number of "Not Applicable" responses, divided by the total number of possible responses, minus the number of "Not Applicable" responses, using the formula:

$$\text{Treatment Fidelity} = \frac{\text{"Yes" responses} - \text{"Not Applicable" responses}}{n \text{ responses} - \text{"Not Applicable" responses}} \times 100$$

Treatment Fidelity for James was 99.1% [range: 96%-100%], for Aimee was 99.3% [range: 96.4%-100%], for Ethan was 99.3% [range: 97.1%-100%], and for Connor was 99.8% [range: 99.0%-100%].

Chapter 4: Results

For each participant, the experiment consisted of a single session of pre-assessment and 22-26 sessions of the experimental preparation. Results of the pre-assessment are presented first to demonstrate that each participant met inclusion criteria (at least five consecutive correct responses with stable Latency to Task Initiation and Total Task Duration on a previously mastered task) prior to entering the experiment. Following pre-assessment data for each participant is the experimental data for each participant, as expressed according to each dependent variable (Latency to Task Initiation, Total Task Duration, Percentage Correct Responses, and No-Responses).

Pre-Assessment

James engaged in the simple task assessment using the same task materials as were used in the experiment. Correct answers followed by delivery of a preferred edible (popcorn), which was the same preferred item used when initially learning the response. On the first five responses, James exhibited consecutive correct responses. The graph demonstrated stable Latency to Task Initiation across trials, demonstrating mastery of the current task in terms of this dependent variable.

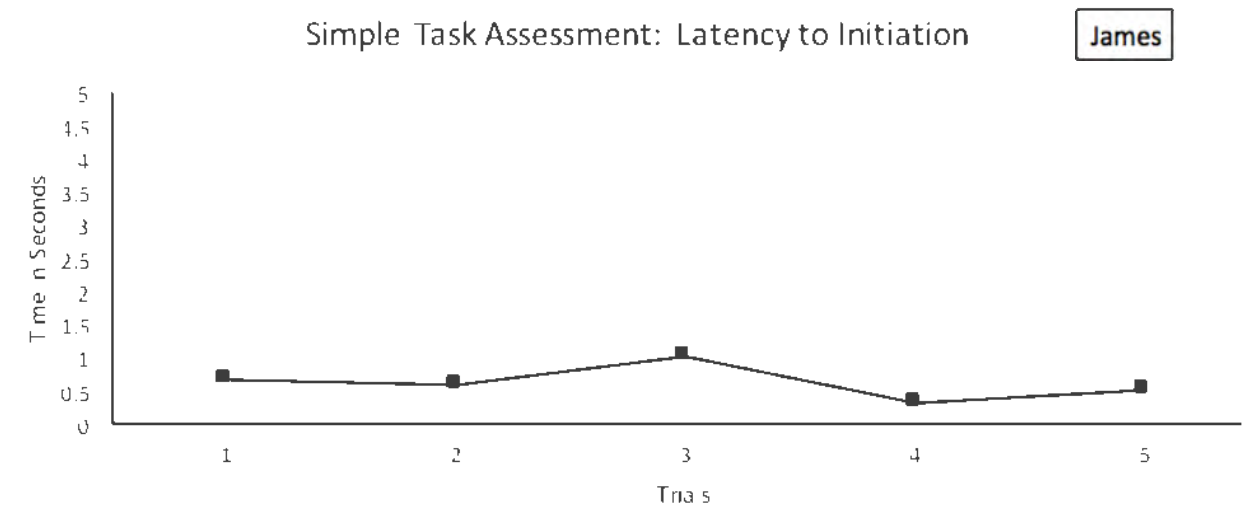

Figure 5: Simple Task Assessment: Latency to Task Initiation data for James

The graph also demonstrated stable Total Task Duration across trials, and therefore, it was determined that the task was mastered in terms of this dependent variable. James was accepted into the experiment.

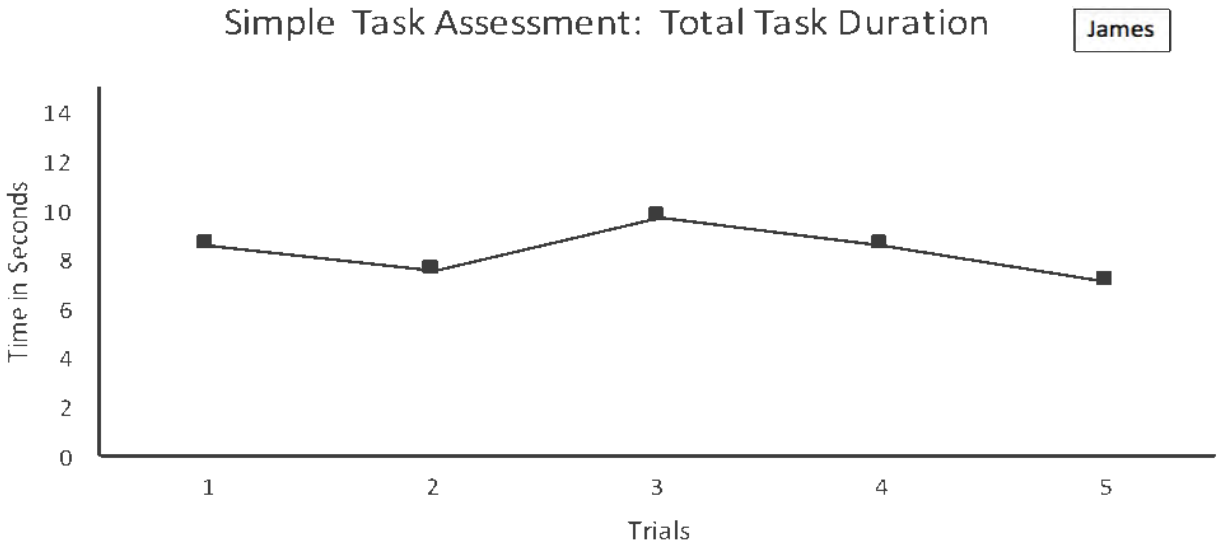

Figure 6: Simple Task Assessment: Total Task Duration data for James

Aimee engaged in the simple task assessment using the same task materials as were used in the experiment. Correct answers were reinforced using the stimulus delivered for reinforcement when the task was originally taught (a video clip of the Teletubbies© on the iPad). On the first six trials, Aimee exhibited consecutive correct responses. The graph demonstrated stable Latency to Task Initiation across trials 2-6, and therefore, it was determined that the task had already been mastered in terms of this dependent variable.

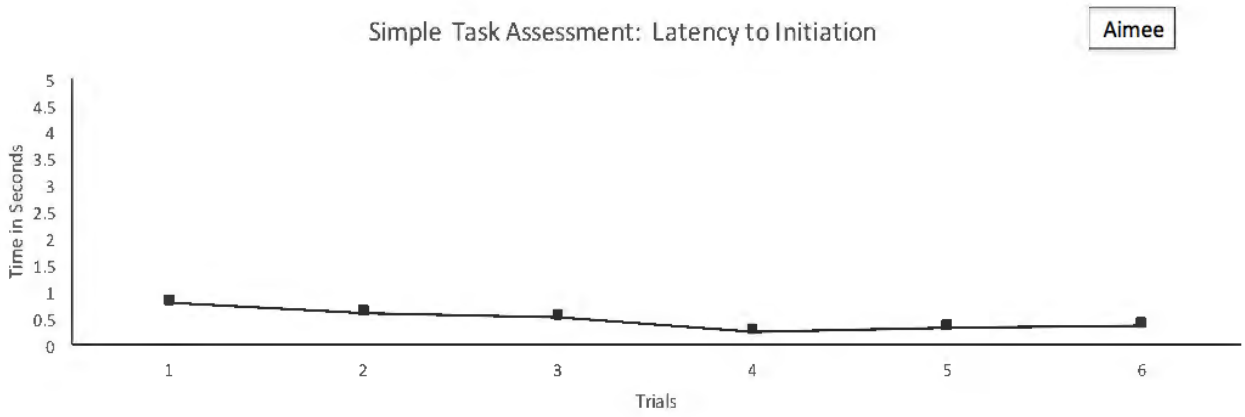

Figure 7: Simple Task Assessment: Latency to Task Initiation data for Aimee

The graph also demonstrated stable Total Task Duration across trials 2-6, and therefore, it was determined that the task had already been mastered in terms of this dependent variable. Aimee was accepted into the experiment.

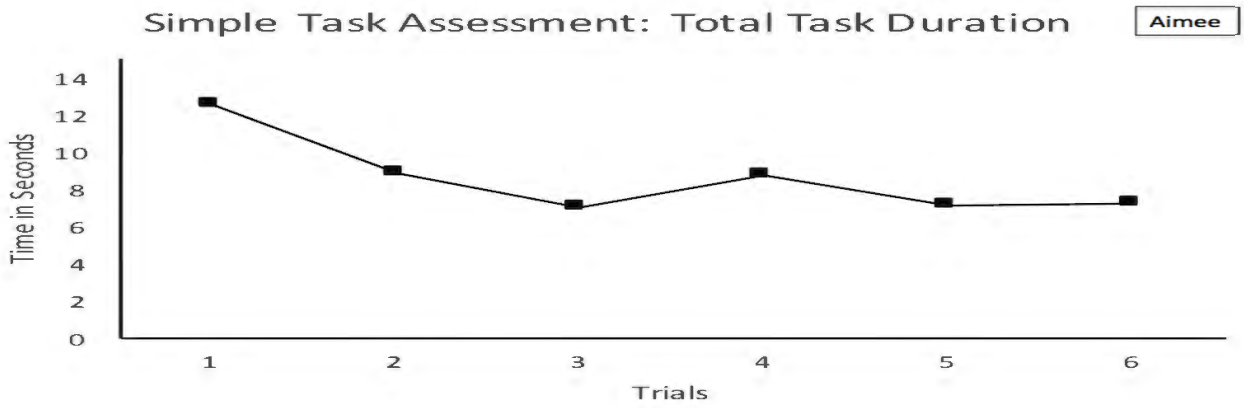

Figure 8: Simple Task Assessment: Total Task Duration data for Aimee

Ethan engaged in the simple task assessment using the same task materials as were used in the experiment. Correct responses were reinforced with the same stimulus delivered for reinforcement as when the task was originally taught (a video clip of a school bus). On the first 13 responses, Ethan exhibited consecutive correct responses. The graph demonstrated stable Latency to Task Initiation across all trials, and therefore, it was determined that this skill was already mastered in terms of this dependent variable.

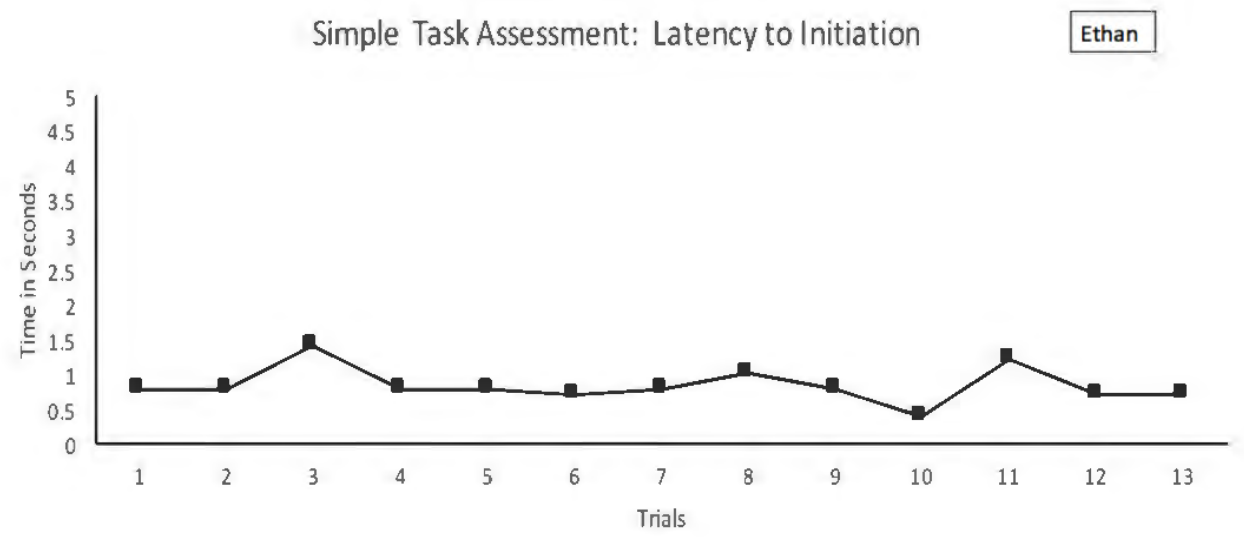

Figure 9: Simple Task Assessment: Latency to Task Initiation data for Ethan

The graph also demonstrated stable Total Task Duration across trials 9-13, and therefore, it was determined that the task was already mastered for this dependent variable. Ethan was

accepted into the experiment.

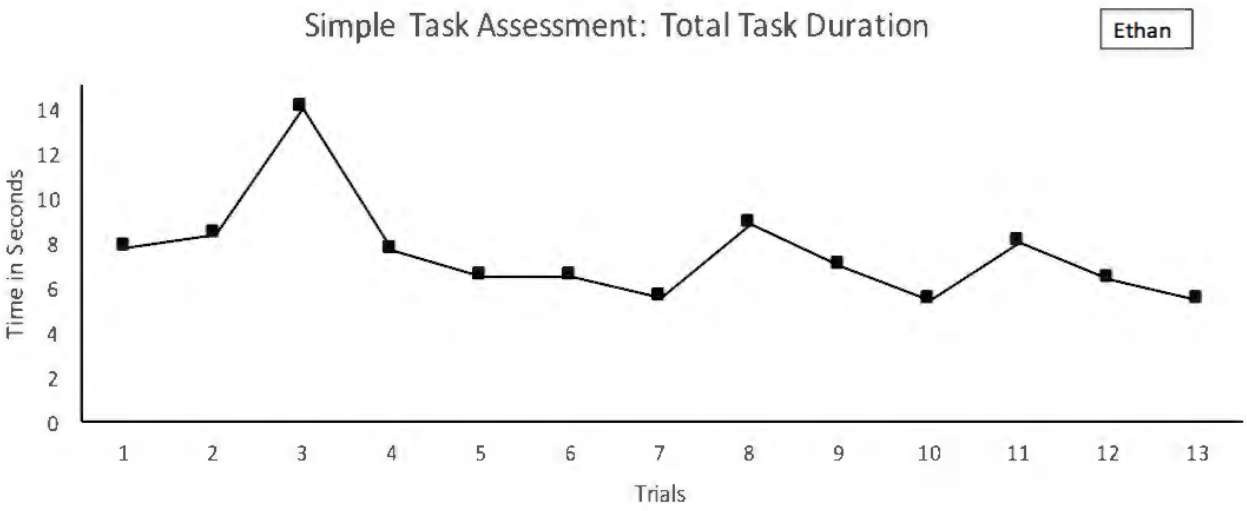

Figure 10: Simple Task Assessment: Total Task Duration data for Ethan

Connor engaged in the simple task assessment using the same task materials as were used in the experiment. Correct responses were reinforced with the same stimulus delivered when the task was originally taught. On the first six trials, Connor exhibited consecutive correct responses. The graph demonstrated stable Latency to Task Initiation across responses, and therefore, it was determined that the task was already mastered in terms of this dependent variable.

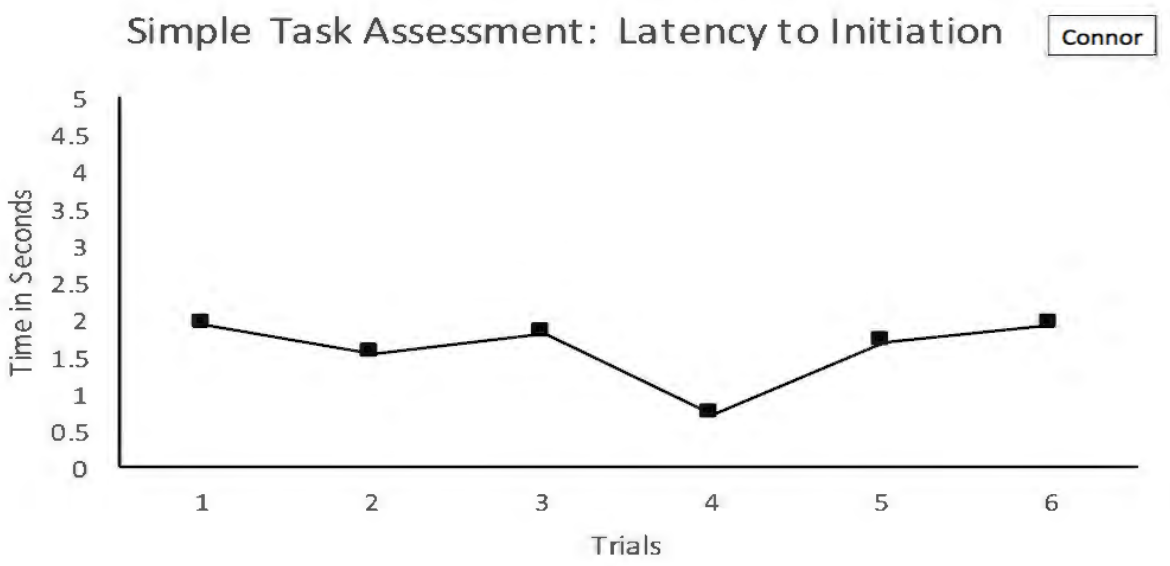

Figure 11: Simple Task Assessment: Latency to Task Initiation data for Connor

The graph also demonstrated stable Total Task Duration across responses, and therefore, it was determined that the task was already mastered in terms of this dependent variable. Connor was accepted into the experiment.

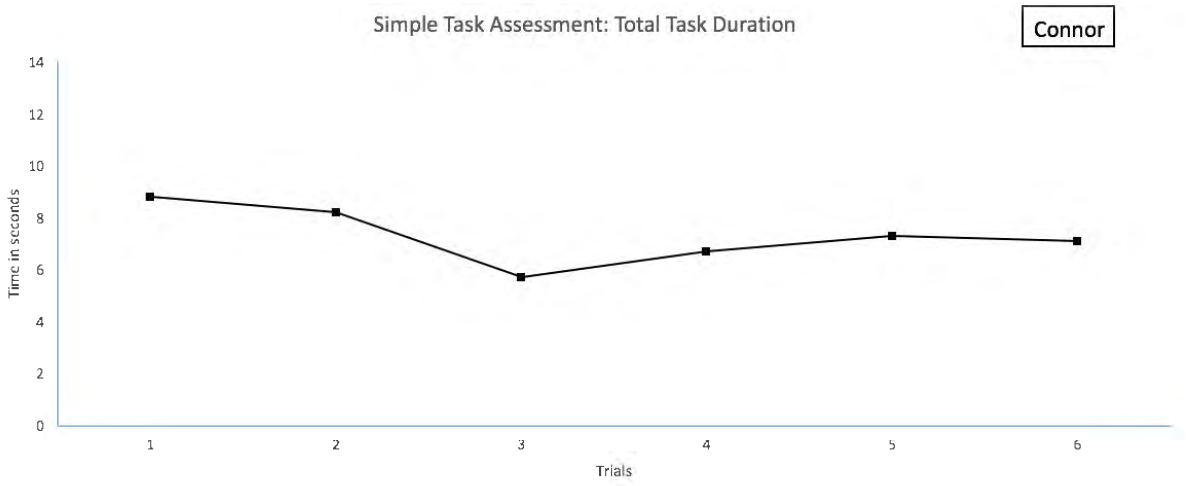

Figure 12: Simple Task Assessment: Total Task Duration data for Connor

Experiment

For the first 10 sessions, James' Latency to Task Initiation data demonstrates undifferentiated data paths with a slight downward trend for the first five data points and a range of latencies from 0.5 s to 1.5 s. James began using two hands to place pieces on the task board after the third session, which shortened both his Latency to Task Initiation and Total Task Duration times. Given that the conditions were undifferentiated, and it did not appear to make a difference in performance whether the choices were offered prior to the session or throughout the session, a third condition, "One Pre-Session Choice" was initiated to determine whether an experimental preparation with lower practitioner response effort could produce similar results in terms of this dependent variable. On session 11, James selected donuts as his reinforcer when given one pre-session choice. In general, this reinforcer appeared to produce the lowest motivation, as demonstrated by

slightly longer Latency to Task Initiation and a greater frequency of No-Responses (see Figure 12). Anecdotally, challenging behavior in the form of elopement from work area was also elevated on session 11. Across sessions, donuts were chosen least frequently through the various choice variations (see Figure 13). It appeared that, when making a selection for a less-preferred item, the restriction of only having one opportunity to make a choice and no opportunities to select a novel reinforcer had an impact on these dependent variables. Across the remaining 11 data points, Latency to Task Initiation for James was also undifferentiated and ranged from 0.8 s to 1.8 s.

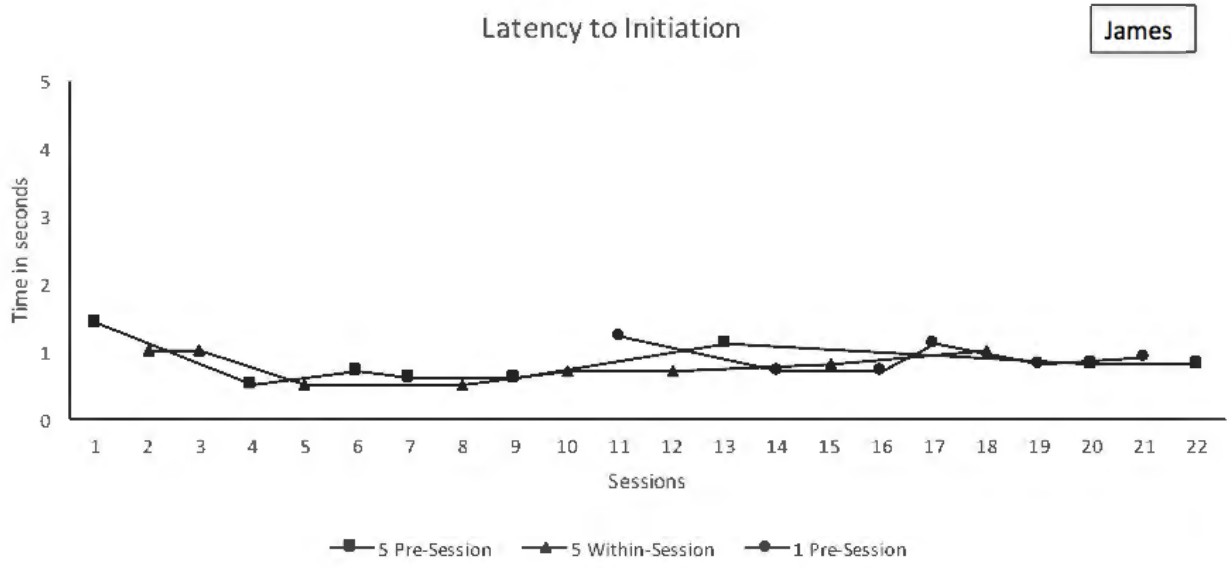

Figure 13: Latency to Task Initiation data for James

For the first ten sessions, James' Total Task Duration data was undifferentiated with a more pronounced downward trend. Total Task Duration was affected more markedly by James' shift to using two hands to place pieces on the task board. Total Task Duration across the first ten data points ranged from 6.1 s to 10.0 s. When the "One Pre-Session Choice" condition was introduced, data paths remained undifferentiated and continued their downward trend, with the exception of session 21, during which a more lengthy Total Task Duration was exhibited.

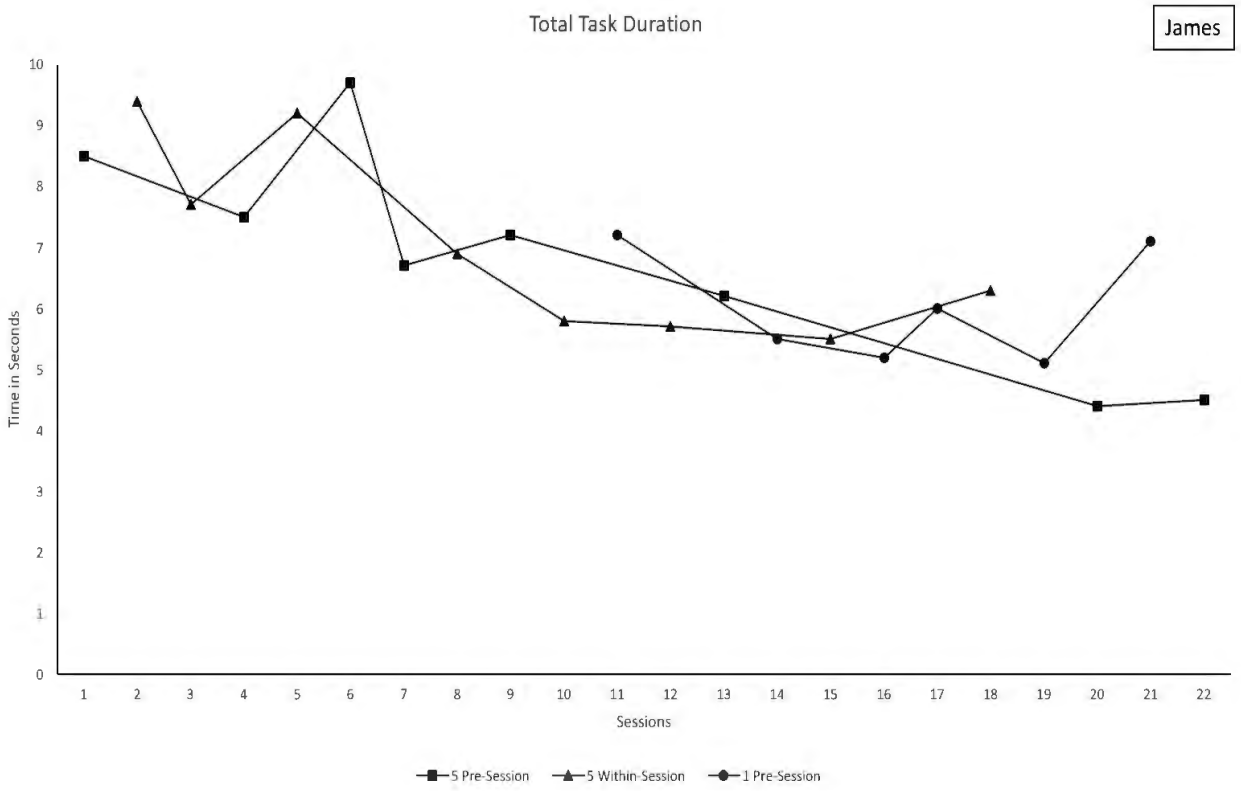

Figure 14: Total Task Duration data for James

Across the first ten data point, James exhibited undifferentiated data for the dependent variable of Percentage Correct Responding across the "Five Pre-Session Choices" and "Five Within-Session Choices" conditions, with percentages ranging from 90% to 100%. When the "One Pre-Session Choice" condition was introduced, Percentage Correct Responses remained undifferentiated and high [range 85%-100%]. Most errors exhibited by James appeared to result from not attending to the stimuli or momentarily being distracted by a noise in the surrounding classrooms.

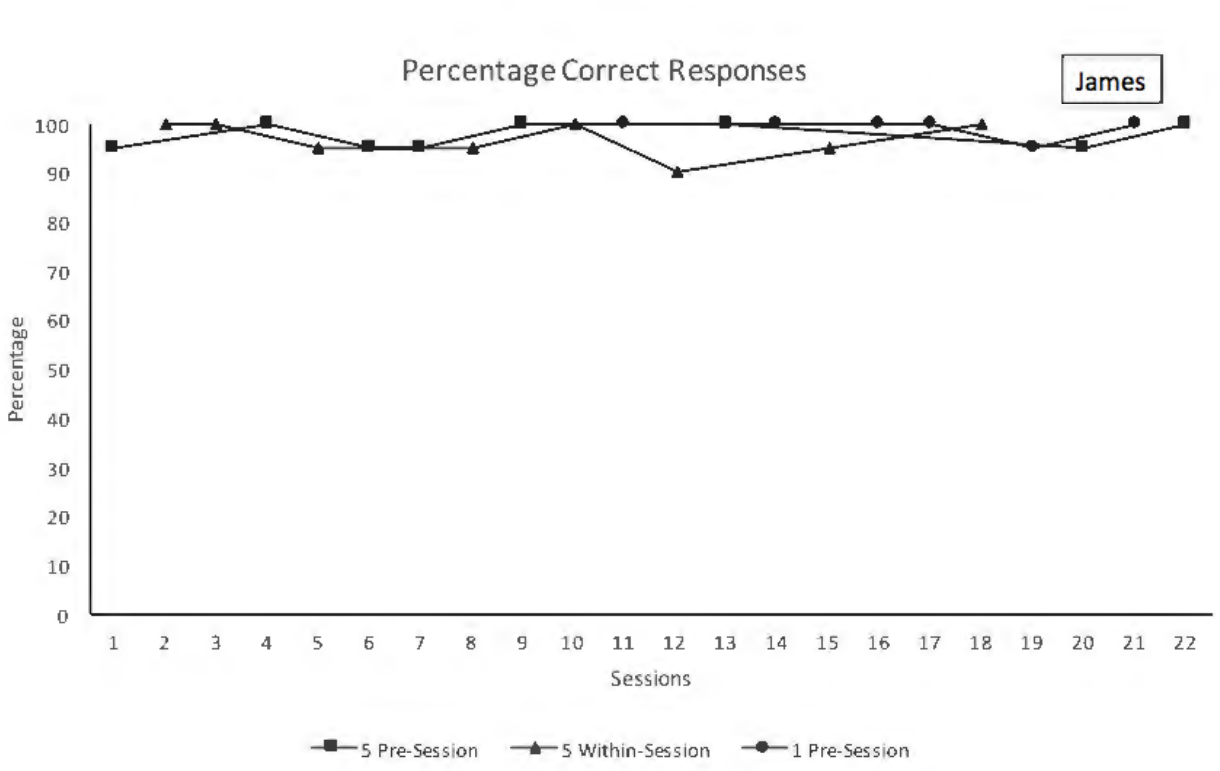

Figure 15: Percentage Correct Responding data for James

For the first ten sessions, James' No-Responses data was undifferentiated across the "Five Pre-Session Choices" condition and the "Five Within-Session Choices"

conditions, with frequency counts ranging from 0-2. When the “One Pre-Session Choice” condition was introduced, data paths remained largely undifferentiated, with the exception of sessions 11, 13, 20. Even with this slight differentiation, there does not appear to be a condition or conditions in which James consistently had elevated frequency of No-Responses, and therefore, the data continues to be characterized as undifferentiated after the introduction of the “One Pre-Session Choice” condition. Frequency count of No-Responses ranged from 0-3 for sessions 11-22.

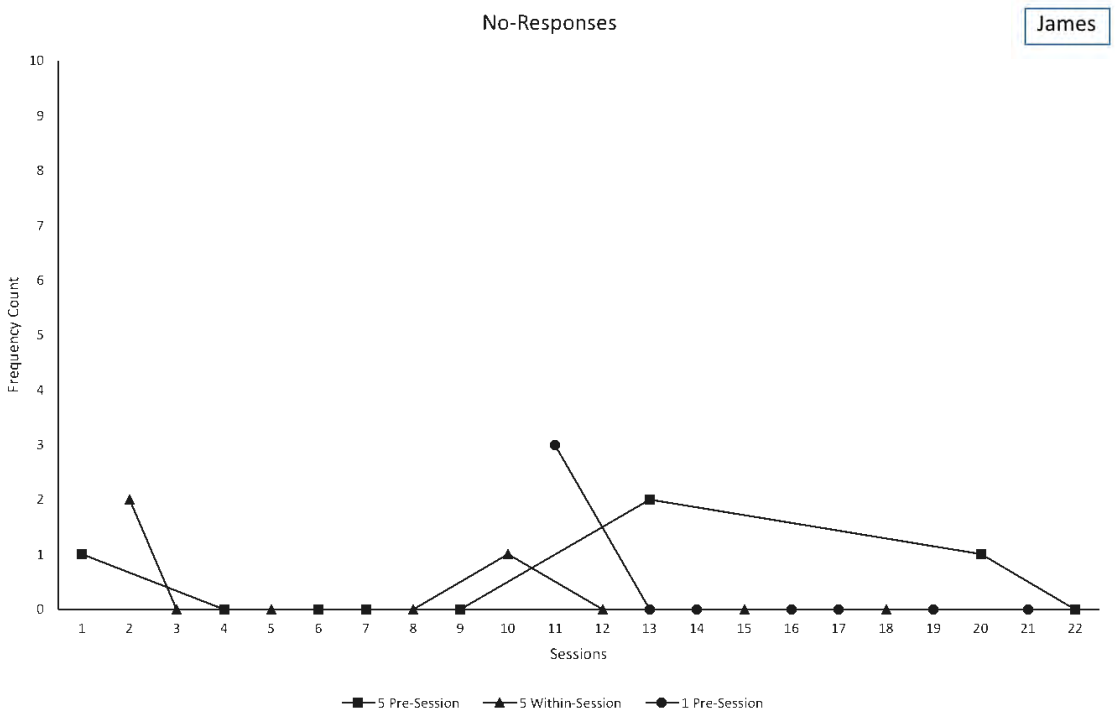

Figure 16: No-Responses data for James

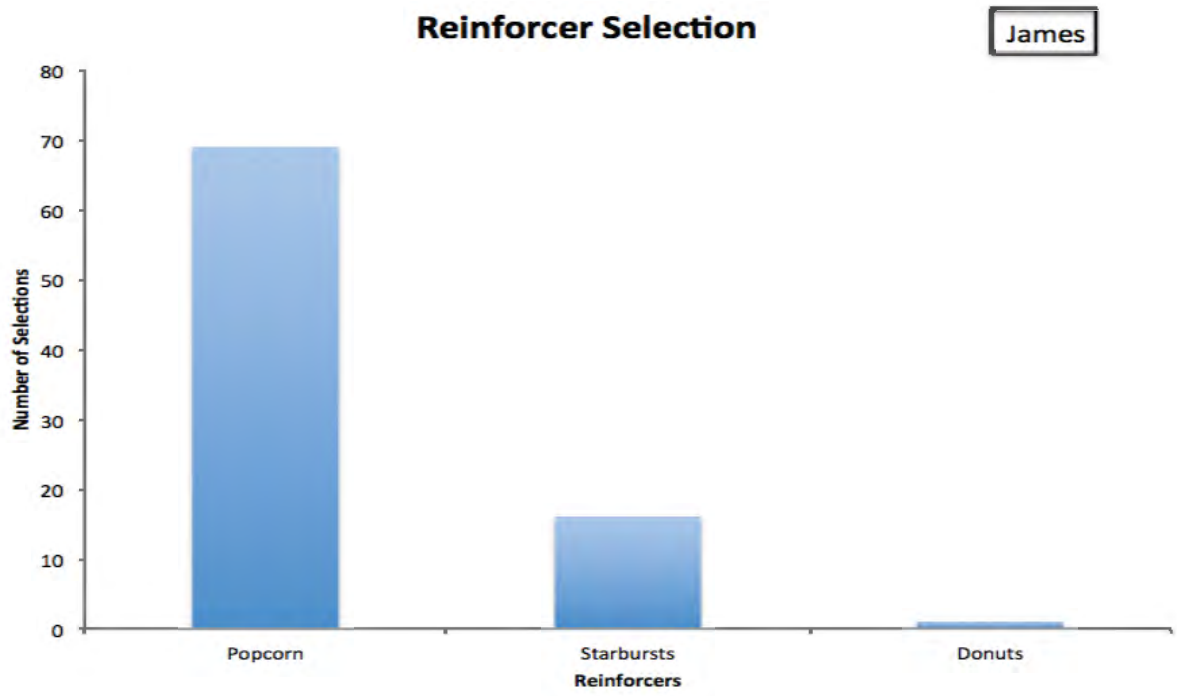

Figure 17: Reinforcer Selection data for James

James chose popcorn as his reinforcer for the majority of sessions (18). See Figure 13 for a graphical representation of James' reinforcer choices.

For the first ten sessions, Aimee's Latency to Task Initiation was undifferentiated, with a very narrow range of latencies from 0.8 s to 1.3 s and no apparent directional trend. Given the undifferentiated nature of the data paths, a "One Pre-Session Choice" condition was introduced to examine the effect of a lower practitioner response effort experimental preparation on all dependent variables. After the introduction of the "One Pre-Session Choice" condition, Latency to Task Initiation remained highly undifferentiated and stable with no apparent directional trend. Latencies ranged from 0.8 s to 1.2 s for sessions 11-22.

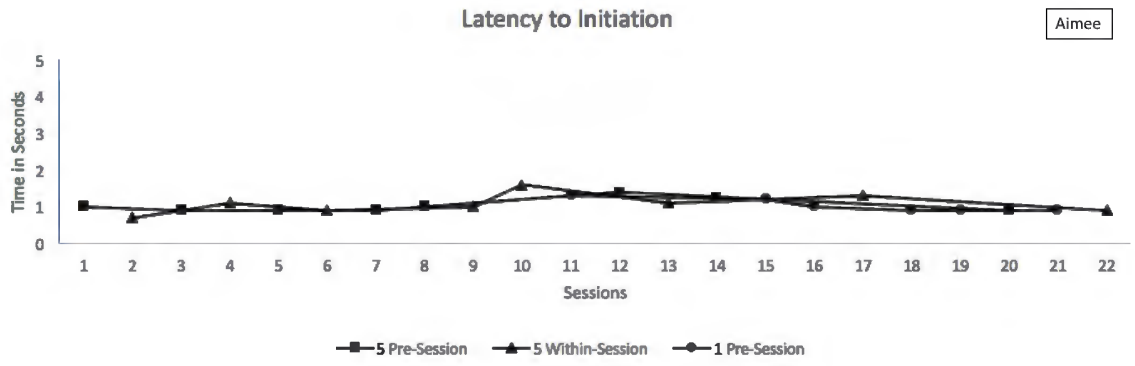

Figure 18: Latency to Task Initiation data for Aimee

For the first ten sessions, Aimee's Total Task Duration data was undifferentiated across the "Five Pre-Session Choices" and "Five Within-Session Choices" conditions with no apparent directional trend. Total Task Duration for these sessions ranged from 6.1 s to 11.5 s. Anecdotally, Aimee engaged in high levels of head banging and crying during session 4, which appears to have lead to elevated Total Task Duration for that session. After the introduction of the "One Pre-Session Choice" condition, data remained undifferentiated, with a narrower range of durations from 5.9 s to 8.1 s.

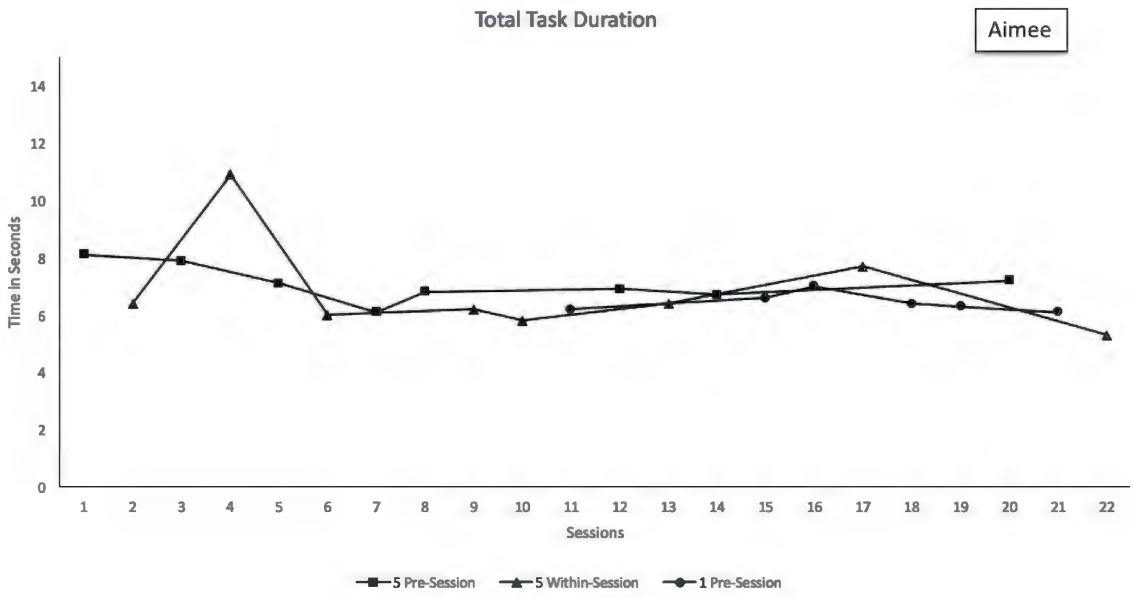

Figure 19: Total Task Duration data for Aimee

For the first ten sessions, Aimee's Percentage Correct Responding was variable but high, ranging from 85% to 100%, and undifferentiated, across the "Five Pre-Session Choices" and "Five Within-Session Choices" conditions. When the "One Pre-Session Choice" condition was introduced, Percentage Correct Responding remained undifferentiated, but became slightly less variable, with the exception of session 22, on which Aimee exhibited her lowest Percentage Correct Responding percentage of 80%. Percentages ranged from 80% to 100% for sessions 11-22.

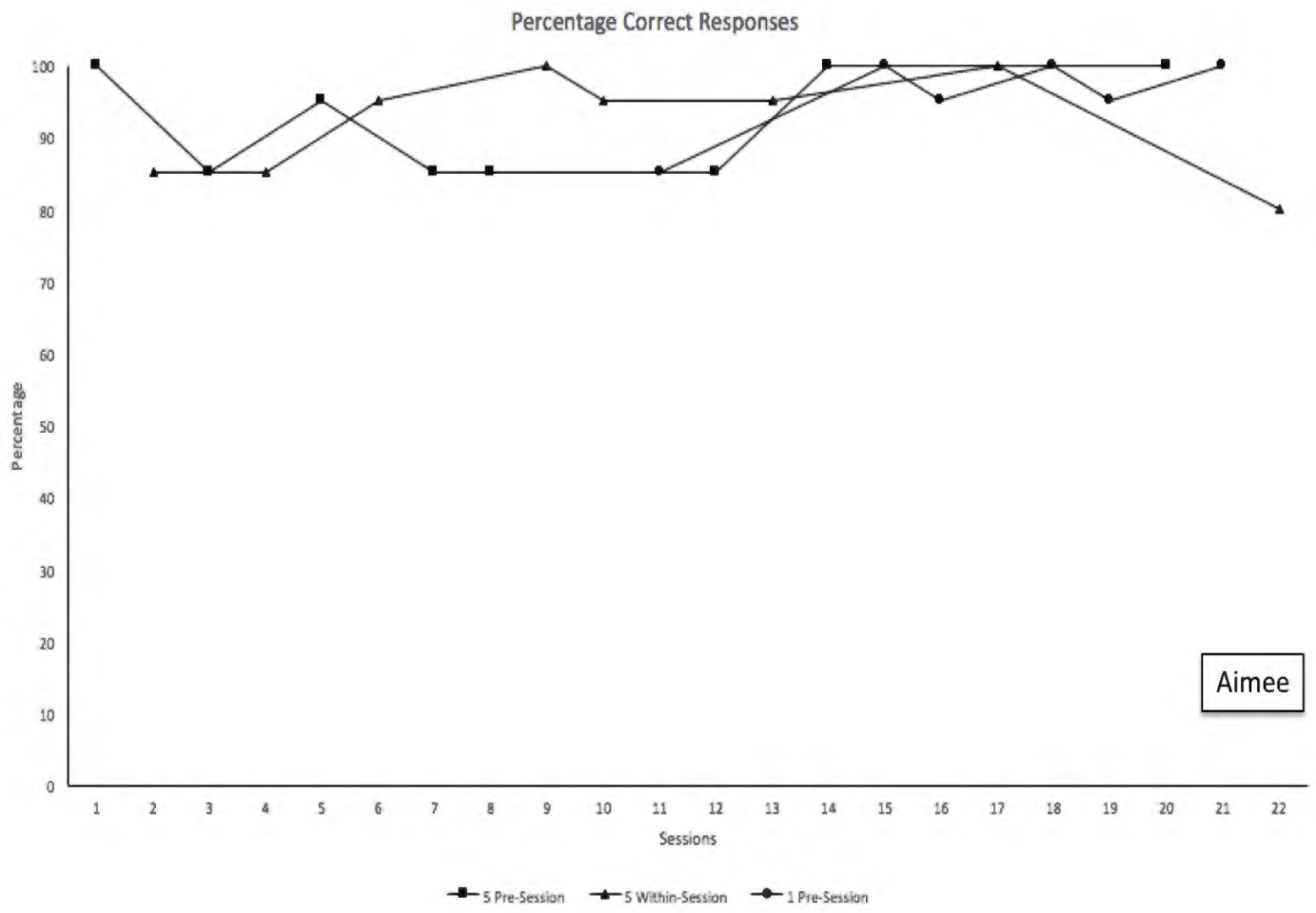

Figure 20: Percentage Correct Responding data for Aimee

For the first ten sessions, Aimee's No-Responses data was undifferentiated and variable across the "Five Pre-Session Choices" and "Five Within-Session Choices" conditions. The frequency count for each session ranged from 0-6 No-Responses for sessions 1-10. As stated above, Aimee demonstrated elevated levels of challenging behavior (head banging and crying) during session 4, which appears to affected her frequency of No-Responses for that session.

After the introduction of the "One Pre-Session Choice" condition, No-Responses data paths remained variable and undifferentiated, with a frequency count range of 0-2.

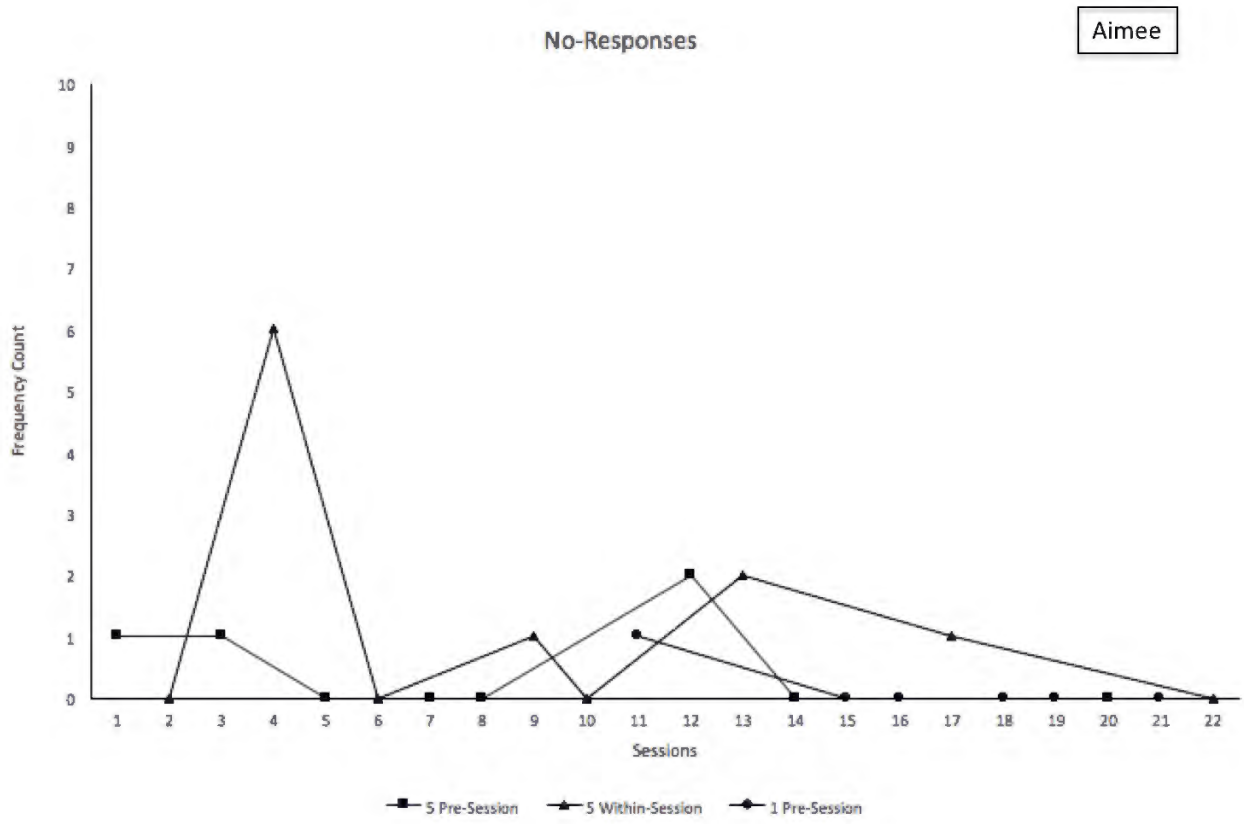

Figure 21: No-Responses Data for Aimee

Aimee selected the video "Line Up" for the majority of sessions (21). See Figure 22 for a graphical representation of Aimee's reinforce selections.

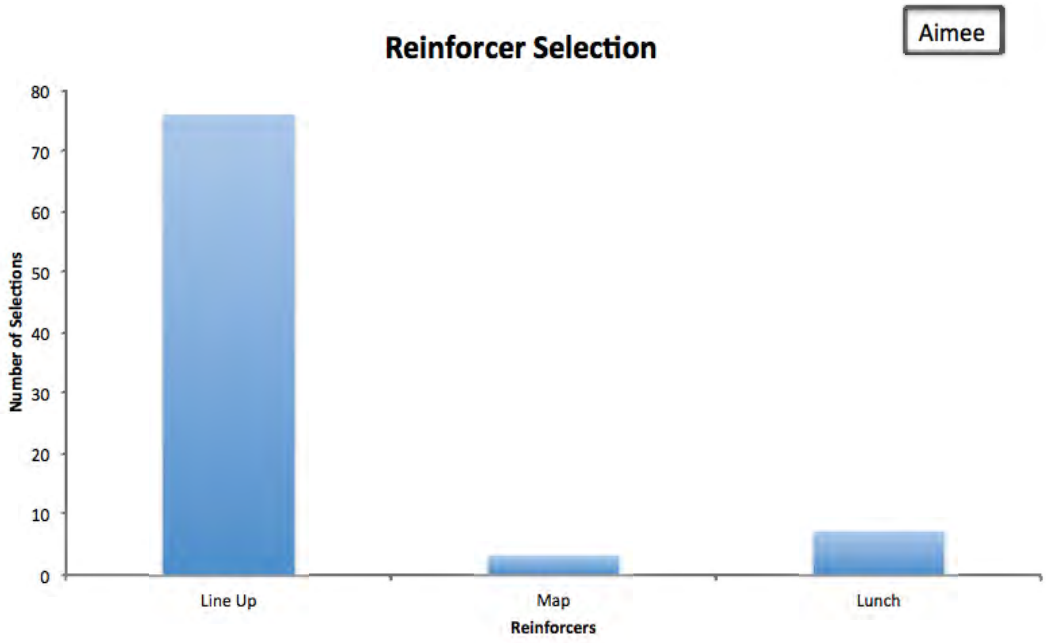

Figure 22: Reinforcer Selection data for Aimee

For the first ten sessions, Ethan's Latency to Task Initiation was undifferentiated across the "Five Pre-Session Choices" and "Five Within-Session Choices" conditions with no apparent directional trend. Latencies ranged from 1.0 s to 1.6 s for sessions 1-10. A "One Pre-Session Choice" condition was introduced to examine the effect of a lower practitioner response effort experimental preparation on dependent variables. After the introduction of the third condition, data paths became more variable, yet remained undifferentiated. The open markers on sessions 15, 16, 17, and 18 indicate that the termination criterion was met for that session and that the session was terminated early. The number next to each open marker indicates the number of trial on which the session was terminated. The number one next to session 17 indicates that not even a single trial was successfully completed prior to meeting the termination criterion. Anecdotally, it is reported that challenging behavior in the form of elopement from the work area and

noncompliance were elevated for all sessions after session 13. After four consecutive days of early termination, it was determined that Ethan would only participate in the experiment if his mother remained in the bedroom. She did not interact with him at all, but it appeared that her presence reduced the amount of challenging behavior, and Ethan was able to complete all 20 trials when she was in the room. Latencies ranged from 0 s (session 17) to 2.1 s for sessions 11-22.

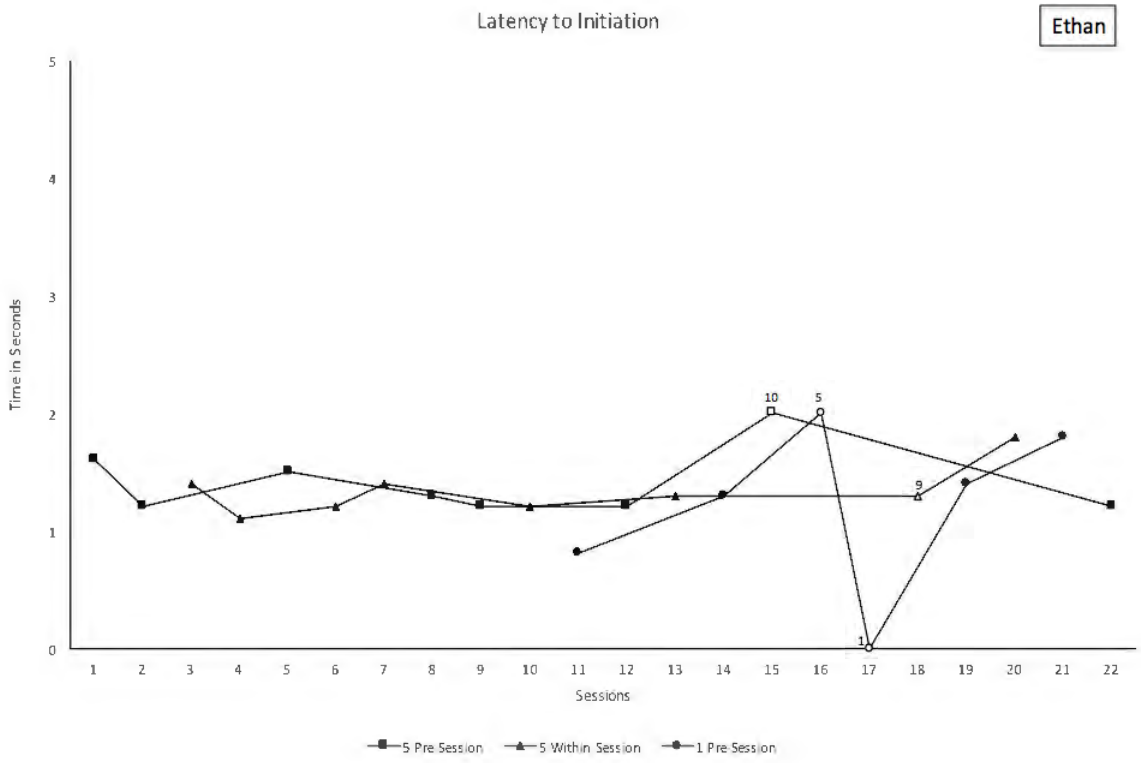

Figure 23: Latency to Task Initiation data for Ethan. Open markers indicate that termination criterion was met during that session. Numbers indicate the number of trial on which the session was terminated. Ethan's mother was present in the room for sessions 19, 20, 21, and 22.

For the first ten sessions, Ethan's Total Task Duration data were highly undifferentiated, stable, and exhibited no directional trend. Durations range from 9.9 s to

12.3 s. After the introduction of the "One Pre-Session Choice" condition, data paths became much more variable but remained undifferentiated and continued to have no apparent directional trend. Sessions were terminated early on sessions 15-18.

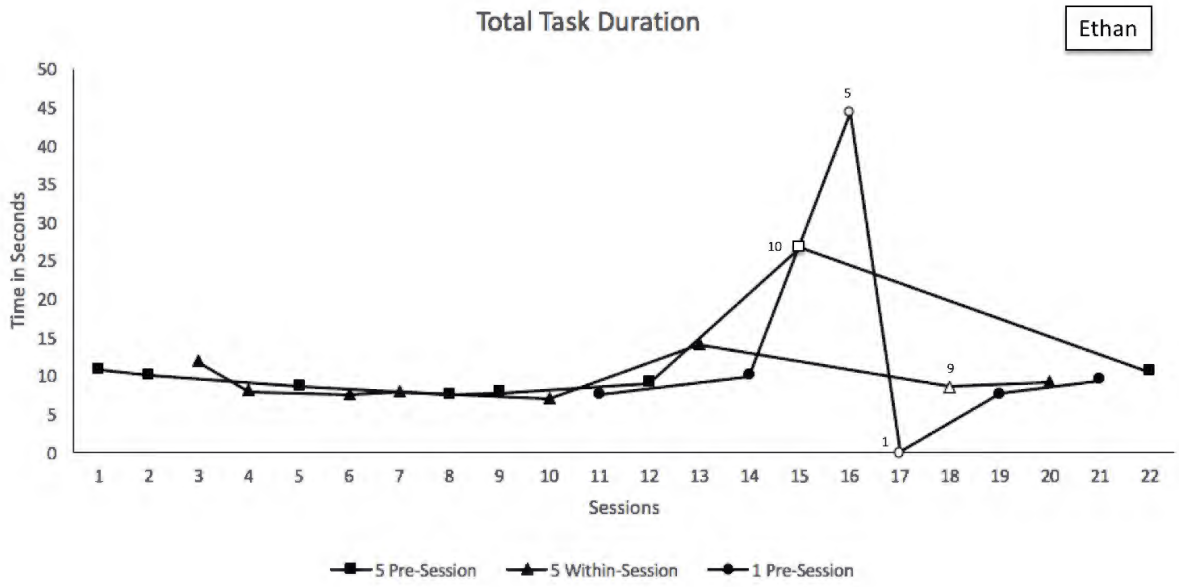

Figure 24: Total Task Duration data for Ethan. Open markers indicate that termination criterion was met during that session. Numbers indicate the number of trial on which the session was terminated. Ethan's mother was present in the room for sessions 19, 20, 21, and 22.

For the first ten sessions, Ethan's Percentage Correct Responding data were variable but high, undifferentiated, and had no apparent directional trend. Percentages ranged from 90%-100%. After the introduction of the "One Pre-Session Choice" condition, data became more variable and remained largely undifferentiated. Once the

third condition was introduced, all sessions in which the "Five Pre-Session Choices" condition were used demonstrate sub-100% percentages. In general, percentages were slightly lower in this condition than in the "Five Within-Session Choices" condition (one session with sub-100% percentage) and the "One Pre-Session Choice" condition (two sessions with sub-100% percentage). Percentages ranged from 0% (session 17) to 100%.

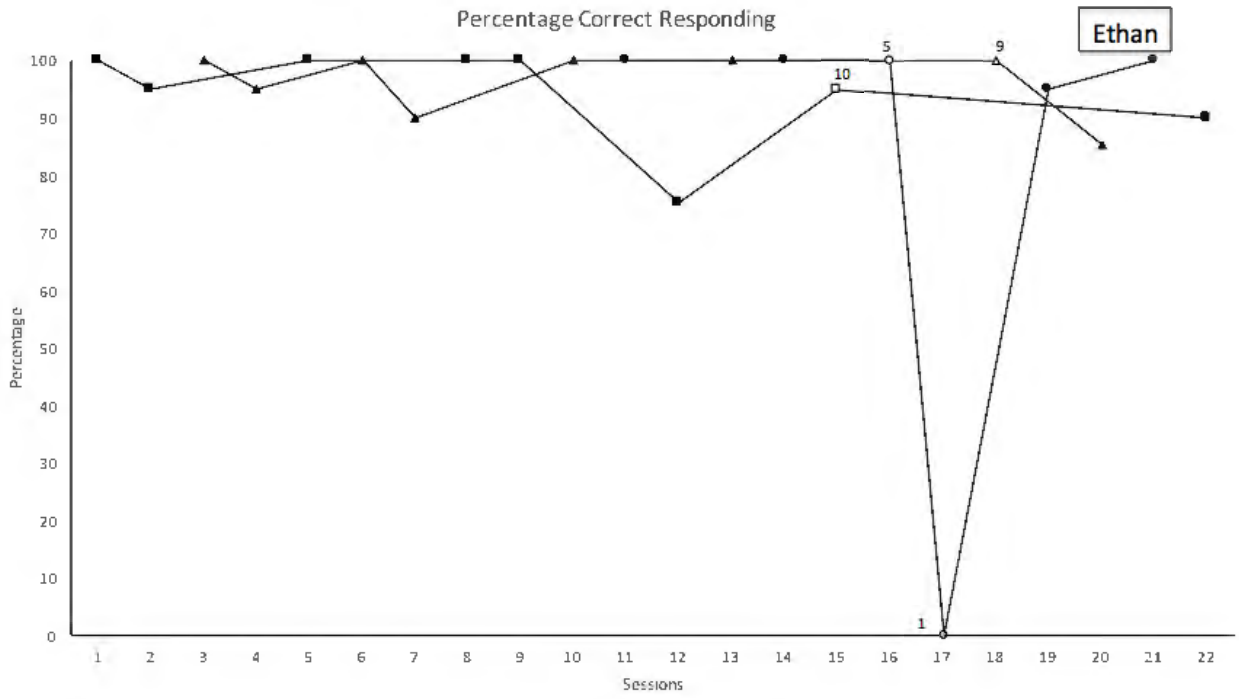

Figure 25: Percentage Correct Responding data for Ethan. Open markers indicate that termination criterion was met during that session. Numbers indicate the number of trial on which the session was terminated. Ethan's mother was present in the room for sessions 19, 20, 21, and 22.

For the first ten sessions, Ethan's No-Responses data was low, slightly variable, and undifferentiated, with no apparent directional trend. Frequency count of No-Responses ranged from 0-1 for sessions 1-10. After the introduction of the "One Pre-Session Choice" condition, data became much more variable, with elevated frequency counts exhibited on sessions 15-18. Increased No-Responses levels, in general, corresponded with sessions on which elevated challenging behavior was demonstrated (sessions 13-18). When Ethan's mother remained in the room for the final four sessions, No-Responses data became low and less variable. Frequency counts ranged from 0-11 No-Responses for sessions 11-22.

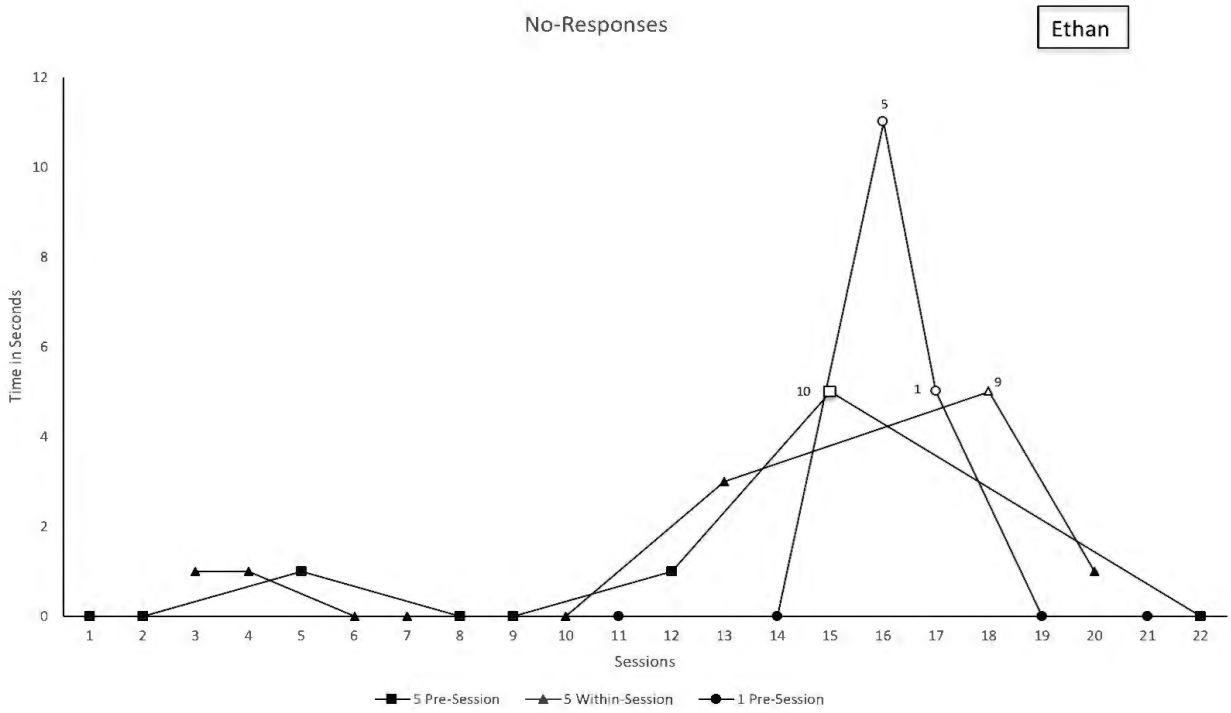

Figure 26: No-Responses data for Ethan. Open markers indicate that termination criterion was met during that session. Numbers indicate the number of trial on which session was terminated. Ethan's mother was present in the room for sessions 19, 20, 21, and 22.

Ethan selected the video, "Elevator" for 55 opportunities, the video "Subway" for 19 opportunities, and the video, "City Bus" for 8 opportunities. See Figure 27 for a graphical representation of Ethan's reinforcer selections.

Figure 27: Reinforcer selections for Ethan

For the first ten sessions, Connor's Latency to Task Initiation data was undifferentiated and low, with no apparent directional trend. Latencies ranged from 0.7-2.0 s. The open marker on session 1 indicates that he met termination criterion for that session, and the number next to the marker indicates the trial on which criterion was met. From sessions 5-10, Connor began eloping from the table during the reinforcement period (i.e., he would complete the task and elope from the table as soon as the video began to play). A "No Reinforcement" phase was initiated to evaluate whether the videos had lost their reinforcer efficacy. This phase was identical to the first ten sessions, with the exception that the videos were not played contingent upon completion of the task. In

other words, the task was presented a total of 21 times consecutively. Correct answers were followed by a neutral "That's it" statement and typical error correction was utilized (see Method). It was hypothesized that if the videos were still reinforcing, that the removal of reinforcement would result in a discernable change in dependent variables. Latencies ranged from 1.0 to 2.1 s. Although the data paths for the dependent variable of Latency to Task Initiation remained stable for the "No Reinforcement" phase, there were changes in other dependent variables, which led to the initiation of the final phase. Similar to other participants, a "One Pre-Session Choice" condition was initiated and the three phases were randomly rotated for the remainder of sessions. Upon the initiation of this phase, data paths remained undifferentiated. Connor met termination criterion on sessions 18, 19, 20, and 21, representing all three conditions.

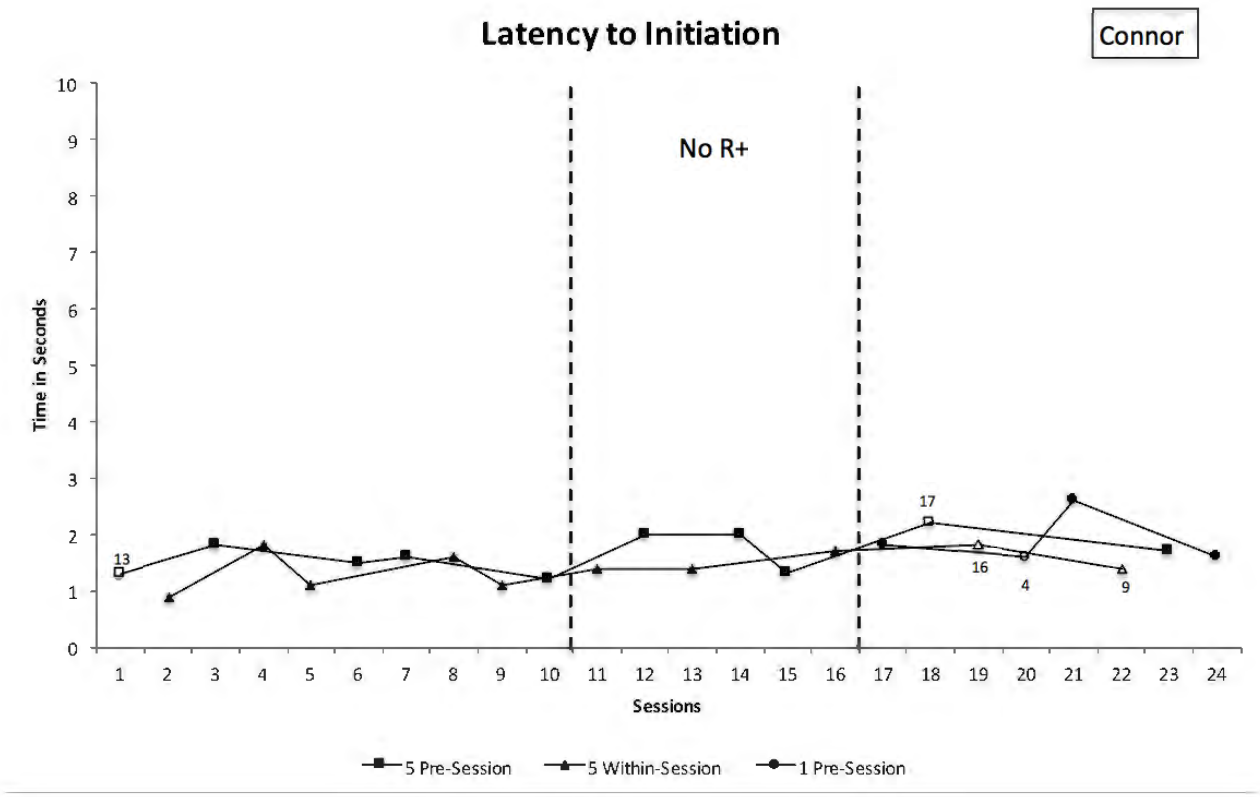

Figure 28: Latency to Task Initiation data for Connor

For the first ten sessions, Connor's Total Task Duration data was undifferentiated, with no apparent directional trend. Durations ranged from 5.8 to 17.3 s for sessions 1-10. A "No Reinforcement" phases was initiated from sessions 11-16 to evaluate the reinforcer efficacy of the videos. Data remained undifferentiated for this dependent variable. Durations ranged from 6.7 to 51.1 s for sessions 11-16. Sessions 12 and 14 (Condition 1) demonstrated Total Task Duration that was considerably higher than durations demonstrated when reinforcement was present. Due, in part, to the change in dependent variable of Total Task Duration when reinforcement was removed, it was determined that the videos still retained moderate reinforcer efficacy, and after six sessions, a "One Pre-Session Choice" condition was introduced to examine the effect of a

lower practitioner response effort experimental preparation on dependent variables. After the introduction of the third phase, data paths remained undifferentiated. Durations ranged from 6.5 to 24.1 s for sessions 17-24.

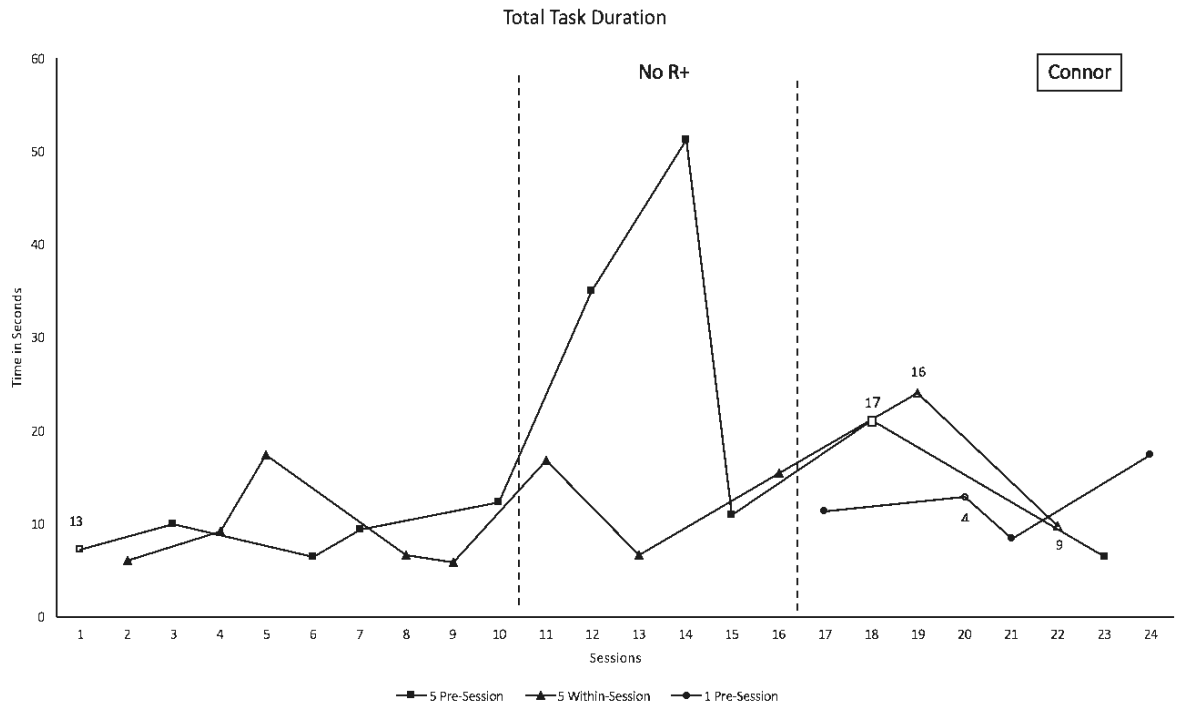

Figure 29: Total Task Duration data for Connor

For the first ten sessions, Connor's Percentage Correct Responding data was undifferentiated and extremely stable, with all sessions representing 100% accuracy. A "No Reinforcement" phase was initiated to evaluate the reinforcer efficacy of the videos. Data paths demonstrated differentiation, with two sessions in the "Five Within-Session Choices" condition demonstrating accuracies below 100% (95% for session 11 and 90% for session 13). Although his accuracy was still very high, exhibiting sub-100% accuracy on a clearly mastered task suggested that the removal of reinforcement had a slight effect

on motivation, and that the videos still retained moderate reinforcer efficacy. A "One Pre-Session Choice" condition was introduced to evaluate the effect of an experimental preparation with lower practitioner response effort on dependent variables. Percent Correct Responding remained at 100% for all remaining sessions. Connor met termination criterion for sessions 18, 19, 20, and 22.

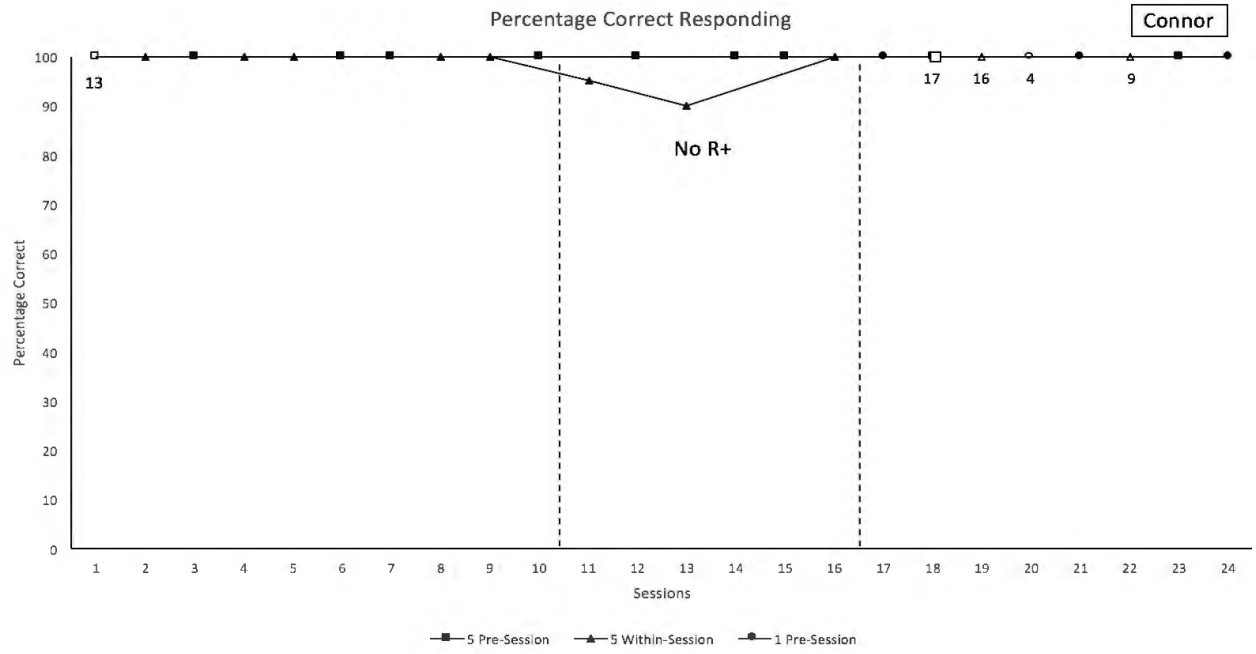

Figure 30: Percentage Correct Responses data for Connor

For the first 10 sessions, Connor's No-Responses data was differentiated and unstable, with a higher frequency of No-Responses in the "Five Pre-Session Choices" condition [range 1-9] than the "Five Within-Session Choices" condition [range 0-1]. A "No Reinforcement" phase was initiated to evaluate the reinforcer efficacy of the videos. Data paths remained undifferentiated for sessions 11-14. On session 15, Connor

demonstrated a very low frequency of No-Responses (one) in the “Five Pre-Session Choices” condition and on session 16, he demonstrated a relatively high frequency of No-Responses in the “Five Within-Session Choices” condition (nine), leading to overlapping data paths. His highest frequency of No-Responses to date (13) was exhibited when reinforcement was removed, and a relatively high frequency of No-Responses in a condition in which No-Responses was typically low (“Five Within-Session Choices” condition; nine), This suggested that the removal of reinforcement had an effect on this dependent variable, and that the videos still had moderate reinforcer efficacy. A “One Pre-Session Choice” condition was introduced to evaluate the effect of an experimental preparation with lower practitioner response effort on dependent variables. Data paths remained undifferentiated, highly variable, and had no apparent directional trend. Connor met termination criterion on sessions 18, 19, 20, and 22. Frequency of No-Responses ranged from 3-13 for sessions 17-24.

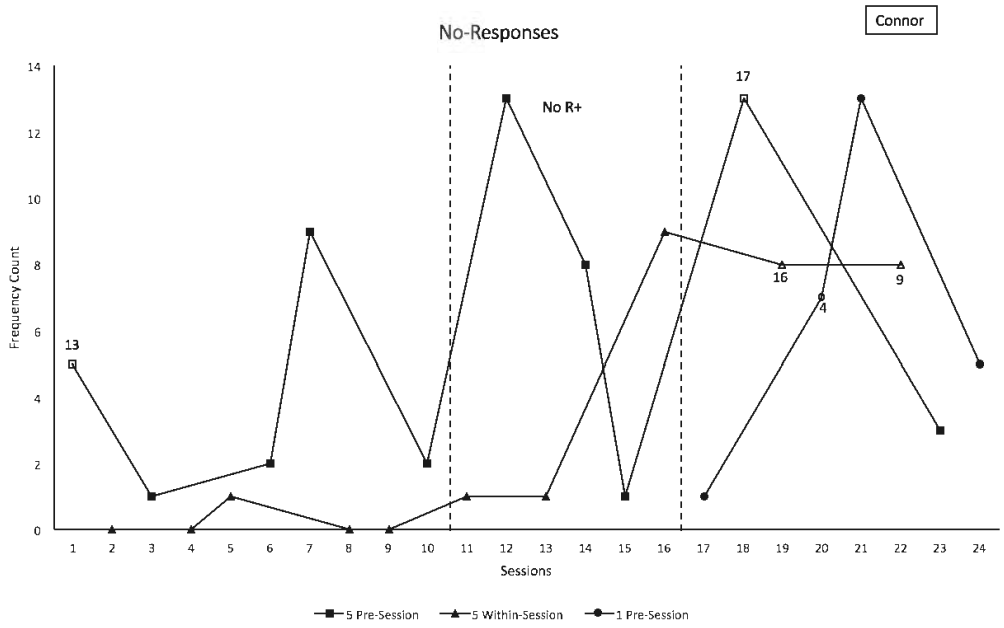

Figure 31: No-Responses data for Connor

Connor selected video, “Eggs” for 60 opportunities, the video, “Planets” for 19 opportunities, and the video, “Pete” for 21 opportunities. See Figure 33 for a graphical representation of Connor’s reinforcer selections.

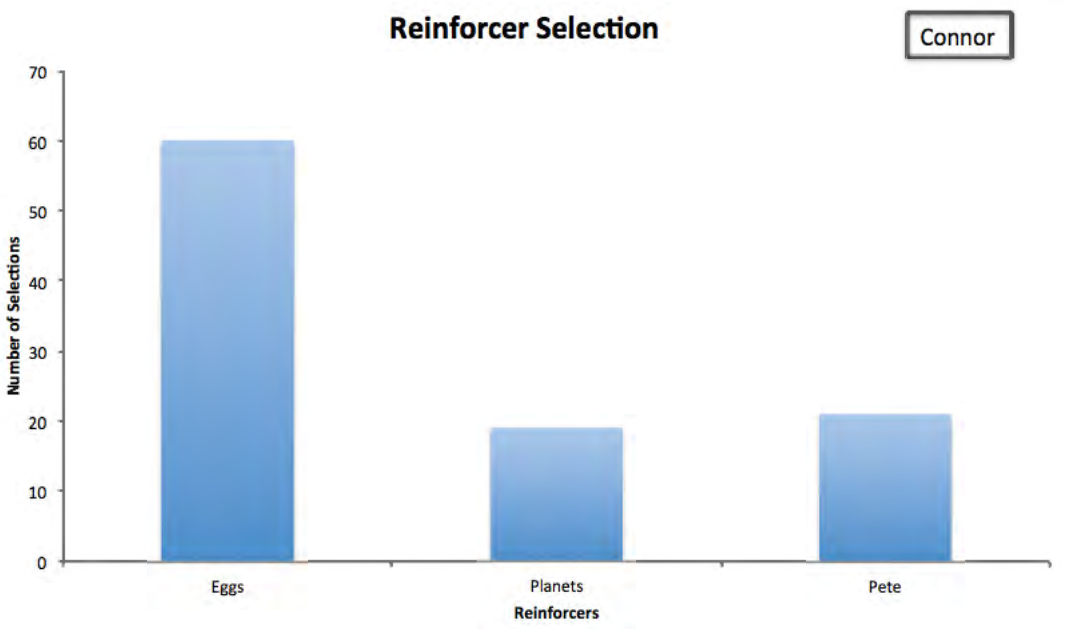

Figure 32: Reinforcer Selection data for Connor

Chapter 5: *Discussion*

The purpose of this experiment was to evaluate whether more frequent opportunities to select reinforcers through a variation of the MSW preference assessment format and response-contingent choice would impact motivation to respond on a mastered task for four participants with developmental disabilities. The rationale of altering the typical structure of an MSW preference assessment, such that choices were spaced evenly throughout a session, rather than clustered together prior to the session, came from the findings of DeLeon et al. (2001), which indicated that brief, but more frequent preference assessment variations led to greater response allocation than infrequent, extensive preference assessment variations. The authors concluded that participants periodically experienced shifts in their preferences, and that more frequent preference assessments were able to capture the most potent reinforcer for each participant on the specific day of the experimental session. The present experimental preparation compared two conditions, a Five Pre-Session Choices condition and a Five Within-Session Choices condition. Although each condition presented the same number of choices, the timing of choice presentations differed, such that the Five Pre-Session Choices preparation resulted in a single designated reinforcer for the duration of the session, while the Five Within-Session Choices preparation resulted in multiple opportunities for participants to select varied reinforcers throughout the session.

Previous research has indicated that repeated presentation of identical stimuli as reinforcers can lead to satiation, a phenomenon that results in lower motivation, often measured in terms of decreased responding (Berlyne, 1960; Fowler, 1971; Glanzer, 1958;

Myers and Miller, 1954). To counter this phenomenon, Egel (1980; 1981) varied reinforcer presentation and demonstrated renewed high levels of responding, the behavioral indicator of motivation in that experiment. The Five Within-Session Choices condition attempted to test this finding in terms of various parameters of responding on a mastered task. Specifically, it was evaluated whether the opportunity to access varied reinforcers through the MSW preference assessment format variation employed in this condition would impact the dependent variables of Latency to Task Initiation, Total Task Duration, Percentage Correct Responses, and No-Responses.

Recently, Elliott and Dillenberger (2014) found that providing a choice between highly preferred reinforcers contingent upon responses in a discrete trial training format, rather than no choice, increased motivation in the form of greater response accuracy and shorter response latency. They concluded that the presence of response-contingent choice opportunities might have created an establishing effect that positively impacted their dependent variables. Their use of a time-based dependent variable (latency), presented an intriguing metric for evaluating motivation, which was incorporated into this experiment in the form of the dependent variables of Latency to Task Initiation and Total Task Duration.

RESEARCH QUESTIONS

Research question 1 queried, "Will offering choices for reinforcers throughout a session lead to greater motivation in the form of lower Latency to Task Initiation and

Total Task Duration, higher Percentage Correct Responses, and lower frequency of No-Responses?"

In terms of the dependent variable of Latency to Task Initiation, James and Aimee demonstrated highly stable, undifferentiated Latency to Task Initiation data across all three conditions. This stability and lack of differentiation indicated that the format of the MSW preference assessment did not have an effect on how quickly these two participants began the task. If conceptualizing high motivation and reinforcer efficacy by low Latency to Task Initiation, practitioners could feel free to choose the preference assessment variation with the least practitioner response effort (the One Pre-Session Choice condition) for these two participants.

Differentiation among conditions would have reflected satiation with the designated reinforcer within a session (i.e., Latency to Task Initiation would have increased throughout the session, only in conditions in which stimulus variation was not allowed). An upward trend across all sessions would have reflected satiation with available reinforcers over time (i.e., Latency to Task Initiation was higher in later sessions, regardless of condition and opportunity to vary stimuli). When examining the Latency to Task Initiation graphs of James and Aimee, it appeared that neither satiation within sessions nor satiation across sessions occurred. In other words, the reinforcers each of these participants selected maintained their efficacy throughout both individual sessions and the duration of the experiment. By examining the bar graphs of their reinforcer selection, it becomes clear that each participant had a strong preference for one specific stimulus (popcorn for James and "Line Up" for Aimee). Practitioners working

with these participants could be confident that said reinforcers were both potent and durable, and it would be likely that preference assessments would not have to be frequent to be effective.

Ethan's Latency to Task Initiation graph demonstrates lack of differentiation across conditions, but increased variation over time. For sessions 15, 16, 17, and 18, Ethan met termination criterion, which indicated that his Latency to Task Initiation exceeded the upward limit of five seconds at least five times consecutively. Given the presence of termination criterion, a clear upward trend is not demonstrated, but it is apparent that Ethan's Latency to Task Initiation increased during sessions 15, 16, 17, and 18, and only stabilized once his mother began to sit in for sessions. Connor's Latency to Task Initiation graph demonstrated similar results, with all three conditions remaining undifferentiated throughout the experiment, but with an increase in sessions meeting termination criterion towards the end of the experiment. In contrast to James and Aimee, it appeared possible that Ethan and Connor experienced satiation with the available reinforcers across sessions, such that no single reinforcer maintained a stable Latency to Task Initiation without exceeding the upward limit. This finding suggested two possibilities: 1) the specific reinforcers selected were not especially durable and/or 2) Connor and Ethan's preferences shifted over time relatively quickly. In both cases, a practitioner working with either of these participants should consider engaging them in more frequent preference assessments and introducing novel stimuli, such that stimuli maintain motivation. However, caution should be exercised when interpreting these results as indicative of satiation, as the number of data points representing increased

Latency to Task Initiation was limited and other explanations for changes in this dependent variable are possible. The lack of differentiation demonstrated indicates that experimental control was not established and therefore, a clear functional relation was not demonstrated.

Results for the dependent variable of Total Task Duration were similar to results of Latency to Task Initiation across participants. James demonstrated very stable data without differentiation, with a downward trend (possibly due to a variation in his hand positioning). Aimee demonstrated stable data without differentiation, other than a single spike in the Five Within-Session Choices condition, which coincided with a session in which challenging behavior was high. Overall, this dependent variable indicates that the reinforcers selected by James and Aimee maintained comparable efficacy across conditions and across time. Connor's Total Task Duration graph is somewhat less clear due to the "No Reinforcement" condition imposed within, but an examination of his later sessions demonstrates similar Total Task Duration across conditions to sessions early in the experiment. Therefore, provided that reinforcement was offered, it appeared that the length of time it took Connor to complete the task was not affected either by choice presentation or the repeated use of the stimuli over time. In other words, this dependent variable did not clearly indicate that Connor experienced satiation with the stimuli over time, but practitioners should still exercise caution with repeated use of the same reinforcers, given Connor's Latency to Task Initiation data. Ethan's Total Task duration graph indicated a significant increase in duration for sessions 15 and 16. Termination criterion was met early in the session for sessions 17 and 18, so a significant increase in

Total Task Duration was less apparent, but would have likely occurred if all trials were conducted. These results suggested that Ethan would benefit from more frequent preference assessment.

In terms of the dependent variable of Percentage Correct Responding, James demonstrated stable and high accuracy with the mastered task across conditions and across time. Again, this dependent variable suggested that reinforcers maintained their potency and durability for James. Aimee's Percentage Correct Responding data was more variable, but still largely undifferentiated. Between sessions 6 and 13, Aimee demonstrated consistently lower accuracy in the Five Pre-Session Choices condition, but this difference disappeared thereafter. Anecdotally, Aimee often appeared to match incorrectly for attention, as she would laugh and make eye contact with the experimenter upon making an incorrect response. Therefore, Percentage Correct Responding may not be an ideal dependent variable for examining Aimee's motivation, as it provided an opportunity for Aimee to seek practitioner attention. Ethan's Percentage Correct Responding data was largely undifferentiated and became more variable over time, with some differentiation between sessions 12 and 22. Although the Five Pre-Session Choices condition demonstrated lower accuracy for all data points once the One Pre-Session Choice condition was introduced, there was some overlap and a data point for which termination criterion was met, so it was difficult to draw clear conclusions about whether this dependent variable indicated lower motivation during the Five Pre-Session Choices condition. Connor demonstrated stable, high, and undifferentiated accuracy when reinforcement was present, but some variability in the Five Within-Sessions Choices

condition when reinforcement was removed. It appears that, when reinforcing stimuli were provided, choice format did not impact Connor's Percentage Correct Responding.

In terms of the dependent variable of No-Responses, James demonstrated low frequency of No-Responses throughout the experiment, with some No-Responses exhibited in each condition. For sessions 13-22, No-Responses only occurred in the Five Pre-Session Choices condition. Caution should be exercised when drawing conclusions about this, however, because higher levels of No-Responses also coincided with parent-reported allergies, perhaps indicating a general lack of focus for those sessions. Aimee's No-Responses data was undifferentiated and variable, with No-Responses demonstrated in all conditions. A spike occurred in session 4, which coincided with high levels of challenging behavior. With the exception of that session, Aimee's affect while watching reinforcer videos was consistently positive, including laughing, smiling, and clapping. It was hypothesized that the occurrence of No-Responses may have been due more to distraction in a busy classroom than lack of motivation. The short duration of the No-Response criterion (5 seconds) made it difficult to differentiate between momentary distraction and low motivation to respond for the reinforcer. In future studies, it was recommended that a longer criterion for No-Responses be utilized, such that clear differences between momentary distraction and delays to respond can be captured. Ethan's No-Responses data showed a somewhat clearer picture of his motivation. Data paths were undifferentiated and stable early in the experiment, and remained undifferentiated, but became more variable towards the end of the experiment. Termination criterion, indicated by five or more consecutive No-Responses, was met in

sessions 15-18, with an experiment high of 11 exhibited in session 16. Once his mother began sitting in the room during sessions, No-Responses were no longer demonstrated, indicating a possible observer effect. Similar to the conclusions drawn about Latency to Task Initiation, it appeared that none of the available reinforcers were potent and durable enough to maintain quick and consistent responding over 22 sessions. For Connor, differentiation did occur for this dependent variable for the first 10 sessions of the experiment, indicating that he was less likely to respond within the time limit when his reinforcer selection was not varied (i.e., the Five Pre-Session Choices condition). After the introduction of the No Reinforcement condition, data paths became undifferentiated and total number of No-Responses increased across sessions in general. Similar to his Ethan, it appeared that the available reinforcers were no longer maintaining responding within the time limit as the experiment progressed. A recommendation drawn from the results of both Ethan and Connor would be to vary stimuli and perform frequent preference assessments. Once again, it is noted that caution should be exercised when hypothesizing whether or not Connor and Ethan's data is indicative of satiation, as experimental control was not established and the number of data points that demonstrated elevated frequency of No-Responses was limited.

Research question 2 queried, "Can time-based dependent variables (Latency to Task Initiation and Total Task Duration) effectively evaluate motivation in a DTT context?" The graphs of all participants across all dependent variables demonstrated lack of differentiation, therefore it is difficult to conclude whether Latency to Task Initiation and Total Task Duration were not effective metrics for evaluating motivation, or whether

the participants simply did not demonstrate clear patterns of motivation across conditions. The upper limit of Latency to Task Initiation (5 s) was problematic in this experiment, as it became difficult to distinguish between momentary distraction and a delay to respond as a result of low reinforcer efficacy. A more effective experimental preparation would arrange the task in such a way that motivation to begin the task could be more clearly conceptualized. For example, if a participant were seated in a chair, and the task were arranged on a table some distance away, Latency to Task Initiation could be measured in terms of the time elapsed between the delivery of the discriminative stimulus and the initiation of the task, which, in this case, would involve the participant standing up and moving his or her body to the task area. Participants who refused to stand up could more clearly be understood to be delaying responding to the task.

Total Task Duration represents a slightly more promising dependent variable to evaluate motivation, from a different perspective than was anticipated. Hypothesis 2 stated that time-based dependent variables (Latency to Task Initiation and Total Task Duration) would effectively evaluate motivating operations in a DTT context, as demonstrated by differentiation between two distinct conditions (i.e., Five Pre-Session Choices condition and Five Within-Session Choices condition). Lack of differentiation indicates that experimental control was not established and that this research question cannot be definitively answered. However, stability of Total Task Duration across sessions for James and Aimee may indicate durability of reinforcer efficacy in general, and increases in duration across sessions for Ethan and Connor may indicate satiation with reinforcers over time. In other words, for James and Aimee, the repeated

presentation of a choice of three reinforcers over time did not have a significant impact on how long it took to complete a mastered task, either within individual sessions or across multiple sessions, but for Ethan and Connor, as reinforcers were presented repeatedly across sessions, the amount of time it took these participants to complete the task increased across multiple data points. One could tentatively conclude that the length of time it took each participant to complete the task reflected how motivated he or she was to obtain the reinforcer, thereby indicating reinforcer efficacy and motivation. On the other hand, although it is possible that the trends of Total Task Duration across these four participants are a clear indication of reinforcer efficacy and motivation, they may also be a function of the setting in which the experiments occurred. James and Aimee, for whom Total Task Duration was stable and undifferentiated across conditions, took part in this experiment at their school, in the presence of their typical lead teacher. It was possible that the presence of the teacher, and the familiar setting, provided a degree of instructional control to the experimenter that was not present in the home setting for Connor and Ethan. James and Aimee may have responded quickly and consistently because of the expectations of their environment and the stimulus control of their teacher's presence. Once again, the lack of experimental control makes it difficult to draw firm conclusions about this research question.

Research question 3 queried, "If differentiation between conditions in a multi-element design does not occur, will an even briefer preference assessment format produce comparable results, so as to reduce response effort for practitioners?" For all four participants, the One Pre-Session Choice condition produced similar results across

dependent variables, demonstrating continued lack of differentiation. Therefore, within the parameters of this experimental preparation (i.e., 21 trials of the specific mastered task(s), it did not appear to matter whether the participant was provided with five choices of stimuli for which to work, or only one. Working within similar DTT conditions, and with comparably complex tasks, practitioners should be able to utilize the preference assessment method with lower response effort (i.e., the One Pre-Session Choice condition).

LIMITATIONS

As is the case with any study, this experiment is not without limitations. First, only two stimulus classes of reinforcers (edibles and videos) were utilized. Ideally, an experiment evaluating the effect of choice timing on reinforcer efficacy would evaluate a variety of stimulus classes of reinforcers, including social reinforcement, tactile reinforcement, and other tangible reinforcement (e.g., toys). In this experiment, the RAISD administered to relevant adults revealed that videos were the most potent and varied class of reinforcers for three of the four participants. Given the inclusion requirement that each participant demonstrate preference for at least five reinforcers of the same stimulus class, videos were chosen for Aimee, Ethan, and Connor, even though a more balanced experiment would have included participants with more varied preferences. Second, two participants engaged in the experiment in the home setting and two participants engaged in the experiment in the school setting, producing the potential for differences due to environment, such as the stimulus control of the teacher's presence

in the room for James and Aimee. Although a variety of settings in an experiment can improve the generalizability of findings, given the undifferentiated nature of the results for all participants, the differences in settings may represent a limitation in terms of drawing conclusions. Third, the time limit chosen for Latency to Task Initiation and No-Responses (5 s) appears to have been too restrictive, producing difficulty in differentiating between momentary distraction and delay to responding due to low motivation. If this experiment were to be replicated or extended, a longer Latency to Task Initiation and No-Response criterion should be selected. Finally, the experimental preparation called for 21 trials (one trial for reinforcer selection and 20 trials for which data was taken) of a simple mastered task. This arrangement produced a session duration that ranged from approximately eight minutes (James) to approximately 30 minutes (Aimee, Ethan, and Connor), based on the stimulus class utilized for reinforcement. Given that satiation with a reinforcer within a session was being examined as a function of choice presentation, these durations may not have been sufficiently long enough to demonstrate satiation within a single session.

FUTURE RESEARCH

This experiment was intended to be an extension of the research by DeLeon et al. (2001) and Elliott and Dillenberger (2011). As such, it was arranged to evaluate the effect of MSW variations that provided either clustered choices and no reinforcer variation or evenly spaced choices, response-contingent choice, and the opportunity for reinforcer variation. This experiment should be replicated with some alterations to determine

whether lack of differentiation was due to the brevity of sessions, preference bias in reinforcers (i.e., one reinforcer was potent enough to be selected almost exclusively, regardless of choice opportunities), or whether choice presentation variation simply had no effect on motivation. It is advised that replications utilize either a larger number of trials or more complex tasks, so as to increase the duration of sessions and elucidate the possible effect of satiation. In addition, it is advised that the criterion used to establish a No-Response be increased so as to more clearly distinguish between momentary distraction and delay to respond due to low motivation. Furthermore, a more naturalistic instructional context (e.g., varied tasks across a longer session) could be valuable in increasing the generalizability of results. Finally, future research should seek to include more varied stimulus classes for reinforcement options, such as social, tactile, and other tangible reinforcers, as it is possible that the potential for satiation varies across stimulus classes.

CONCLUSIONS

A general lack of differentiation of data paths in this experiment makes it difficult to draw conclusions about the effects of choice timing on motivation, due to a failure to establish experimental control. However, changes in dependent variables over time for Ethan and Connor suggest that for some participants, frequent preference assessment and stimulus variation were important for maintaining motivation in the form of low Latency to Task Initiation and Total Task Duration, low frequency of No-Responses, and high accuracy. For other participants, such as James and Aimee, preference bias for a specific

reinforcer were strong enough to allow said stimulus to be utilized repeatedly across many sessions. Practitioners should be cognizant of the behavioral trends of the individuals with whom they work, so that they can recognize possible satiation with reinforcers, as evident by changes in the dependent variables examined in this experiment, and adjust treatment accordingly.

Appendix

Citation		
Assesses (circle one)	**Choice**	**Preference**
Number of Participants		
Ages of Participants		
Diagnoses of Participants		
Stimuli used in preference assessment		
Type of preference assessment		
Dependent Variables		
Experimental Design		
Findings (Circle)	**Positive**	**Mixed**
Describe Findings		

Figure 33: Coding sheet for literature review

Table 5 *Coding sheet for Interobserver agreement and Treatment Fidelity*

E displayed 3 choices	**Choice**								
Y N									
E displayed 3 choices	**Choice**								
Y N									
E displayed 3 choices	**Choice**								
Y N									
E displayed 3 choices	**Choice**								
Y N									
E displayed 3 choices	**Choice**								
Y N									
Trial 0	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 1	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 2	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**

Trial 3	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 4	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 5	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 6	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 7	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**

Trial 8	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 9	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 10	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 11	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 12	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**

Trial 13	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 14	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 15	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 16	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 17	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**

Trial 18	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 19	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**
Trial 20	**E placed pieces in front of P**	**E stated SD before P touched pieces**	**E did not prompt other than tapping paper every 10 seconds**	**If latency was longer than 5 seconds, said, "Oops, let's try again", removed pieces and restored pieces**	**If P made a matching mistake, E corrected (displayed correct answer) and restored pieces**	**When P terminated task, E said, "That's it" in a neutral voice**	**E presented R+ for approximately 30 seconds or until edible was consumed**	**Latency to initiation (time elapsed from termination of SD to initiation of the task in seconds)**	**Total task duration (time elapsed from initiation to termination of the task in seconds)**

Table 5: Fidelity and Interobserver Agreement document (Five Pre-Session Choices condition). Other condition documents were identical except for spacing of reinforcer selections.

Table 6

Complete IOA and Fidelity Scores for all participants

			James		
	Fidelity	IOA Latency	IOA Duration	IOA Accuracy	IOA No-Responses
J_C1_D6	100.0	85.6	99.5	100.0	100.0
J_C1_D7	100.0	87.9	98.3	100.0	100.0
J_C2_D10	96.0	73.0	94.6	100.0	100.0
J_C2_D12	98.3	84.4	96.5	95.0	100.0
J_C2_20	99.1	66.0	96.6	100.0	100.0
J_C3_D14	99.0	63.0	96.7	100.0	100.0
J_C2_D15	100.0	38.0	94.6	95.0	100.0
J_C3_D16	100.0	49.6	96.2	100.0	100.0
Mean:	**99.1**	**68.4**	**97.3**	**98.8**	**100.0**
			Aimee		
	Fidelity	IOA Latency	IOA Duration	IOA Accuracy	IOA No-Responses
A_C2_D5	96.4	84.3	85.3	90.0	100.0
A_C2_D6	100.0	79.0	95.0	100.0	100.0
A_C1_D8	99.0	75.4	96.0	100.0	100.0
A_C2_D13	100.0	93.9	96.0	100.0	100.0
A_C1_D14	100.0	94.5	97.0	100.0	100.0
A_C3_D16	99.0	89.0	96.0	100.0	100.0
A_C3_D18	100.0	97.5	94.9	100.0	100.0
Mean:	**99.3**	**89.1**	**94.5**	**98.8**	**100.0**
			Ethan		
	Fidelity	IOA Latency	IOA Duration	IOA Accuracy	IOA No-Responses
E_C1_D8	98.6	81.4	97.6	100.0	100.0
E_C1_D9	100.0	81.3	99.0	100.0	100.0
E_C1_D12	100.0	70.0	90.0	96.0	100.0
E_C2_D13	100.0	82.4	98.6	100.0	100.0
E_C3_D14	98.6	73.0	99.3	100.0	100.0
E_C1_D16	100.0	98.8	100.0	100.0	100.0
E_C3_D19	100.0	83.0	85.0	100.0	100.0
E_C2_D20	99.1	93.0	98.6	100.0	100.0
Mean:	**99.3**	**81.9**	**98.1**	**98.8**	**100.0**
			Connor		
	Fidelity	IOA Latency	IOA Latency	IOA Duration	IOA No-Responses
C_C1_D1	100.0	96.4	97.3	100.0	100.0
C_C2_D2	99.0	42.4	97.0	100.0	100.0
C_C1_D3	100.0	73.1	93.8	100.0	100.0
C_C2_D9	100.0	74.6	99.0	100.0	100.0
C_C1_D12	100.0	90.9	99.7	100.0	100.0
C_C2_D13	100.0	81.7	99.6	90.0	100.0
C_C3_D17	100.0	92.1	100.0	100.0	100.0
C_C3_D21	99.1	93.0	98.6	100.0	100.0
Mean:	**99.8**	**80.5**	**98.1**	**98.8**	

References

Ahearn, W.H., Clark, K.M., DeBar, R., & Florentino, C. (2005). On the role of preference in response competition. *Journal of Applied Behavior Analysis, 38,* 247-250. doi: 10.1901/jaba.2005.36-04

Baer, R.A., Tishelman, A.C., Degler, J.D., Osnes, P.G., & Stokes, T.F. (1992). Effects of self- vs. experimenter-selection of rewards on classroom behavior in young children. *Education and Treatment of Children, 15,* 1-14.

Bambara, L.M., Koger, F., Katzer, T., & Davenport, T.A. (1995). Embedding choice in the context of daily routines: An experimental case study. *Journal of the Association for Persons with Severe Handicaps, 20,* 185-195.

Berlyne, D.E. *Conflict, arousal, and curiosity.* New York: McGraw-Hill, 1960.

Bowers, D.S., Clement, P.W., Fantuzzo, J.W., & Sorensen, D.A. (1985). Effects of teacher-administered and self-administered reinforcers on learning disabled children. *Behavior Therapy, 16,* 357-369.

*Call, N.A., Trosclair-Lasserre, N.M., Findley, A.J., Reavis, A.R., & Shilingsburg, M.A. (2012). Correspondence between single versus daily preference assessment outcomes and reinforcer efficacy under progressive-ratio schedules. *Journal of Applied Behavior Analysis, 45,* 763-777. doi: 10.1901/jaba.2012.45-763.

*Call, N.A., Shillingsburg, M.A., Bowen, C.N., Reavis, A.R., & Findley, A.J. (2013). Direct assessment of preferences for social interactions in children with autism. *Journal of Applied Behavior Analysis, 46,* 821-826. doi: 10.1002/jaba.69

Cannella, H.I., O'Reilly, M.F., & Lancioni, G.E. (2004). Choice and preference assessment research with people with severe to profound developmental disabilities: A review of the literature. *Research in Developmental Disabilities, 26,* 1-15. doi:10.1016/j.ridd.2004.01.006

*Clay, C.J., Samaha, A.L., Bloom, S.E., Bogoev, B.K., & Boyle, M.A. (2013). Assessing preference for social interactions. *Research in Developmental Disabilities, 34,* 362-371. doi: 10.1016/j.ridd.2012.07.028

Cole, C.L., & Levinson, T.R. (2002). Effects of within-activity choices on the challenging behavior of children with severe developmental disabilities. *Journal of Positive Behavior Interventions, 4,* 29-37. doi: 10.1177/109830070200400106

Conyers, C., Doole, A., Vause, T., Harapiak, S., Yu, D.C.T., & Martin, G.L. (2002). Predicting the relative efficacy of three presentation methods for assessing preferences of persons with developmental disabilities. *Journal of Applied Behavior Analysis, 35,* 49-58. doi:10.1901/jaba.2002.35-49

*Davies, G., Chand, C., Yu, C.T., Martin, T.L., & Martin, G.L. (2013). Evaluation of multiple-stimulus preference assessment with adults with developmental disabilities. *Education and Training in Autism and Developmental Disabilities, 48,* 269-275.

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

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

DeMarse, T.B., Killeen, P.R., & Baker, D. (1999). Satiation, capacity, and within-session responding. *Journal of the Experimental Analysis of Behavior, 72,* 407-423.

Didden, R., Korzilius, H., Sturmey, P., Lancioni, G., & Curfs, L.M.G. (2008). Preference for water-related items in Angelman syndrome, Down syndrome, and non-specific intellectual disability. *Journal of Intellectual and Developmental Disability, 33,* 59-64. doi**:**10.1080/13668250701872126

Dunlap, G., DePerczel, M., Clarke, S., Wilson, D., Wright, S., White, R., & Gomez, A. (1994). Choice making to promote adaptive behavior for students with emotional and behavioral challenges. *Journal of Applied Behavior Analysis, 27,* 505-518.

Dunlap, G., Kern-Dunlap, L., Clarke, S., & Robbins, F.R. (1991). Functional assessment, curricular revision, and severe behavior problems. *Journal of Applied Behavior Analysis, 24,* 387-397.

Dyer, K. (1987). The competition of autistic stereotyped behavior with usual and specially assessed reinforcers. *Research in Developmental Disabilities, 8,* 607-626.

Dyer, K., Dunlap, G., & Winterling, V. (1990). Effects of choice making on the serious problem behaviors of students with severe handicaps. *Journal of Applied Behavior Analysis, 23,* 515-524.

Egel, A.L. (1980). The effects of constant vs. varied reinforcer presentation on responding by autistic children. *Journal of Experimental Child Psychology, 30,* 455-463.

Egel, A.L. (1981). Reinforcer variation: Implications for motivating developmentally disabled children. *Journal of Applied Behavior Analysis, 14,* 345-350

*Elliott, C., & Dillenburger, K. (2014). The effect of choice on motivation for young children on the autism spectrum during discrete trial teaching. *Journal of Research in Special Education.* doi: 10.1111/1471-3802.12073

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

Fisher, W., Piazza, C.C., Bowman, L.G., Hagopian, L.P., Owens, J.C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491-498.

Fisher, W., Thompson, R., Piazza, C., Crosland, K., & Gotjen, D. (1997). On the relative reinforcing effects of choice and differential consequences. *Journal of Applied Behavior Analysis, 30,* 423-438.

Fowler, H. Implications of sensory reinforcement. In R.H. Glaser (Ed.), *The nature of reinforcement.* New York: Academic Press, 1971.

Glanzer, M. (1958). Curiosity, exploratory drive, and stimulus satiation. *Psychological Bulletin, 56,* 302-315.

*Graff, R.B., & Larsen, J. (2011). The relation between obtained preference value and reinforcer potency. *Behavioral Interventions, 26,* 125-133. doi: 10.1002/bin.325

*Hall, J., Morgan, R.L., Salzberg, C.L. (2014). Job-preference and job-matching assessment results and their association with job performance and satisfaction among young adults

with developmental disabilities. *Education and Training I Autism and Developmental Disabilities, 49,* 301-312.

*Kang, S., O'Reilly, M.R., Fragale, C., Aguilar, J.M., Rispoli, M., & Lang, R. (2011). Evaluation of the rate of problem behavior maintained by different reinforcers across preference assessments. *Journal of Applied Behavior Analysis, 44,* 835-846. doi: 10.1901/jaba.2011.44-835

*Karsten, A.M., Carr, J.E., & Lepper, T.L. (2011). Description of a practitioner model for identifying preferred stimuli with individuals with autisms spectrum disorders. *Behavior Modification, 35,* 347-369. doi: 10.1177/0145445511405184

Kearney, C.A., & McKnight, T. (1997). Preference, choice, and persons with disabilities: A synopsis of assessments, interventions, and future directions. *Clinical Psychology Review, 17,* 217-238.

*Kelly, M.A., Roscoe, E.M., Hanley, G.P., & Schlichenmeyer, K. (2014). Evaluation of assessment methods for identifying social reinforcers. *Journal of Applied Behavior Analysis, 47,* 113-135. doi: 10.1002/jaba.107

Kennedy, C. H. (2005). Single-case designs for educational research. Boston, MA: Allyn and Bacon.

Kennedy, C.H., & Haring, T.G. (1993). Teaching choice making during social interactions to students with profound multiple disabilities. *Journal of Applied Behavior Analysis, 26,* 63-76.

*Kenzer, A.L., & Bishop, M.R. (2011). Evaluating preference for familiar and novel stimuli across a large group of children with autism. *Research in Autism Spectrum Disorders, 5,* 819-825. doi:10.1016/j.rasd.2010.09.011

*Kenzer, A.L., Bishop, M.R., Wilke, A.E., & Tarbox, J.R. (2013). Including unfamiliar stimuli in preference assessments for young children with autism. *Journal of Applied Behavior Analysis, 46,* 689-694. doi: 10.1002/jaba.56

*Keyl-Austin, A.A., Samaha, A.L., Bloom, S.E., & Boyle, M.A. (2012). Effects of preference and reinforcer variation on within-session patterns of responding. *Journal of Applied Behavior Analysis, 45,* 657-641. doi: 10.1901/jaba.2012.45-637

*Lanovaz, M.J., Rapp, J.T., Ferguson, S. (2012). The utility of assessing musical preference before implementation of noncontingent music to reduce vocal stereotypy. *Journal of Applied Behavior Analysis, 45,* 845-851. doi: 10.1901/jaba.2012.45-845

Lancioni, G.E., Bellini, D., & Oliva, D. (1993). Building choice opportunities within a robot-assisted occupational program: A case study. *Behavioral Residential Treatment, 8,* 219-226.

Lancioni, G.E., O'Reilly, M.F., & Emerson, E. (1996). A review of choice research with people with severe and profound developmental disabilities. *Research in Developmental Disabilities, 17,* 391-411.

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36,* 407-414.

Lim, L., Browder, D.M., & Bambara, L. (2001). Effects of sampling opportunities on preference development for adults with severe disabilities. *Education and Training in Mental Retardation and Developmental Disabilities, 36,* 188-195.

*Lough, C.L., Rice, M.S., & Lough, L.G. (2012). Choice as a strategy to enhance engagement in a colouring task in children with autism spectrum disorders. *Occupational Therapy International, 19,* 204-211. doi: 10.1002/oti.1337

Machaliecek, W., O'Reilly, M., Chan, J.M., Rispoli, M., Lang, R., Davis, T., & Langthorne, P. (2009). Using videoconferencing to support teachers to conduct preference assessments with students with autism and developmental disabilities. *Research in Autism Spectrum Disorders, 3,* 32-41. doi:10.1016/j.rasd.2008.03.004

*Mangum, A., Frederick, L., Pabico, R., Roane, H. (2012). The role of context in the evaluation of reinforcer efficacy: Implications for preference assessment outcomes. *Research in Autism Spectrum Disorders, 6,* 158-167. doi:10.1016/j.rasd.2011.04.001

Matson, J.L., Bamburg, J.W., Cherry, K.E., & Patlawskij, R.R. (1999). A validity study on the Questions About Behavioral Function (QABF) scale: Predicting success for self-injury, aggression, and stereotypies. *Research in Developmental Disabilities, 20,* 163-176.

Matson, J.L., Bielecki, J., Mayville, E.A., Smalls, Y., Bamburg, J.W., & Baglio, C.S. (1999). The development of a reinforcer choice assessment scale for persons with severe and profound mental retardation. *Research in Developmental Disabilities, 20,* 379-384.

Mithaug, D.E., & Hanawalt, D.A. (1978). The validation of procedures to assess prevocational task preferences in retarded adults. *Journal of Applied Behavior Analysis, 11,* 153-162.

Mithaug, D.E., & Mar, D.K. (1980). The relation between choosing and working prevocational tasks in two severely retarded young adults. *Journal of Applied Behavior Analysis, 13,* 177-182.

Myers, A.K., & Miller, N.E. (1954). Failure to find a learned drive based on hunger: Evidence for learning motivated by "exploration". *Journal of Comparative and Physiological Psychology,* 47, 428-436.

Newton, J.S., Ard, W.R., & Horner, R.H. (1993). Validating predicted activity preferences of individuals with severe disabilities. *Journal of Applied Behavior Analysis, 26,* 239-245.

*Nuernberger, J.E., Smith, C.A., Czapar, K.N., & Klatt, K.P. (2012). Assessing preference for social interaction in children diagnosed with autism. *Behavioral Interventions, 27,* 33-44. doi: 10.1002/bin.1336

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

*Phillips, K.J., & Mudford, O.C. (2011). Effects of noncontingent reinforcement and choice of activity on aggressive behavior maintained by attention. *Behavioral Interventions, 26,* 147-160. doi: 10.1002/bin.32

Reid, D.H., & Parsons, M.B. (1991). Making choice a routine part of mealtimes for persons with profound mental retardation. *Behavioural Residential Treatment, 6,* 591-603.

*Rispoli, M., Lang, R., Neely, L., Camargo, S., Hutchins, N., Davenport, K., & Goodwyn, F. (2013). A comparison of within- and across- activity choices for reducing challenging

behavior in children with autism spectrum disorders. *Journal of Behavioral Education, 22,* 66-83. doi: 10.1007/s10864-012-9164-y

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

Romaniuk, C., & Miltenberger, R.G. (2001). The influence of preference and choice of activity on problem behavior. *Journal of Positive Behavior Interventions, 3,* 152-159. doi: 10.1177/109830070100300303

Roscoe, E.M., Fisher, W.W., Golver, A.C., & Volkert, V.M. (2006). Evaluating he relative effects of feedback and contingent money for staff training of stimulus preference assessments. *Journal of Applied Behavior Analysis, 39,* 63-77. doi: 10.1901/jaba.2006.7-05

*Russo, S.R., Tincani, M., & Axelrod, S. (2014). Evaluating open-ended parent reports and direct preference assessments to identify reinforcers for young children with autism. *Child and Family Behavior Therapy, 36,* 107-120. doi: 10.1080/07317107.2014.910732

*Sellers, T.P., Bloom, S.E. Samaha, A.L., Dayton, E., Lambert, J.M., & Keyl-Austin, A.A. (2013). Evaluation of some components of choice making. *Journal of Applied Behavior Analysis, 46,* 455-464. doi: 10.1002/jaba.46

Seybert, S., Dunlap, G., & Ferro, J. (1996). The effects of choice making on the problem behaviors of high school students with intellectual disabilities. *Journal of Behavioral Education, 6,* 49-65.

Smaby, K., MacDonald, R.P.F., Ahearn, W.H., & Dube, W.V. (2007). Assessment protocol for identifying preferred social consequences. *Behavioral Interventions, 22,* 311-318. doi: 10.1002/bin.242

*Snyder, K., Higbee, T.S., & Dayton, E. (2012). Preliminary investigation of a video-based stimulus preference assessment. *Journal of Applied Behavior Analysis, 45,* 413-418. doi: 10.1901/jaba.2012.45-413

Stancliffe, R.J. (1995). Assessing opportunities for choice-making: A comparison of self- and staff-reports. *American Journal on Mental Retardation, 99,* 418-429.

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

Tullis, C.A., Cannella-Malone, H.I., & Fleming, C.V. (2012). Multiple stimulus without replacement preference assessments: An examination of the relation between session number and effectiveness. *Journal of Developmental and Physical Disabilities, 24,* 337-345. doi: 10.1007/s10882-012-9273-5

*Verschuur, R., Didden, R., Van Der Meer, L., Achmadi, D., Kagohara, D., Green, V.A., Lang, R., & Lancioni, G.E. (2011). Investigating the validity of a structured interview protocol for assessing the preferences of children with autism spectrum disorder. *Developmental Neurorehabilitation, 14,* 366-371. doi: 10.3109/17518423.2011.606509

*Weldy, C.R., Rapp, J.T., & Capocasa, K. (2014). Training staff to implement brief stimulus preference assessments. *Journal of Applied Behavior Analysis, 47,* 214-218. doi: 10.1002/jaba.98

Windsor, J., Piche, L.M., & Locke, P.A. (1994). Preference testing: A comparison of two presentation methods. *Research in developmental Disabilities, 15,* 439-455.

Zhou, L., Iwata, B.A., Goff, G.A., & Shore, B.A. (2001). Longitudinal analysis of leisure-item preferences. *Journal of Applied Behavior Analysis, 34,* 179-184. doi: 10.1901/jaba.2001.34-179