COMPARISON OF PARAMATER HEIRARCHY ASSESSMENTS WITH ARBITARY AND CLINICALLY RELEVANT BEHAVIOR

by

ROSE M. MORLINO

(Under the Direction of Joel Ringdahl)

ABSTRACT

Although extinction is a common and effective treatment component included in behavior reduction procedures, adverse side effects may make its inclusion impractical to implement. Treatments that include concurrent reinforcement schedule arrangements may provide an alternative to the inclusion of extinction. A recent review by Trump et al. (2020) found that concurrent schedules without extinction interventions were successful in reducing challenging behavior, but the results of the interventions were idiosyncratic. Kunnavatana et al. (2018) assessed individual and relative sensitivity to the parameters of reinforcement used in concurrent schedule-based interventions to determine a parameter hierarchy prior to developing an intervention for challenging behavior. However, they used arbitrary behaviors during the assessment. Thus, the current study compared the results of a parameter hierarchy assessment with arbitrary behaviors to a parameter hierarchy assessment with clinically relevant behaviors to determine if the results would align. Results show no alignment between the two assessments.

INDEX WORDS: Concurrent Schedules, Reinforcement Parameters, Parameter Sensitivity

COMPARISON OF PARAMATER HEIRARCHY ASSESSMENTS WITH ARBITARY AND CLINICALLY RELEVANT BEHAVIOR

by

ROSE M. MORLINO

B.S., James Madison University, 2016

M.S., University of Georgia, 2018

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2023

COMPARISON OF PARAMATER HEIRARCHY ASSESSMENTS WITH ARBITARY AND CLINICALLY RELEVANT BEHAVIOR

by

ROSE M. MORLINO

Major Professor: Joel E. Ringdahl
Committee: Kevin M. Ayres
Rachel R. Cagliani
Scott P. Ardoin

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2023

DEDICATION

I dedicate this dissertation to my parents and my fiancé. Mom and Dad, I would not be who I am today without your love and guidance. Thank you for your endless support, love, and advice that made this degree and dissertation possible. Chris, you have been by my side throughout this entire process. There is no way I could've completed this dissertation without your encouragement, love, and comedic relief. Thank you for believing in me. I love you all!

ACKNOWLEDGEMENTS

Anyone who has pursued a PhD knows this journey is not possible without the abundant support of faculty, colleagues, and friends. First, I would like to thank my advisor, Dr. Joel Ringdahl, for his mentorship, advice, and most importantly, dad jokes throughout my academic career. I would also like to thank the rest of my committee, Dr. Rachel Cagliani, Dr. Kevin Ayres, and Dr. Scott Ardoin, for their encouragement and support with this project and degree.

Next, I would like to acknowledge all the wonderful CABER staff members who helped me along the way. Thank you to all the CRES, CA, and Clinic staff for your endless encouragement and laughter throughout my master's and doctoral program. No one else will understand the bond we share and the love we have for the students and clients we serve. I cannot wait to see the amazing impact you will all have on this world.

Lastly, I would not be the person I am today without the support of the trusted colleagues I get to call friends. To my soulmate, Dr. Jessica Herrod, there are not enough words to express my gratitude for your support throughout my academic career and personal life. The world is infinitely better with you in it, and I am eternally blessed to have you in my life. Kavya Kandarpa, thank you for always answering my late-night spiral phone calls and supporting all my life decisions. Holly Long, thank you for cosigning all of my bad decisions and for being a shoulder to lean on. Kadijah Quinland, thank you for being a safe space for all my intrusive thoughts and for being such a wonderful and supportive friend. Dr. Karla Zabala-Snow, thank you for the wonderful memories and laughter that kept me sane. Who knew that a PhD would lead to lifelong friendship with such amazing people, I love you all so much!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Individuals with intellectual and developmental disabilities are more likely to engage in challenging behavior than children with other disabilities and children without disabilities (Pope et al., 2010). Emerson (2001) defined challenging behavior as behaviors that endanger the safety of the individual or those around them, or behaviors that can deny an individual access to their community. Examples of challenging behaviors may include aggression, self-injury, or property destruction (Pope et al., 2010). With the potential for harm and loss of community access, it is important to develop effective interventions that decrease challenging behavior. One common intervention known to reduce challenging behavior is differential reinforcement of alternative behavior (DRA). This procedure typically involves withholding reinforcement for challenging behavior (i.e., extinction) and providing reinforcement for a specified alternative behavior (Cooper et al., 2019).

Many examples of DRA procedures can be found in the literature. For example, Carr and Durand (1985) used a DRA procedure to reduce challenging behavior exhibited by four school-aged children. Researchers taught participants to request help or attention, while actively ignoring challenging behavior (i.e., challenging behavior produced no programmed consequence). Results indicated that reinforcing a functionally equivalent communicative response while withholding consequences for challenging behavior reduced challenging behavior and increased appropriate behavior. Petscher et al. (2009) conducted a literature review on DRA-based intervention for challenging behavior published between 1977 and 2007. Researchers reviewed 116 articles that evaluated DRA procedures. Researchers coded whether the included articles provide empirical support for the use of DRA procedures to treat various behavior topographies and behavior functions using the Task Force criteria for single subject design (Task Force, 1995). The Task Force criteria specify that at least 10 single-subject studies must use a good experimental design, compare the treatment to another treatment, conduct the study using a treatment manual, clearly specify participant characteristics, and demonstrate treatment effectiveness (Task Force, 1995) Of the 116 articles included, 40 met the Task Force criteria. These 40 articles provide empirical support for the use of DRA to treat challenging behavior (i.e., aggression, self-injury, and property destruction) and food refusal. According to the Petscher et al., many of the remaining 76 articles reported successful decreases of challenging behavior using DRA procedures; however, the designs of the studies did not meet the Task Force criteria for inclusion in the empirical support analysis of the review. This demonstrates that DRA is an effective intervention for the treatment of challenging behavior.

Although DRA has been demonstrated to be a successful intervention approach, the extinction component is not always possible to implement. Extinction procedures can produce undesirable side effects and therefore inclusion of an extinction component in an intervention is not always feasible (Cooper et al., 2019). These undesirable side effects include initial increases in challenging behavior and the emergence of other challenging behaviors (Lerman et al., 1999). In addition to undesirable side effects, extinction is not always implemented with fidelity, which may result in intermittent reinforcement of challenging behavior (Vollmer et al., 2020). Behaviors such as aggression and self-injury that are maintained by attention cannot always be ignored due to the severity of the behavior (Vollmer et al., 2020). This lapse in procedural fidelity can hinder the effectiveness of DRA procedures (St. Peter Pipkin et al., 2010). Due to these issues, it is important to evaluate effective alternative intervention approaches that can be utilized when extinction is not possible. One alternative intervention approach involves utilizing concurrent schedules of reinforcement, without an extinction component.

Concurrent Schedule-Based Interventions

Concurrent schedules of reinforcement are multiple component schedules that include more than one simultaneously available schedule of reinforcement (Cooper et al., 2019). When developing interventions using concurrent schedules, researchers and practitioners must understand the variables that impact response allocation to different reinforcement schedules. The matching law (Herrnstein, 1961) quantifies response allocation under concurrent schedule arrangements. Matching law was first described by Herrnstein (1961) using the equation $\frac{B_1}{B_1+B_2} = \frac{R_1}{R_1+R_2}$. In this equation, B represents the rate of responding (the subscripts denote the two behavior alternatives) and R represents the obtained reinforcement rate. Herrnstein explains that proportional response rate is proportional to obtained reinforcement rate. That is, in a concurrent schedule arrangement, the relative rate of response allocation will match the relative rate of reinforcement. This does not mean that individuals only respond to the better reinforcement schedule, but instead responding is distributed between the schedules based on the reinforcement available (Cuvo et al., 1998). Although matching law provides a good quantification of responding in concurrent schedule arrangements, it does not always yield precise outcomes Baum (1974) explains that bias (e.g., preference) affects responding during concurrent reinforcement schedules. Bias is not accounted for by the equation described by Herrnstein (1961). To address this imprecision, Baum (1974) proposed a new equation using natural logarithms, the generalized matching equation: $\log(\frac{B_1}{B_2}) = a\log(\frac{R_1}{R_2}) + \log b$. In this equation, B represents response frequency, R represents reinforcement frequency, a represents the slope, and *b* represents the intercept. Researchers and practitioners can use the general relations predicted by matching law when developing concurrent schedule arrangements in the treatment of challenging behavior. To do so, researchers and practitioners must understand the different variables that impact response allocation.

Reinforcement Manipulations

One way to guide response allocation in the desired direction is to manipulate aspects of the reinforcement associated with each component of the concurrent schedule. Several reinforcement aspects have been demonstrated to impact response allocation in the existing literature, including reinforcement magnitude, reinforcement rate, reinforcer quality, immediacy of reinforcement, delivery of arbitrary reinforcers, and combinations of these aspects. Each of these is discussed in greater detail in the paragraphs that follow.

Magnitude

The magnitude of reinforcement refers to the quantity of the reinforcer delivered (i.e., duration or amount; Hoch et al, 2002a). In Experiment 1 of their study, Athens and Vollmer (2010) manipulated the magnitude of reinforcement (measured in duration of access) for challenging behavior and appropriate behavior exhibited by two individuals with developmental disorders. Specifically, Athens and Vollmer compared a concurrent schedule arrangement with equal magnitudes of reinforcement for challenging and appropriate behavior to a concurrent schedule arrangement with higher magnitude of reinforcement for appropriate behavior relative to challenging behavior. Both participants in this experiment allocated responding to appropriate behavior when it resulted in a higher magnitude of reinforcement than challenging behavior. Similarly, Boyle et al. (2020) manipulated the duration of reinforcement for a child with autism spectrum disorder (ASD) to reduce challenging behavior. Researchers found that a concurrent schedule that provided a longer duration of reinforcement for appropriate behavior relative to challenging behavior resulted in the individual allocating responding to appropriate behavior. These two experiments demonstrated that challenging behavior could be reduced without the inclusion of extinction if a higher magnitude of reinforcement was provided for appropriate behavior.

Rate

Rate of reinforcement refers to how often a reinforcer is provided (Neef et al., 2001). The rate of reinforcement can be manipulated using interval schedules of reinforcement. In interval schedules, reinforcement is provided after the first response emitted when the fixed or variable duration has elapsed (Cooper et al., 2019). Martens et al. (2016) evaluated the effects of reinforcement rate on on-task and off-task behaviors of preschoolers. Teachers were instructed to provide attention for on-task or off-task behaviors on variable interval (VI) schedules of reinforcement. Researchers evaluated the duration of on-task and off-task behaviors in three conditions: equal VI schedules for both behaviors, VI schedule that favored on-task behavior, and VI schedule that favored off-task behavior. Results indicated that on-task behavior was greater when the VI schedule was denser for on-task behavior than off-task behavior. These findings demonstrate the utility of manipulating rate to guide response allocation from challenging behavior to an appropriate behavior.

Quality

The quality of a reinforcer refers to the individual preference for the reinforcing stimuli (Hoch et al., 2002a). Hoch et al. (2002a) evaluated the play behavior of three children with ASD. Researchers compared solitary play with peer play when the quality of toys was manipulated. When toy quality was held constant (i.e., both areas had a highly preferred toy), response allocation was to the solitary play area. When the researcher manipulated the quality by placing the high-quality toy in the area with the peer and a low-quality toy in the solitary play area, response allocation shifted to the area with the peer. These results indicated that manipulating quality impacted response allocation. Further evidence that the quality of reinforcement impacts response allocation was demonstrated by Kunnavatana et al. (2018). As part of their study, a DRA without extinction intervention was developed based on response allocation observed under various concurrent schedule arrangements that targeted arbitrary (i.e., non-clinically relevant) responses. During DRA without extinction, Kunnavatana et al. manipulated the quality of reinforcement received contingent on appropriate behavior and challenging behavior exhibited by three individuals with disabilities. These arrangements resulted in participants allocating responding to appropriate behavior when that behavior produced a higher quality reinforcer relative to the reinforcer related to challenging behavior.

Immediacy

Immediacy refers to the length of time between an individual engaging in the behavior and receiving the reinforcement for the behavior (Neef et al., 2001). Horne and Day (1991) evaluated the immediacy of a break from instruction during functional communication training (FCT). If the participant exchanged a picture card, they received a break after a 20 s delay in one condition or 1 s delay in another condition. Challenging behavior always resulted in an immediate break. Results indicated that the participant's response allocation favored challenging behavior during the condition with a 20 s delay for communication, but response allocation shifted to communication during the 1 s delay condition. This finding suggests that immediacy of reinforcement impacts response allocation. Athens and Vollmer (2010) also manipulated the immediacy of reinforcement following appropriate behavior and challenging behavior in

Experiment 3 of their study. Response allocation favored appropriate behavior when the reinforcer was delivered immediately as opposed to a delay to reinforcement for challenging behavior.

Arbitrary Reinforcers

Carr and Durand (1985) demonstrated that providing the functional reinforcer (i.e., the reinforcer that maintains challenging behavior) for an appropriate behavior reduces challenging behavior. However, including an arbitrary reinforcer (i.e., a non-functional reinforcer) may influence response allocation. Lalli et al. (1999) used positive reinforcement to treat escape-maintained challenging behavior. Researchers provided a preferred edible for appropriate behavior while challenging behavior still resulted in a break. Results indicated that the use of arbitrary reinforcement increased appropriate behavior and decreased challenging behavior. In a more recent study, Berth et al. (2019) used an arbitrary reinforcer to treat escape-maintained food refusal. Researchers found that providing an arbitrary reinforcer (e.g., preferred toys) for food acceptance increased food acceptance and decreased food refusal for one participant.

Combinations

Although individual reinforcement manipulations are useful, creating a combination of reinforcer manipulations within one component of the concurrent schedule (e.g., a longer break with an arbitrary reinforcer vs. a shorter break only) may have a greater impact on response allocation. In a review of the literature, Trump et al. (2020) found that a combination of reinforcement parameters was the most successful schedule arrangement.

Combinations of Reinforcement Parameters. Research shows that combining multiple parameters of reinforcement (i.e., magnitude, rate, quality, and immediacy) in one component of the concurrent schedule impacts the effectiveness of interventions. Peterson et al. (2009)

manipulated both the quality and magnitude of reinforcement to shift allocation towards appropriate behavior. The results showed that response allocation favored the schedule with the longest duration and highest quality reinforcer. Similarly, Athens and Vollmer (2010) combined magnitude and quality with immediacy of reinforcement to treat challenging behavior in Experiment 4 of their study. Athens and Vollmer (2010) found results similar to Peterson et al. (2009) demonstrating that response allocation favored the component with the best quality, highest magnitude, and most immediate reinforcement.

Combinations of Functional and Arbitrary Reinforcers. To impact response allocation in concurrent schedules, researchers can also combine the functional reinforcer with an arbitrary reinforcer. Hoch et al. (2002b) provided a break with a preferred activity for appropriate behavior during treatment for escape-maintained challenging behavior. This reinforcement manipulation resulted in an increase in appropriate behavior and a decrease in challenging behavior that was maintained in follow-up sessions. Davis et al. (2012) also provided a break with a preferred activity for appropriate behavior and found results similar to Hoch et al. (2002b) with increases in appropriate behavior and decreases in challenging behavior that maintained during schedule thinning.

Participant Variables

Research also suggests there are participant variables that impact response allocation. One variable that may impact response allocation is individual preference. Hanley et al. (1997) evaluated individual preference for treatment packages. Researchers used FCT and noncontingent reinforcement to decrease challenging behavior of two individuals. Once both treatments effectively reduced challenging behavior, researchers evaluated individual preference of treatment in a concurrent schedule arrangement. Both participants demonstrated a preference for one treatment package. These results indicated that preference impacted response allocation in a concurrent schedule arrangement.

Another example of individual preference impacting response allocation is shown by Ringdahl et al. (2016). In this study, researchers evaluated individual preference for communication modalities during FCT. Researchers trained individuals to communicate using two communication modalities. Once responding was stable across both communication modalities and challenging behavior decreased by 80% for the participants that engaged in challenging behavior, the researchers arranged a concurrent schedule. In the concurrent schedule, both modalities were simultaneously available, and reinforcement was provided for communication with either modality. Rate, magnitude, quality, and immediacy of reinforcement were held constant throughout the study. Results indicated that all participants exhibited a clear preference for one communication modality as evidenced by greater response allocation to one option over the other.

Uniformity

Research shows that manipulating the reinforcement variables mentioned previously impacts response allocation, but these variables do not have a uniform effect on every individual's behavior. For example, Athens and Vollmer (2010) found that manipulating the magnitude of reinforcement was effective at shifting response allocation to favor appropriate behavior over challenging behavior. However, manipulating the magnitude of reinforcement did not change response allocation in a study conducted by Briggs et al. (2019). Similarly, in Martens et al. (2016) manipulating the rate of reinforcement was successful at shifting allocation from off-task behavior to on-task behavior. Manipulating the rate of reinforcement, however, was unsuccessful at shifting allocation to appropriate behavior in a study by Borrero et al.

(2010). When providing an arbitrary reinforcer with the functional reinforcer, Berth et al. (2019) only decreased challenging behavior and increased appropriate behavior for one participant. For the other four participants, researchers had to implement extinction for challenging behavior to decrease challenging behavior.

Trump et al. (2020) analyzed the effect of different reinforcement parameters in a literature review on differential reinforcement without extinction (i.e., concurrent schedules). Researchers found that manipulating a combination of reinforcement parameters had a positive effect in 100% of studies ($N = 10$); providing the functional reinforcer with an arbitrary reinforcer had a positive effect in 86% of studies ($N = 7$); manipulating the immediacy of reinforcement had a positive effect in 66% of studies ($N = 3$); manipulating the quality of reinforcement had a positive effect in 60% of studies ($N = 5$); providing just an arbitrary reinforcer had a positive effect in 50% of studies ($N = 12$); manipulating the rate of reinforcement had a positive effect in 29% of studies ($N = 14$); and manipulating the magnitude of reinforcement had a positive effect in 11% of studies ($N = 9$). The idiosyncratic results across reinforcement parameters suggest that parameter manipulations may not be universally applicable to all individuals.

Intervention Design

Given the individual factors that impact the efficacy of concurrent schedule-based interventions, arbitrarily selecting a parameter to manipulate may decrease the efficacy of the treatment. Researchers have examined ways to determine individual sensitivity to parameters to determine the most efficacious reinforcement manipulation. Neef et al. (1994) evaluated the effects of rate, quality, and immediacy of reinforcement and response effort on choice behavior for six children with learning and behavior difficulties. Participants were presented with two problems on a computer screen. Both problems were correlated with one of four dimensions: reinforcement rate, reinforcement quality, reinforcement immediacy, and response effort. Results indicate that participants were sensitive to different reinforcement and response parameters. In an extension of this study, Neef et al. (2001) evaluated these same dimensions with 11 adolescents with emotional and behavioral difficulties. Again, participants were presented with two sets of math problems correlated with different response and reinforcement parameters. The findings of this study replicate findings from Neef et al. (1994). Eight participants' behavior was sensitive to one parameter and three participants' behavior was equally sensitive to more than one parameter.

In a more recent study, Kunnavatana et al. (2018) evaluated individual sensitivity to reinforcement parameters using arbitrary behavior as a basis for developing an intervention for challenging behavior. Prior to the intervention, researchers assessed parameter sensitivity using arbitrary responses. In Experiment 1, researchers manipulated one parameter of reinforcement while keeping the other parameters constant. In the quality analysis, pressing one button resulted in a high-quality reinforcer, and pressing another button resulted in a low-quality reinforcer. In the magnitude analysis, pressing one button resulted in a large magnitude of reinforcement, and pressing another button resulted in a small magnitude of reinforcement. In the immediacy analysis, pressing one button resulted in immediate access to reinforcement, and pressing another button resulted in delayed access to reinforcement. Results indicate that two of the three participants allocated responding to the highest quality, largest magnitude, or most immediate reinforcement schedule. One participant only showed differentiation in the quality analysis. The two participants whose behavior was sensitive to all parameters participated in Experiment 2. In Experiment 2, parameters were compared to determine the parameter the participants' arbitrary behavior was most sensitive to. The study included three conditions: 1) magnitude versus immediacy; 2) immediacy versus quality; and 3) magnitude versus quality. Results indicated that both participants' behavior was most sensitive to quality. Following the parameter sensitivity assessments, researchers developed an intervention for challenging behavior within a concurrent schedule arrangement. In this arrangement, challenging behavior resulted in low-quality reinforcement and appropriate behavior resulted in high-quality reinforcement. Results indicated that this arrangement was effective in reducing challenging behavior and increasing appropriate behavior. However, the research design used in this study did not allow for the comparison of intervention effectiveness at the individual level. Specifically, the researchers used a multiple baseline across participants to assess the effectiveness of the parameter participants' behavior was most sensitive to and least sensitive to on decreasing challenging behavior. A multiple baseline across participants design only allows for evaluation of inter-subject replication of treatment effects (Ledford & Gast, 2018). Though this may be a useful design for some interventions, this design does not permit evaluation of intra-subject replication of effects, which is necessary to evaluate the effectiveness of parameter manipulation that participants' behavior is least and most sensitive to.

In addition to the research design limitation, the parameter sensitivity assessment was conducted using an arbitrary response while the intervention was developed to treat challenging behavior. Kunnavatana et al. (2018) are assuming that the parameter that drives allocation with arbitrary responses is the same parameter that will drive allocation with clinically relevant behaviors such as challenging behavior and appropriate behavior. This assumption follows the logic of other assessments conducted prior to interventions. For example, stimulus preference assessments are often conducted to identify potential reinforcers to use in behavior interventions (Hagopian et al., 2004). These assessments are conducted using contrived choice arrangements.

Individuals are allowed to choose between various stimuli and data is collected on choice making. The stimuli that are selected the most are then used as potential reinforcers for appropriate behavior during interventions. The assumption is that choices made in contrived settings will translate to more natural settings. However, this assumption does not follow the logic of other assessments used to design intervention for challenging behavior. Typically, such assessments are designed to determine the reinforcers maintaining challenging behavior (i.e., functional analyses) by directly measuring the effect of antecedents and consequences directly on those challenging behaviors (e.g., Iwata et al., 1982/1994). Such assessments do not focus on the effects of antecedents and consequences on arbitrary behavior. By conducting the parameter sensitivity assessment with arbitrary responses, the logic of the Kunnavatana et al. (2018) study varies from the logic set forth by previous assessments of challenging behavior. This different approach has two potential outcomes: 1) the approach is successful, and researchers can develop a successful intervention for challenging behavior without further reinforcing challenging behavior during the assessment phase; or 2) the approach is unsuccessful, and researchers must conduct further assessments before developing a successful intervention for challenging behavior. More research is needed to evaluate the assumptions made by Kunnavatana et al. (2018) and to evaluate effective and efficient ways to determine sensitivity to parameters of reinforcement.

CHAPTER 2

LITERATURE REVIEW

Effective treatment to reduce challenging behavior often include an extinction component (Corte et al., 1971; Iwata et al., 1990). However, extinction can produce undesirable side effects (Lerman et al., 1999) including extinction bursts and response variability. An extinction burst is the initial increase in behavior when reinforcement is removed (Cooper et al., 2019). In a review of the literature, Lerman et al. (1995) found that extinction bursts occurred in 26% of studies included in the review. When extinction was combined with another intervention (e.g., differential reinforcement), an extinction burst only occurred in 12% of studies. In another review, Lerman et al. (1999) found that extinction bursts occurred in 39% of all studies included in the review. Extinction bursts occurred less (15% of studies) when extinction was combined with another intervention. Although the prevalence of extinction bursts in this review was low, the occurrence of an extinction burst with severe challenging behavior (e.g., self-injury) can be dangerous. Another undesirable side effect of extinction is response variability. Lerman et al. (1999) found that aggression occurred following the removal of reinforcement as part of an intervention for self-injury in 22% of studies reviewed. Though these effects might be transient, both extinction bursts and response variability can be dangerous for the individual or the implementer.

Along with undesirable side effects, the inclusion of an extinction component may hinder treatment outcomes if high fidelity cannot be guaranteed (i.e., intermittent delivery of reinforcers for challenging behavior; Vollmer et al., 2020). Lapses in fidelity have been shown to decrease the success of interventions that include extinction, such as differential reinforcement of alternative behavior (DRA; St. Peter Pipkin et al., 2010). In a study examining types of treatment fidelity errors in the context of DRA, St. Peter Pipkin et al. (2010) found that errors of commission (i.e., providing the reinforcer for challenging behavior- a fidelity breach related to extinction) resulted in increased challenging behavior and decreased appropriate behavior. Thus, if high procedural fidelity seems unlikely, interventions that do not include an extinction component, such as concurrent schedules of reinforcement, may need to be considered to guard against poor clinical outcomes.

To create an effective intervention without an extinction component, researchers have manipulated parameters of reinforcement within a concurrent schedule-based intervention. In an early example of this approach, Lalli et al. (1999) implemented a two-component concurrent reinforcement schedule that included an arbitrary reinforcer (i.e., a reinforcer not related to the function of challenging behavior, in this case a preferred edible) for appropriate behavior and the functional reinforcer (i.e., the reinforcer determined to be functionally related to challenging behavior, in this case a break from instruction) for challenging behavior exhibited by individuals with disabilities. In the concurrent schedule arrangement, both the arbitrary and functional reinforcer were available on a fixed ratio (FR) 1 schedule of reinforcement. If the participant engaged in the appropriate behavior of compliance with demands, a preferred edible was provided. If the participant engaged in challenging behavior, a 30 s break from instruction was provided. This arrangement led to a decrease in challenging behavior and increased appropriate behavior. In a more recent example, Rogalski et al. (2020) manipulated the magnitude of the same reinforcer provided for challenging behavior and appropriate behavior. The functional reinforcer (i.e., escape from demands) was provided on an FR 1 schedule under four reinforcement arrangements: a) 0 s for appropriate behavior and 30 s for challenging behavior; b) 30 s for both appropriate and challenging behavior; c) 240 s for appropriate behavior and 10 s for challenging behavior; and d) 90 s for appropriate behavior and 10 s for challenging behavior. When providing a higher magnitude of reinforcement for appropriate behavior, researchers observed a decrease in challenging behavior and an increase in appropriate behavior.

Other reinforcement parameter manipulations described in the literature include quality (Athens & Vollmer, 2010; Briggs et al., 2019; Kunnavatana et al., 2018), denser reinforcement schedules (Borrero et al., 2010; Brown et al., 2020; Kelley et al., 2002; Martens et al., 2002; Worsdell et al., 2000), and delay (Athens & Vollmer, 2010; Kunnavatana et al., 2018). Interventions that utilize concurrent reinforcement schedules may also include a combination of reinforcement parameters. Combination of reinforcement parameters may include providing the functional reinforcer and an arbitrary reinforcer (Davis et al., 2012; Davis et al., 2018; Hoch et al., 2002; Peterson et al., 2009) or manipulating both the magnitude and quality of reinforcement (Briggs et al., 2019; Peterson et al., 2009) in a single component of the concurrent schedule. In each of these studies, concurrent schedule arrangements were successfully implemented to achieve a reduction in challenging behavior without the need for an extinction component.

Other examples of concurrent schedule arrangements can be found in the differential reinforcement without extinction literature. Differential reinforcement without extinction is best conceptualized as a concurrent reinforcement schedule arrangement (Athens & Vollmer, 2010). Trump et al. (2020) conducted a review of the literature on differential reinforcement without extinction articles published between 1961 and 2015. Researchers reviewed a total of 32 articles, which included 109 experiments. Of the 109 experiments, 34 evaluated equal reinforcement arrangements across both components of the concurrent reinforcement schedule. Components of the concurrent schedule arrangement differed in the remaining 74 experiments. In at least one component of the concurrent schedule, 14 experiments used a denser reinforcement schedule, 12 used an arbitrary reinforcer, 10 used a combination of reinforcement parameters, nine used magnitudes of reinforcement, seven used a functional plus arbitrary reinforcer, five manipulated the quality of reinforcement, and three used a delay to reinforcement.

Trump et al. (2020) found that the most effective parameter manipulation was a combination of reinforcement parameters; 100% of the 10 studies that used a combination of reinforcement parameters produced positive effects. A combination of reinforcement parameters could include a simultaneous manipulation of magnitude and density or quality and magnitude within one component of the concurrent schedule. Of the remaining parameters, 86% of the seven studies that used functional plus arbitrary reinforcement produced positive effects, 66% of the three studies that used delay to reinforcement produced positive effects, 60% of the five studies that used quality of reinforcement produced positive effects, 50% of the 12 studies that used an arbitrary reinforcer produced positive effects, 29% of the 14 studies that used denser reinforcement schedules produced positive effects, 12% of the 34 studies that used equal reinforcement parameters produced positive effects, and 11% of the nine studies that used magnitude of reinforcement produced positive effects. The results of their review show effective behavior change can be produced without the use of extinction.

Although the review by Trump et al. (2020) showed that manipulating the parameters of reinforcement within a concurrent schedule arrangement produces successful behavior change, the review did not provide examination of why or how researchers selected the parameter(s) of reinforcement manipulated in the reviewed studies. The current review replicated Trump et al. (2020) while constraining the search to only concurrent reinforcement schedules (i.e., excluding differential reinforcement of other behavior) and extended the review by assessing the rationale for the reinforcement parameter manipulation. Thus, the purpose of the current review was to evaluate the reported rationale for selecting the manipulated reinforcement parameter(s) in concurrent schedule-based interventions for challenging behavior and to examine the reinforcement parameters of concurrent schedules used to treat challenging behavior.

Method

Search

The first author conducted four multi-database searches using a search engine. The databases included in the search were APA PsycArticles, APA PsycInfo, Child Development & Adolescent Studies, ERIC, Family & Society Studies Worldwide, MEDLINE with Full Text, Professional Development Collection, and SocINDEX with Full Text. The search keywords were 1) differential reinforcement AND without extinction, 2) differential reinforcement of alternative behavior AND without extinction, 3) functional communication training AND without extinction, and 4) 'concurrent schedule' AND reinforcement. The first and third authors examined all academic journal articles written in English that met the following inclusion criteria: a) conducted with human participants, b) used concurrent schedules of reinforcement, c) used a different reinforcement arrangement for each component of the concurrent schedule, and d) included reinforcement for challenging behavior in one component of the concurrent schedule arrangement. The results of these two independent searches identified a similar set of included articles. Authors disagreed on the inclusion of four articles; however, after clarification of inclusion criteria, they agreed the articles met inclusion criteria.

Coding

Articles that met inclusion criteria were coded by the first author based on the following variables: participant demographics, setting, dependent variables, functional analysis results, type of preference/reinforcer assessment, reinforcement parameter manipulate, rationale for the reinforcement parameter selected, and treatment effects. The categories for coding participant demographics, setting, dependent variables, functional analysis results, and reinforcement parameter manipulated were based on the categories used in Trump et al. (2020). Additionally, the first author conducted an analysis of rigor on included articles.

Participant Demographics and Setting

Participant demographic variables included age, sex, and diagnosis. Participant age was categorized into a range category: 0-5 years old, 6-12 years old, 13-21 years old, and 22-44 years old. Participant sex was coded as male, female, or not reported. Participant diagnosis was coded into the following categories: 1) autism spectrum disorder (ASD), pervasive developmental delay (PDD), or developmental delay; 2) mild/moderate intellectual disability; 3) severe/ profound intellectual disability; 4) multiple diagnoses; 5) other; or 6) not reported. The setting was coded into the following categories: participant home/group home, hospital/inpatient clinic, day treatment, outpatient clinic, school, or not reported.

Dependent Variables and Functional Analysis Results

Dependent variables were coded in two categories: challenging behavior and replacement behavior. Challenging behaviors were coded as aggression, self-injury, disruption/property destruction, elopement/flopping, off-task, inappropriate vocalizations, or multiple challenging behaviors. The challenging behavior categories of food refusal and stereotypy included in Trump et al. (2020) were excluded in the current review because no studies included those behaviors. Additionally, inappropriate vocalizations were added as a category. Replacement behaviors were coded as the appropriate behavior targeted during the study and included communication, compliance/cooperation, both, and other. Functional analysis results were coded as attention, escape, tangible, automatic, multiple functions, or not reported.

Preference Assessments

Preference and reinforcer assessments were coded using a numerical scoring system. Studies were scored as 0 if a preference assessment was not specified but items/activities were used as reinforcers during at least one component of the concurrent schedule intervention. Studies were scored as 1 if verbal report was used to determine the items/activities used as reinforcements. A verbal report included reports from participants, parents/caregivers, or teachers about the participant's most preferred items. Studies were scored as 2 if a systematic preference assessment was used. A systematic preference assessment includes evaluating participant preferences using assessments such as a paired choice preference assessment (Fisher et al., 1994) or allowing the participant to choose the preferred item/activity in the context of the study (e.g., a choice board presented prior to a session). Studies that did not use any items/activities as reinforcers, were coded as not applicable.

Reinforcement Parameters and Rationale

The reinforcement parameter manipulated in the concurrent schedule arrangement was coded for each identified study. Parameters included magnitude, density, quality, delay, arbitrary, functional plus arbitrary, and combination. *Magnitude* was coded if the study manipulated the duration or amount of reinforcement. *Density* was coded if the study manipulated the density of reinforcement using a variable- or fixed-ratio or variable- or fixed-interval schedule of reinforcement. *Quality* was coded if the study manipulated the preference of items used during reinforcement (e.g., low preferred items versus high preferred items). *Delay* was coded if the study manipulated the immediacy of the reinforcement following the target behavior. *Arbitrary* was coded if the study included a reinforcer that was not identified in the functional analysis as relevant to challenging behavior. Functional plus arbitrary was coded if the study included both an arbitrary reinforcer and the reinforcer identified in the functional analysis in a single component. *Combination* was coded if the study used more than one reinforcement parameter in a single component of the concurrent schedule (e.g., magnitude plus quality or functional reinforcer plus arbitrary reinforcer in a single component). For studies that included more than one concurrent schedule arrangement, the reinforcement parameters associated with each concurrent schedule were coded separately. For example, if a study used an ABACAC design where phase B was a concurrent schedule manipulating magnitude and phase C was a concurrent schedule manipulating quality, magnitude and quality were coded separately.

Rationale for the reinforcement parameter selected was coded using a numerical scoring system. Studies were scored as 0 if no rationale was provided. Studies were scored as 1 if a verbal rationale was provided (e.g., the article stated that an arbitrary reinforcer was used because teachers reported more compliance/cooperation when the participant worked for a tablet). Studies were scored as 2 if a systematic assessment was used to determine the reinforcement parameter selected (i.e., the author manipulated multiple parameters to determine the parameter the participant was most sensitive to). A systematic assessment of reinforcement parameters could occur prior to intervention (i.e., parameter sensitivity assessment) or during intervention.

Treatment Effects

The researcher coded the immediacy of treatment effects and the replication of treatment effects across subsequent phases. For immediacy of treatment effects, the researcher counted the number of sessions of the prescribed intervention necessary for challenging behavior to decrease and appropriate behavior to increase such that appropriate behavior is occurring more than challenging behavior and this trend is stable and/or moving in a therapeutic direction for at least three consecutive sessions. The prescribed intervention was defined as the intervention that researchers hypothesized would impact behavior. The number of sessions needed were categorized into a range of categories: 3 to 7 sessions, 8 to 12 sessions, 13 to 17 sessions, 18 or more sessions, and no effect. Studies that used different measurement systems for appropriate and challenging behavior (e.g., percent of trials for appropriate behavior and rate for challenging behavior) were labeled not applicable for evaluating treatment effects because the change in each behavior could not be compared.

For studies where the immediacy of effect was assessed, the replication of those treatment effects was coded using a numerical coding system. Studies were scored as 0 if the treatment effects did not replicate in subsequent phases. Studies were scored as 1 if the treatment effects replicated but the effect was smaller than the first phase. Studies were scored as 2 if the treatment effects replicated and the effect was equal to or larger than the first phase. Studies that did not allow for the replicability of treatment effects based on the design were scored not applicable.

Analysis of Rigor

The researcher conducted an analysis of rigor for 11 articles identified that were not included in the recent Trump et al. (2020) review (Berth et al., 2019; Bishop et al, 2013; Boyle et al., 2020; Briggs et al., 2019; Brown et al., 2020; Davis et al., 2018; Dowdy & Tincani, 2020; Dowdy et al., 2018; Kunnavatana et al., 2018; Rogalski et al., 2020; Stuesser & Roscoe, 2020). The rigor analysis was conducted using the single-case analysis and review framework (SCARF;

Ledford et al., 2020) tool; the same tool Trump et al. (2020) used to assess the article quality (see Trump et al., 2020 for an analysis of rigor for the remaining 11 articles included in this review). The SCARF tool analyzes single-case designs across three domains: rigor, quality and breadth of measurement, and primary outcomes. In the rigor domain, studies are rated based on reliability data, procedural fidelity data, and sufficient data points in each condition using yes/no questions. In the quality and breadth of measurement domain, studies are rated based on study descriptions (i.e., participant demographic, dependent variables, condition procedures) and measures of generalization, maintenance, and social and ecological validity using yes/no questions with a few 0 to 4 rating scales. In the primary outcomes domain, primary, generalization, and maintenance outcomes are rated on a numerical scale from 0 to 4. Scores are populated based on ratings in each domain and automatically graphed by the SCARF tool.

Interrater Reliability (IRR)

An independent review screened all articles identified in the initial search using the inclusion criteria. Independent reviewers also coded 39.1% of the included articles using an identical coding template and coded 36.3% of the articles included in the analysis of rigor using the SCARF tool for the purpose of calculating interrater reliability (IRR). An agreement for inclusion or exclusion of a study was scored when the first author and reviewer included or exclude the same study. An agreement for coding included articles was scored when the researcher and the reviewer coded the same result in an individual cell of the coding sheet and SCARF tool. Interrater Reliability was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. The IRR for article inclusion was 100%. The average IRR for demographic and descriptive variables (i.e., setting, dependent variable, functional analysis results, and preference assessments) was 91.9% (range, 75% to

100%). The IRR for the reinforcement parameter manipulated and the rationale for reinforcement parameter selected were both 87.5%. The IRR for treatment effects was 91.7%. The average IRR for the analysis of rigor with the SCARF tool was 84.8% (range, 78.3% to 93.5%).

Results

The initial database searches with the keywords resulted in a total of 988 articles published between 1961 and 2020. All articles were screened using the inclusion criteria, resulting in 12 articles. The first author conducted an ancestral search of the included articles that yielded 33 additional articles. After the inclusion criteria was applied to the articles from the ancestral search, an additional 10 articles were added to the review. Thus, a total of 22 articles, containing 73 experiments, which further included 82 parameter analyses, were included in the current review. Table 1 provides an overview of participants, reinforcement parameter manipulations, and rationale scores.

Participant Demographics and Setting

The 22 articles in this review included a total of 72 participants. Of the 72 participants, 45 (62.5%) were male, 20 (27.8%) were female, and seven (9.7%) were not reported. For participant age ranges, 15 (20.8%) were between 0 to 5 years old, 37 (51.4%) were between 6 and 12 years old, 14 (19.4%) were between 12 and 21 years old, and six (8.3%) were between 22 and 44 years old. Most participants had multiple diagnoses (31 participants, 43%). Of the remaining participants, 23 (31.9%) were diagnosed with ASD, PDD, or developmental disability; eight (11.1%) were diagnosed with severe/profound intellectual disability; two (2.7%) were diagnosed with mild/moderate intellectual disability; six (8.3%) had other diagnoses; and two (2.7%) were not reported.

The two most common setting locations were schools (29 experiments, 40.2%) or outpatient clinics (20 experiments, 27.8%). Of the remaining experiments, 12 (16.7%) were conducted in hospitals/inpatient clinics; five (6.9%) were conducted in day treatment programs; three (4.1%) were conducted in the participants home/group home; and three (4.1%) were not reported.

Dependent Variables and Functional Analysis Results

Dependent variables were coded separately for challenging behavior and replacement behavior. Regarding challenging behavior, half of the participants engaged in multiple topographies of challenging behavior (36 participants 50%). Of the remaining participants, 11 (15.3%) engaged in aggression; nine (12.5%) engaged in property destruction/disruption; seven (9.7%) engaged in self-injury; five (6.9%) engaged in off-task behavior; three (4.2%) engaged in elopement/flopping; and one (1.4%) engaged in inappropriate vocalizations. For replacement behaviors, cooperation with instruction was selected for 37 (51.3%) of the participants, a communicative response was selected for 26 (36.1%) of the participants, and nine (12.3%) earned reinforcement for cooperating with instructions and exhibiting a communicative response. No experiments used replacement behaviors other than communication and/or cooperation.

The most common function of challenging behavior exhibited by participants was escape (34 participants, 47.2%). Of the remaining participants, 21 (29.2%) engaged in challenging behavior that served multiple functions (all were social positive plus social negative), nine (12.5%) engaged in challenging behavior to access tangible items, three (4.2%) engaged in challenging behavior to access attention, and five (6.9%) were not reported. There were no participants that engaged in challenging behavior that was maintained by automatic reinforcement.

Preference Assessments

Most of the articles included in the current review (12 articles, 54.5%) reported the use of a systematic preference assessment. The systematic preference assessments include paired-choice, multiple stimulus without replacement, and free-operant preference assessments. Of the remaining articles, five (22.7%) used verbal report, three (13.6%) did not report any preference assessment, and two (9%) did not use any items so a preference assessment was not applicable.

Reinforcement Parameters and Rationale

In the 73 experiments identified for the current review, there were 82 reinforcement parameter analyses. The most common parameter was an arbitrary reinforcer, used 26 (31.7%) times. Of the remaining analyses, denser schedule was each used 15 (18.3%) times; a combination of reinforcement parameters in a single component was used 13 (15.9%) times; magnitude was used 11 (13.4%) times; functional plus arbitrary was used eight (9.7%) times; quality was used seven (8.5%) times; and delay was used two (2.4%) times.

Although reinforcement parameter manipulations were reported in all of the articles, only eight (36.4%) articles included some rationale for the parameter selected for the study. Six articles (27.3%) included verbal reports for their rationale. Examples of verbal reports included creating an intermittent schedule of reinforcement to decrease but not eliminate the challenging behavior (Brown et al., 2020) and reporting that the challenging behavior already received some magnitude of reinforcement (Boyle et al., 2020). Finally, two articles (9.1%) included a systematic rationale for the reinforcement parameter manipulated.

Treatment Effects

Of the 73 experiments included in the current review, the researcher coded 43 (58.9%) for the immediacy of treatment effects. Most of the experiments (22 experiments, 51.2%)

demonstrated an immediate effect between 3 to 7 sessions. For the remaining experiments, three (6.9%) demonstrated an effect between 8 to 12 sessions; three (6.9%) demonstrated an effect between 13 to 17 sessions; three (6.9%) demonstrated an effect in 18 or more sessions; and 12 (27.9%) did not demonstrate an effect. The breakdown of treatment effects based on the rationale for selecting the parameter evaluated is summarized in Table 2. For the 52 experiments that scored a 0 for rationale, 17 (32.7%) demonstrated an effect in three to seven sessions; three (5.6%) demonstrated an effect in eight to 12 sessions; three (5.6%) demonstrated an effect in 13 to 17 sessions; three (5.6%) demonstrated an effect in 18 or more sessions; five (9.6%) did not demonstrated an effect; and 21 (40.4%) were not coded due to incomparable data for appropriate and challenging behavior. For the 16 experiments that scored a 1 for rationale, four (25%) demonstrated an effect in three to seven sessions; seven (43.8%) did not demonstrate an effect; and five (31.6%) were not coded due to the measurement differences for appropriate and challenging behavior. Of the 5 experiments that scored 2 for rationale, one (20%) demonstrated an effect in three to seven sessions and four (80%) were not coded due incomparable results for each behavior.

Based on the immediacy of treatment effects, the researcher coded 31 (72.1%) experiments for replication of treatment effects. A little over half of the experiments (16 experiments, 51.6%) demonstrated a replication of treatment effects with an equal to or larger effect compared to the previous phase of the same condition. Of the remaining experiments, seven (22.6%) experiments demonstrated a replication of treatment effects with a smaller effect compared to the previous phase of the same condition, three (9.7%) did not demonstrate a replication of treatment effects, and five (16.1%) did not include an opportunity to demonstrate replication based on the design used.

Analysis of Rigor

The researcher evaluated 31 experiments (11 articles) not included in Trump et al (2020) using the SCARF tool. Results from the SCARF tool are shown in Figure 1. The first panel depicts the primary outcomes and overall quality and rigor. Studies with positive effects and high quality and rigor appear in the top right quadrant. Studies with non-effect or negative effects and low quality and rigor appear in the bottom left quadrant. Of the 31 experiments, 17 (54.8%) were scored as having positive effects, but only three were also scored as having high quality and rigor. The second panel depicts the generalized outcomes and quality and rigor of generalization measurement. Only two experiments assessed generalization (Bishop et al.,2013; Boyle et al., 2020). Both experiments received high generalized outcome scores, but only one experiment had a high quality and rigor of generalization measurement (Bishop et al., 2013). The third panel depicts the maintenance outcomes and latency of maintenance measurement. Eight experiments assessed maintenance (Dowdy & Tincani, 2020; Dowdy et al., 2018; Stuesser & Roscoe, 2020). All eight experiments received high maintenance outcome scores and high quality and rigor of maintenance measurement.

Table 3 depicts the results of the SCARF analysis by reinforcement parameter. Of the 31 experiments included in the SCARF analysis, there were 40 reinforcement parameter manipulations. Sixteen (40%) experiments used an arbitrary reinforcer, 12 (30%) used a magnitude manipulation, five (12.5%) used a quality manipulation, four (10%) used a combination of reinforcement parameters, two (5%) used a denser schedule manipulation, and one (2.5%) used functional plus arbitrary reinforcement. No experiments analyzed using the SCARF tool included a delay to reinforcement manipulation. The SCARF tool accounts for demonstrations of effect when calculating the experimental outcomes. Most of the negative or minimal effects were due to a lack of three demonstrations of effect. Table 4 depicts the combined SCARF results from the current review and the review conducted by Trump et al. (2020).

Discussion

The current review examined reinforcement parameter manipulations described in 22 articles reporting on concurrent schedule arrangements. This review partially replicates the review conducted by Trump et al. (2020) by examining behavior analytic articles that use concurrent schedules without extinction to treat challenging behavior. This review also extends the review conducted by Trump et al. (2020) by including articles published between the years 2015 and 2020. The current review found that an arbitrary reinforcer was the most common reinforcement parameter manipulated. Lalli et al. (1999) used positive reinforcement to treat escape-maintained challenging behavior. In their study, researchers compared providing an edible for compliance with tasks while still providing a break for challenging behavior to providing a break for both compliance and challenging behavior. Researchers found that providing the arbitrary reinforcer for appropriate behavior increased compliance and decreased challenging behavior. Although arbitrary reinforcers were the most common parameter manipulation, the SCARF analysis found that only 50% of experiments with arbitrary reinforcers. These results match the SCARF analysis results found in the Trump et al. (2020) review.

As shown by the current review and the review conducted by Trump et al. (2020), concurrent-based schedules without extinction can be effective interventions to reduce challenging behaviors; however, this effectiveness is not the same for all participants. Of the 43 experiments coded for treatment effects in the current review, 12 (27.9%) experiments did not demonstrate any treatment effects. Looking at the combined SCARF analysis results from the Trump et al. (2020) review and the current review, four of the seven reinforcement parameters (magnitude, denser schedule, quality, and arbitrary) had a negative or minimal effect for at least 50% of experiments. This effectiveness may increase when researchers provide a systematic rationale for the parameter they selected to manipulate. Only two studies in the current review conducted a systematic analysis of parameters (i.e., provided a systematic rationale for parameter selection). Kunnavatana et al. (2018) conducted multiple parameter sensitivity assessments before developing a concurrent schedule-based intervention to treat challenging behavior. The first assessment evaluated each participant's sensitivity to quality, magnitude, and immediacy of reinforcement individually using arbitrary responses. In each evaluation, pressing one button resulted in one reinforcement schedule (i.e., high quality item, large magnitude, or immediate reinforcement) and pressing another button resulted in a different reinforcement schedule (i.e., low quality item, small magnitude, or delayed reinforcement). Researchers evaluated sensitivity to each parameter in a reversal design where the reinforcement schedule associated with each button was flipped. The second assessment was used to evaluate the parameter participants were most sensitive to when multiple parameters were identified in the first assessment. This assessment was conducted similarly to the first assessment, except each button was associated with a combination of reinforcement parameters (i.e., button one resulted in low magnitude reinforcement delivered immediately and button two resulted in high magnitude reinforcement with a delay to reinforcement). Following the parameter sensitivity assessments, researchers conducted a multiple baseline across participants analysis of concurrent schedules to treat challenging behavior. Researchers also compared the participant's most sensitive parameter to a less sensitive parameter. Results of this experiment showed that not all parameters were effective at reducing challenging behavior and maintaining appropriate behavior and that the parameter participants' behavior was most sensitive to was effective at reducing challenging behavior and increasing appropriate behavior. Briggs et al. (2019) compared participant responding when challenging behavior and appropriate behavior resulted in the same reinforcement, when appropriate behavior resulted in a higher magnitude of reinforcement, when appropriate behavior resulted in a higher quality of reinforcement, and when appropriate behavior resulted in a combined higher magnitude and quality of reinforcement. Researchers found that all participants engaged in less challenging behavior and more compliance when compliance resulted in a better schedule of reinforcement. However, the optimal schedule of reinforcement was different for each participant. Although only one of these two studies were coded for immediacy of treatment effects, the article coded demonstrated treatment effects immediately (within three to seven sessions). This finding suggests that conducting parameter sensitivity assessments with each participant may increase the immediacy of effect and overall effectiveness of a concurrent schedule-based intervention; however, more research is needed to determine if basing intervention on the outcomes of a parameter sensitivity assessment increases the likelihood of effectiveness.

Limitations

There were a few limitations in the current review. First, concurrent schedule-based interventions that used equal schedules for appropriate behavior and challenging behavior were excluded from the current review. Although this criterion was included to specifically evaluate the differences in parameter manipulation, it may have been useful to evaluate concurrent schedules with the equal schedule components. Second, to assess the treatment effects, the researcher relied solely on the visual analysis of each graph. Because the researcher relied on visual analysis, the definition for immediacy of treatment effects was not based on a numerical value (e.g., 90% reduction in challenging behavior compared to baseline). The definition was instead based on an observable shift in the data where appropriate behavior was occurring more than challenging behavior. Due to the definition of immediacy of treatment effect, many experiments were excluded because the measurement systems used for challenging behavior and appropriate behavior were not comparable. For example, depicting appropriate behavior as percent of trials and challenging behavior as rate makes it impossible for the observer to determine if appropriate behavior is happening more than challenging behavior. Finally, this review only includes published experiments, so the conclusions drawn may be skewed due to a publication bias. Experiments with negative or minimal results are less likely to be published (Tincani & Travers, 2019), so it is possible that concurrent scheduled-based interventions that were not successful were not get published. If unsuccessful interventions are left out of evaluations, the conclusions drawn about the success of that intervention are biased.

Future Research

Given the increasing number of studies utilizing concurrent schedule-based interventions and demonstrating their efficacy, more research is needed to determine the long-term effect of these interventions. Researchers should evaluate the maintenance and generalization of concurrent schedule-based interventions. Based on the SCARF results from the current review, only two of the 31 experiments assessed generalization of effects and only eight of the 31 experiments assessed maintenance of effects. Those that assessed generalization and maintenance demonstrated that the behavior change generalized to other therapists and context as well as maintained over time; however, more research is needed in this area. Researchers should also look at how the schedule of reinforcement can be thinned within a concurrent schedule arrangement. This focus is especially important for the practicality of concurrent schedule-based interventions. A dense schedule of reinforcement may interfere with learning outcomes and take a lot of time and effort to implement. Thus, it is important for researchers to examine ways to thin the schedule that does not hinder treatment effects. Future research should assess individual break points in effectiveness as the reinforcement schedule is thinned.

In addition to assessing generalization, maintenance, and schedule thinning, more research is needed to determine which reinforcement parameters will be effective for individual clients/participants. As shown in the current review, only some reinforcement parameters were effective at increasing appropriate behavior and decreasing challenging behavior, and the effective parameter was different for most participants. There is some evidence to suggest that conducting a parameter sensitivity assessment prior to developing a concurrent schedule-based intervention may increase intervention efficacy. Using arbitrary behavior, Kunnavatana et al. (2018) conducted a reinforcer parameter sensitivity assessment and reinforcer parameter hierarchy assessment with three participants as a basis for designing intervention for challenging behavior. The resulting intervention was effective at shifting allocation from challenging behavior to appropriate behavior. While this study provided evidence for the efficacy of pre-intervention parameter sensitivity and hierarchy assessments, it was conducted using arbitrary behaviors. Researchers assumed that the reinforcer parameter identified as impacting an arbitrary behavior was the same parameter that would impact challenging behavior. This assumption does not follow the logic of other assessments used to develop interventions for challenging behavior. In such assessments, researchers/practitioners typically arrange contingencies such that consequences are provided following the challenging behavior as opposed to arbitrary behavior as a proxy for challenging behavior. In the assessments used by Kunnavatana and colleagues

(2018), it was possible that arbitrary behaviors were sensitive to changes in one parameter while clinically relevant behavior was sensitive to a different parameter. Basing the concurrent schedule arrangement on a parameter that was relevant in a non-clinical context might limit success in clinical contexts. Future research should examine the effects of conducting parameter hierarchy assessments with arbitrary behaviors and clinically relevant behaviors to determine if the assessments yield the same results. If the results of both assessments align, the results will provide supporting evidence for the use of a parameter hierarchy assessment utilizing arbitrary behavior prior to an intervention for challenging behavior. The use of a pre-intervention assessment with arbitrary behaviors would allow practitioners to get to treatment without adding to the reinforcement history for challenging behavior and without implementing intervention in what might amount to a trial and error approach. If the results of the assessments do not align, the results will provide supporting evidence for conducting the assessment utilizing clinically relevant behavior (i.e., the target behavior).

Given the need to evaluate assessments to determine effective reinforcer parameter arrangements for concurrent schedules, the purpose of the current study was to compare results from a parameter hierarchy assessment with arbitrary behaviors to a parameter hierarchy assessment with clinically relevant behaviors. The current study evaluated the following research questions:

1. Can parameter assessments identify a parameter hierarchy for both arbitrary and clinically relevant behaviors?
2. If so, will the outcomes of parameter hierarchy assessments conducted with arbitrary behavior align with outcomes of the same assessment conducted with clinically relevant behavior when a component of the concurrent schedule includes reinforcement for challenging behavior?

Table 1

Type of Reinforcement Parameter and Rationale Score

Citation	*N*	Reinforcement parameter	Rationale
Adelinis et al. (2001)	1	Arbitrary	0
Athens & Vollmer (2010)	7	Magnitude, Quality, Delay, and Combination	0
Berth et al., (2019)	5	Arbitrary	1
Bishop et al., (2013)	3	Arbitrary	0
Borrero et al., (2010)	3	Density (VI schedule)	0
Boyle et al., (2020)	1	Magnitude	1
Briggs et al., (2019)	4	Magnitude, Quality, and Combination	2
Brown et al., (2020)	2	Density (VI schedule)	1
Davis et al., (2012)	4	Functional plus arbitrary	0
Davis et al., (2018)	1	Functional plus arbitrary	0
Dowdy & Tincani (2020)	2	Arbitrary	0
Dowdy et al., (2018)	2	Arbitrary	1
Hoch et al., (2002b)	3	Functional plus arbitrary	0
Kelley et al., (2002)	3	Density (VI schedule)	1
Kunnavatana et al., (2018)	3	Magnitude and Quality	2
Lalli et al., (1999)	5	Arbitrary	0
Martens et al., (2016)	2	Density (VI schedule)	1
Peterson et al., (2009)	7	Combination (functional, arbitrary, magnitude, and quality)	0
Piazza et al., (1997)	3	Arbitrary	0
Rogalski et al., (2020)	3	Magnitude	0
Stuesser & Roscoe (2020)	4	Arbitrary	0
Worsdell et al., (2000)	5	Density (FR schedule)	0

Note. VI = Variable Interval; FR = Fixed Ratio. Rationale score of 0 = no rationale provided, rationale score of 1 = verbal rationale, and rationale score of 2 = systematic assessment as rationale.

Table 2

Immediacy of Treatment Effects by Rationale Score

Rationale Score	Number of sessions needed to demonstrate treatment effects					
	3 to 7	**8 to 12**	**13 to 17**	**18 and up**	**No effect**	**Not Applicable**
0	17	3	3	3	5	21
1	4	0	0	0	7	5
2	1	0	0	0	0	4

Table 3

Experimental Outcome by Reinforcement Parameter Manipulation

Reinforcement Parameter	*N*	Positive Effect (%)	Negative/Minimal Effect (%)
Magnitude	12	66.7	33.3
Density	2	0	100
Quality	5	40	60
Arbitrary	16	50	50
Functional plus arbitrary	1	100	0
Combination	4	25	75

Note. Experimental outcomes calculated using the Single-Case Analysis and Review Framework.

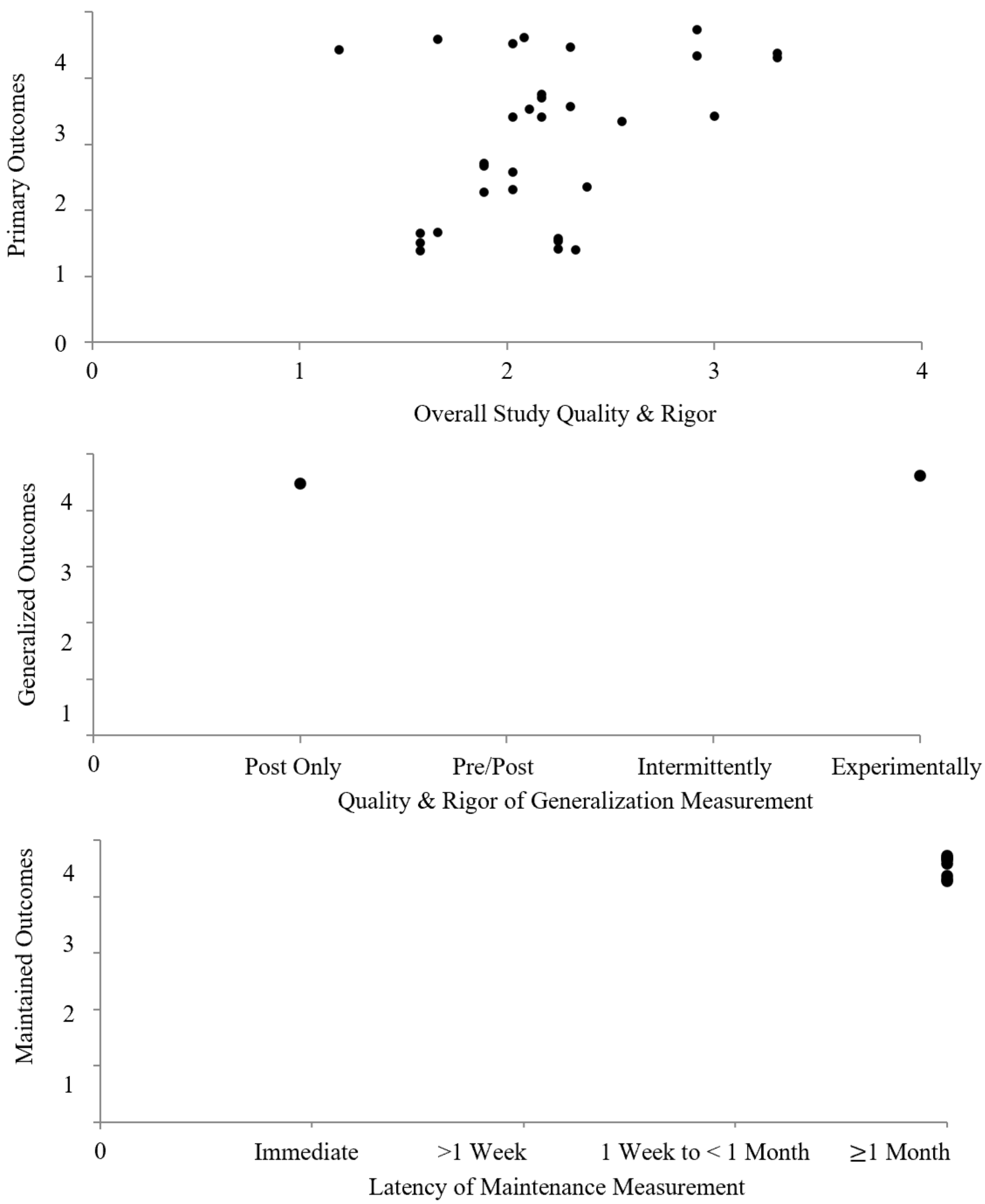
Table 4

Combined Experimental Outcomes by Reinforcement Parameter

Reinforcement Parameter	*N*	Positive Effect (%)	Negative/Minimal Effect (%)
Magnitude	21	42.9	57.1
Density	16	25	75
Quality	10	50	50
Delay	3	66.7	33.3
Arbitrary	28	50	50
Functional plus arbitrary	8	87.5	12.5
Combination	14	75.6	21.4

Note. Experimental outcomes calculated using the Single-Case Analysis and Review Framework from the current review combined with outcomes from Trump et al. (2020, p. 7).

Figure 1
Single-case Analysis and Review Framework Results



Note. Each data point represents an experiment. Studies including multiple experiments are all depicted on separate data points. The top panel depicts the primary outcome scores, the middle panel depicts the generalized outcome scores, and the bottom panel depicts the maintenance outcome scores.

CHAPTER 3

METHOD

Participants, Settings, and Materials

Four school-aged children were recruited for participation in this study. Participants were recruited from a local public elementary school and a university-based outpatient clinic to address behavioral concerns. Participants were included if they met the following criteria: a) were between 3 and 17 years old at the time of enrollment, b) had a diagnosis or special education eligibility of intellectual or developmental disabilities or developmental delay, c) engaged in challenging behavior (e.g., aggression, disruption, screaming), and d) demonstrated sensitivity to changes in at least two reinforcement parameters in the parameter sensitivity screener (described later in this chapter).

Travis was an 8-year-old White, Non-Hispanic male diagnosed with ASD. He communicated vocally using full sentences. He was referred to the outpatient behavior clinic for assessment and treatment of aggression and received behavior analytic services two hours per week. Prior to the current study, a behavior analyst administered the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP). Travis scored a 5 (out of 15) in manding; 10 (out of 15) in tacting; 9.5 (out of 15) in listener skills; 12 (out of 15) in visual perceptual skills and matching-to-sample; 13.5 (out of 15) in independent play skills; 2 (out of 15) in social skills and social play; 7 (out of 10) in motor imitation; 10 (out of 10) in vocal imitation; 5 (out of 5) in spontaneous verbal behavior; 5 (out of 10) in listener responding by function, feature, and class;

3 (out of 10) in intraverbal; 3 (out of 10) in group skills; 3.5 (out of 10) in linguistics; and 5 (out of 5) in reading, writing, and math. Travis's overall VB-MAPP score was 103.5 (out of 170).

Aliyah was a 4-year-old Black, Non-Hispanic female with a special education eligibility of autism. She communicated with gestures and picture card exchange. She received special education services in a self-contained preschool classroom for students who engage in severe challenging behavior at a local public elementary school. Prior to the current study, school personnel conducted the Adaptive Behavior Assessment System-Third edition (ABAS-3) and the Developmental Profile 3 (DP-3). On the ABAS-3, Aliyah's practical domain standard score was 52 (3.2 standard deviations below the mean), and her social domain standard score was 54 (3.07 standard deviations below the mean). On the DP-3, Aliyah's cognitive domain standard score was less than 50 (more than 3.33 standard deviations below the mean), adaptive behavior domain standard score was 56 (2.93 standard deviations below the mean), and social-emotional domain standard score was less than 50 (more than 3.33 standard deviations below the mean).

Carson was a 4-year-old White, Non-Hispanic male with a special education eligibility of significant developmental delay. He communicated vocally in full sentences. He was referred to the outpatient behavior clinic for assessment and treatment of screaming and aggression and received behavior analytic services one hour per week. Prior to the current study, public school personnel conducted the Preschool Evaluation Scale-2nd edition (PES-2) and the Developmental Profile-4 (DP-4) with Carson. On the PES-2, Carson's Cognitive Thinking standard score was 10 (0 standard deviations below the mean), Self-Help Skills standard score was 9 (0.33 standard deviations below the mean), and Social/Emotional standard score was 5 (1.67 standard deviations below the mean). On the DP-4, Carson's Cognitive standard score was 97 (0.2 standard deviations below the mean), Adaptive Behavior standard score was 101 (0.07 standard deviations below the mean), and Social/Emotional standard score was 67 (2.2 standard deviations below the mean). Based on these assessments, Carson has a significant developmental delay in the social-emotional domain.

Christian was a 7-year-old White, Non-Hispanic male diagnosed with ASD. He communicated vocally in full sentences. He was referred to the outpatient behavior clinic for assessment and treatment of screaming and received behavior analytic services two hours per week. Following initial assessment, the clinic staff determined that his clinical goals did not align with the study goals. Thus, his participation in the study was ended after the preference assessment and a clinical course unrelated to the current study was pursued.

All sessions were conducted in a treatment room at the outpatient clinic (Travis and Carson) or in a designated area of the special education classroom (Aliyah). The outpatient clinic treatment room included a table, chairs, and session specific stimuli (e.g., leisure items, microswitches, and laminated colored paper). The treatment room was also equipped with cameras and a microphone for video/audio recording and observation. The designated area of the special education classroom was either a section of the classroom play area (separated with a divider) or at Aliyah's work table. Both classroom areas included session specific stimuli (e.g., leisure items, microswitches, communication device, and laminated colored paper) and a computer setup for video recording.

Pre-Experimental Assessments

Stimulus Preference Assessments

The researcher interviewed Travis's caregiver and Aliyah's teachers to determine leisure items to include in the preference assessment. Once a list of six potentially preferred items was generated, the researcher conducted a preference assessment. For Carson, the researcher used data from a recently conducted paired-choice preference assessment on file at the university outpatient clinic. The high-preferred item identified was toy construction trucks and the low-preferred item identified was a rubber playground ball. For Aliyah, the research conducted a paired-choice preference assessment (Fisher et al., 1992). Items were presented in pairs and the participant was instructed to select one item. *Selection* was defined as reaching for or pointing to an item. Following selection, Aliyah was given 30 s to engage with the selected item. The assessment continued until each item had been presented with each of the other items. Trained graduate students collected data using a paper data sheet to mark which item was selected. After all items were paired with each other, the researcher calculated the percentage of trials for which item was selected. The percentage of trials an item was selected was used to determine relative preference. An item was defined as high preferred if it was selected between 80% and 100% of trials and an item was defined as low preferred if it was selected in 40% of trials or less.

For Travis, the researcher conducted a response-restriction analysis (Hanley et al., 2003) to determine stimulus preference due to challenging behavior during the paired-choice preference assessment. The researcher arranged all items in a semi-circle in front of the participant and told Travis he may engage with any item. Sessions were 5 min, and trained graduate students collected data on the duration of engagement with each item using continuous 5 s partial interval recording. *Engagement* was defined as touching an item with one or both hands for any amount of time. The researcher calculated percent engagement by dividing the number of intervals with engagement by the total number of intervals and dividing by 100. If an item was selected for 60% of intervals across two sessions, the item was removed from the array for subsequent sessions. Sessions continued until no items were left in the array, little to no interaction (20% or less of intervals) occurred, or challenging behavior occurred with the remaining items. Items were ranked based on the percentage of interaction across all sessions that item was presented. The two items with the highest percent of interaction were considered high preferred and the two items with the lowest percent of interaction were considered low preferred.

Reinforcer Assessment

The researcher conducted a reinforcer assessment (concurrent operants assessment; see DeLeon et al., 2001) to determine if the magnitudes of the stimuli identified via the preference assessment functioned as reinforcers. Specifically, the researcher evaluated four consequence arrangements: 1) high preferred item and low magnitude (15 s) versus no consequence; 2) high preferred item and high magnitude (90 s) versus no consequence; 3) low preferred item and low magnitude (15 s) versus no consequence; and 4) low preferred item and high magnitude (90 s) versus no consequence. Prior to the evaluation, the research selected academic tasks from their current educational plan. For Travis, the task was completing simple math equations. For Aliyah, the task was receptive identification of common objects. For Carson, the task was tracing pre-k sight words. During the evaluation, two identical tasks were placed on the table in front of the participant. Behind one task was an item identified in the preference assessment and a visual aid to depict the magnitude of reinforcement available (e.g., timer or poker chips). There was nothing placed behind the second task. At the start of the trial, the researcher described the consequences available for completing each task and instructed the participant to select one of the tasks. *Selection* was defined as pointing to or touching a task or vocally stating the task name. After completion of the task, the researcher provided the consequence associated with the task selected. *Task completion* was defined as engaging in the target response with no more than a model prompt. Each session consisted of 10 trials. The researcher or trained graduate students collected data on *selection* and *task completion* each trial.

Functional Analysis

The researcher conducted a functional analysis to determine the variables maintaining challenging behavior for Travis and Aliyah. Prior to the study, the clinical team at the university outpatient clinic conducted a functional analysis with Carson and determined his behavior was maintained by escape from academic demands. *Challenging behavior* was individually defined for each participant. For Travis and Aliyah, the target challenging behavior was aggression. For Travis, *aggression* was defined as wrapping one or both arms around another person and pushing with enough force to alter the person's original position. For Aliyah, *aggression* included hitting or kicking and was defined as the participant's hand or foot contacting another person with enough force to produce an audible sound, and biting, defined as the participant's mouth opening and closing on another person's body with enough force to leave teeth marks or redness on the skin. All attempts at engaging in aggression were also scored. For Carson, the target challenging behavior was *screaming*, defined as vocalizations above a normal conversational level for any length of time.

For Travis, the researcher conducted a pairwise functional analysis based on the test-control design functional analysis described by Iwata et al. (1994b), but with a tangible test condition. Due to the nature of Travis's aggression (i.e., long duration), the researcher selected a pairwise functional analysis to minimize carryover effects across test conditions. The included test conditions were determined based on researcher observation and parent interviews. The functional analysis included a tangible test condition and a control condition. All sessions were 5 min. Before each tangible test session, the researcher provided access to a high preferred item. At the start of the session, the researcher removed the item but still provided attention. Contingent on the occurrence of challenging behavior, the researcher provided the preferred item for 30 s. During each session of the control condition, the researcher provided access to preferred items and attention. There were no programmed consequences for challenging behavior. During both conditions, the researcher or trained graduate students collected data on the duration per occurrence of aggression. The researcher calculated the total duration of aggression by adding all the durations per occurrence of aggression for the session and reported the data as percent of session time.

For Aliyah, a latency-based functional analysis based on procedures described by Thomason-Sassi et al. (2011) was conducted. The researcher selected a latency-based functional analysis to decrease the number of response-reinforcer pairings for challenging behavior due to the setting. All sessions were conducted by the classroom teacher, who was a trained behavior analyst. The functional analysis included three test conditions (attention, tangible, and escape) and a control condition. The control condition was 5 min during which the teacher provided access to preferred items and attention and refrained from delivering instructions. There were no programmed consequences for challenging behavior. The test conditions ended following the programmed consequence for the first instance of aggression or when 5 min elapsed if no aggression occurred. Prior to each session of the attention test condition, the teacher provided attention for 2 min. At the start of the session, the teacher removed attention and provided a moderately preferred item. The teacher stated, “I have to work, but you can play with this.” Contingent on the occurrence of aggression, the teacher provided attention for 30 s before moving on to the next session of the assessment. Before each session of the tangible test condition, the teacher provided access to a high preferred item. At the start of the session, the teacher removed the item but still provided attention. Contingent on the occurrence of aggression, the teacher provided the preferred item for 30 s before moving on to the next session of the assessment. During each session of the escape test condition, the teacher presented academic demands. The teacher used three-step prompting (i.e., verbal prompt, model prompt, and physical prompt) when placing demands. Contingent on the occurrence of aggression, the teacher removed demands for 30s before moving on to the next session of the assessment. Across all conditions, the researcher or trained graduate student collected data on the occurrence of aggression. The researcher calculated the latency to aggression by subtracting the time aggression occurred by the time the teacher presented the relevant antecedent.

Interobserver Agreement and Procedural Fidelity

An independent, trained second observer collected procedural fidelity data and data for the purpose of calculating interobserver agreement (IOA) on a range of 35% to 100% across all phases and participants. The researcher calculated IOA for the paired choice preference assessment using trial-by-trial agreement. An agreement was scored if data collectors marked the same item selection each trial. An agreement percentage was calculated by dividing the number of trials with agreement by the total number of trials and multiplying by 100. The researcher calculated IOA for the response-restriction analysis, reinforcer assessments, and functional analyses using interval-by-interval agreement with 10 s intervals. An agreement was scored if data collectors scored the same occurrence or nonoccurrence of a behavior within the same 10 s interval. An agreement percentage was calculated by dividing the number of intervals with agreement by the total number of intervals and multiplying by 100. Procedural fidelity was calculated by dividing the number of procedural steps completed correctly by the total number of steps and multiplying by 100. Table 5 depicts the IOA and procedural fidelity coefficients for the pre-experimental assessments.

Experimental Assessments

Dependent Variables and Measurement

The primary dependent variable (DVs) assessed during the screener and hierarchy sensitivity assessments included button pressing (screener and arbitrary), and challenging behavior and appropriate behavior (clinically relevant). During the screener, arbitrary, and clinical assessments, two behaviors produced reinforcement. For the screener and arbitrary assessment, *Behavior 1* was defined as pressing the blue microswitch with enough force to produce the pre-recorded sound (Carson and Aliyah) or touching the blue button (Travis). *Behavior 2* was defined as pressing the red microswitch with enough force to produce the pre-recorded sound (Carson and Aliyah) or touching the red button with one or both hands (Travis). During the parameter hierarchy assessment with clinically relevant behaviors, B*ehavior 1* was challenging behavior and *Behavior 2* was appropriate behavior. *Challenging behavior* was defined in the same manner as in the functional analysis. For Travis and Carson, it was measured as duration and reported as percent of session (screener and arbitrary) or as cumulative trials during which the behavior occurred (clinically relevant). For Aliyah, it was measured as a count and reported as responses per minute (screener and arbitrary) or as cumulative trials during which the behavior occurred (clinically relevant). *Appropriate behavior* was defined as emitting a communicative response individually defined for each participant. For Travis, a vocal request was defined as saying a three-to-four-word phrase such as, “I want toys please.” For Carson, a vocal request was defined as saying the phrase, “I want a break please.” For Aliyah, a picture card request was defined as picking up the picture card and placing it in the researcher’s hand.

Appropriate behavior was measured as a count and reported as cumulative trials during which the behavior occurred (clinically relevant).

For both parameter hierarchy assessments, the researcher or trained graduate students collected data on the occurrence of Behavior 1 and Behavior 2 using the Countee application on a tablet (Piec & Hernández, 2016). The researcher calculated a cumulative record of response allocation by adding the number of trials the participant engaged in Behavior 1 or Behavior 2 each session to the number of trials the participant engaged in Behavior 1 or Behavior 2 in the preceding session.

Interobserver Agreement and Procedural Fidelity

An independent, trained second observer collected procedural fidelity data and data for the purpose of calculating interobserver agreement (IOA) on a range of 35% to 100% of sessions across all phases and participants. The researcher calculated IOA for the parameter sensitivity screeners and parameter hierarchy assessments using interval-by-interval agreement with 10 s intervals. An agreement was scored if data collectors scored the same occurrence or nonoccurrence of a behavior within the same 10 s interval. An agreement percentage was calculated by dividing the number of intervals with agreement by the total number of intervals and multiplying by 100. Procedural fidelity was calculated by dividing the number of procedural steps completed correctly by the total number of steps and multiplying by 100. Table 6 depicts the IOA and procedural fidelity coefficients for the experimental assessments.

Experimental Design

The researcher used a concurrent schedule, embedded in an ABAB reversal design to evaluate the effect of reinforcement parameters on two participant behaviors in the parameter hierarchy assessments. In phase A, one behavior resulted in delivery of a high magnitude/delayed reinforcement and one behavior resulted in delivery of an immediate/low magnitude reinforcement, both on a fixed-ratio (FR) 1 schedule. In phase B, the reinforcement schedule associated with each behavior was reversed. A functional relation was demonstrated when phase A resulted in stable response allocation, that response allocation was stable and opposite during phase B, and that effect was replicated across another implementation of both A and B.

Parameter Sensitivity Screening Procedures

The purpose of the parameter sensitivity screening was to determine if participants' behavior was sensitive to individual manipulations in magnitude, quality, and immediacy of reinforcement. Response allocation that favored pressing the button associated with high magnitude, high quality, or immediate reinforcement indicated sensitivity to that parameter. The researcher evaluated individual sensitivity using arbitrary behaviors (i.e., non-clinically relevant behavior) in a two-component concurrent schedule. Participants were presented with two colored buttons and instructed to choose one. Each button was associated with a different schedule of reinforcement. Following either button touch or switch press exhibited by the participant, the researcher provided the related consequence. Challenging behavior did not result in any programmed consequences; however, if challenging behavior occurred immediately preceding the delivery of the consequence for pressing the microswitch, the researcher implemented a 5 s change over delay and re-presented the choice. Sessions consisted of one trial. The trial began with the presentation of the button options and ended after the reinforcement interval. The researcher or trained graduate students collected data on participant response allocation for each trial. Table 7 provides a summary of the consequences and parameter manipulations for each phase of the parameter sensitivity assessments.

Exposure trials. Prior to each phase in the following comparisons, the researcher conducted six exposure trials (three with each button) following the procedures described by Kunnavatana et al. (2018). The researcher placed the button in front of the participant with the relevant consequence (i.e., high or low preferred stimulus and visual aid for reinforcement magnitude and immediacy) behind the item and prompted the participant to press the button. Once the participant pressed the button, the researcher provided the consequence associated with the button. An additional six exposure trials (three with each button) were conducted for Aliyah. For these trials, the researcher placed both buttons with relevant consequences in front of Aliyah and prompted her to press the button associated with a specific consequence while simultaneously pointing to the correct button. Once Aliyah pressed the correct button, the researcher provided the relevant consequence associated with the button.

Magnitude. The magnitude sensitivity assessment was conducted to determine if participants' behavior was sensitive to changes in the duration of reinforcement. The magnitude values selected were based on the study conducted by Kunnavatana et al. (2018). The high magnitude was 90 s and the low magnitude was 15 s. Quality and immediacy of reinforcement were kept constant throughout the magnitude manipulations. That is, a high-quality item was delivered immediately upon pressing a button.

Quality. The quality sensitivity assessment was conducted to determine if participants' behavior was sensitive to changes in the relative preference of the stimulus provided during reinforcement. The quality of an item was determined using the data collected from the preference assessment. Magnitude and immediacy of reinforcement were kept constant, so the item was delivered immediately for 30 s upon pressing a button.

Immediacy. The immediacy assessment was conducted to determine if participants' behavior was sensitive to changes in the time between engaging in a behavior and the delivery of the reinforcer. The immediate reinforcement was a 0 s delay and the delayed reinforcement was a 90 s delay. The magnitude and quality of reinforcement were kept constant at 30 s and high quality.

Parameter Hierarchy Assessment Procedures

The purpose of the experimental assessments was to compare the results from a parameter hierarchy assessment with arbitrary behavior to a parameter hierarchy assessment with clinically relevant behavior. The order of the parameter hierarchy assessments was counterbalanced across participants to control for sequencing effects. The parameter hierarchy assessments were conducted in a two-component concurrent schedule arrangement. Sessions consisted of 5 trials. A trial was the presentation of the choice options and the reinforcement interval following this presentation. Due to the length of the assessment and the clinical need to implement an intervention, sessions in the second phase A and phase B were decreased to 1 trial for Carson and Aliyah.

Arbitrary Behaviors. The arbitrary behavior used for this assessment was button pressing. For the comparison, there were two buttons available: a blue button and a red button. Each button was associated with a reinforcement schedule based on the phases described below. The researcher placed two buttons and visual aids depicting the appropriate consequence behind the buttons (i.e., colored poker chips for reinforcement magnitude and strips of laminated paper for reinforcement immediacy) in front of the participant. The researcher presented a contingency statement for the concurrent schedule (e.g., "If you press the blue button, you have to wait for your toy, but you will get it for a long time. If you press the red button, you can have your toy immediately but for a short time"). There were no programmed consequences for challenging behavior in this assessment; however, if challenging behavior occurred immediately preceding the delivery of the consequence for pressing the microswitch, the researcher implemented a 5 s change over delay.

Exposure trials. Prior to each phase in the following comparison, the researcher conducted six exposure trials (three with each microswitch). The researcher placed the microswitch in front of the participant with the relevant consequence behind the item and prompted the participant to press the button. Once the participant pressed the button with enough force to produce the pre-recorded output, the researcher provided the consequence associated with the button. An additional six exposure trials were conducted for Aliyah. For these trials, the researcher placed both buttons with relevant consequences in front of Aliyah and prompted her to press the button associated with a specific consequence while simultaneously pointing to the correct button. Once Aliyah pressed the correct button, the researcher provided the relevant consequence associated with the button.

Magnitude versus Immediacy. This comparison evaluated if the participants' behavior was more sensitive to magnitude or immediacy of reinforcement in the context of arbitrary behaviors. There were two consequences evaluated: immediate, low-magnitude reinforcement and high magnitude, delayed reinforcement. The magnitude and immediacy values were identical to the values used in parameter sensitivity screener. In phase A, the consequence for pressing the *blue microswitch/button* was an immediate, low magnitude reinforcement and the consequence for pressing the *red microswitch/button* was a high magnitude, delayed reinforcement for Travis and Aliyah. For Carson, the consequence for pressing the *red microswitch* was an immediate, low magnitude reinforcement and the consequence for pressing the *blue microswitch* was a high magnitude, delayed reinforcement. In phase B, these consequences were switched. A high-quality item was used to keep the quality constant.

Clinically Relevant Behavior Procedures. The parameter sensitivity assessment with clinically relevant behavior was conducted the same as the assessment with arbitrary behavior. However, for this assessment, the researcher presented the choice of engaging in clinically relevant behaviors. The clinically relevant behaviors used in this assessment were challenging behavior and communication. Each behavior was associated with a different schedule of reinforcement described below in the comparison procedures.

Functional Communication Training. Prior to the parameter hierarchy assessment with clinically relevant behaviors, the researcher used functional communication training (FCT) to teach the participant a communicative response. Prior to each session, the researcher presented a relevant contingency statement (e.g., "if you want your toys, you can hand me the picture card"). At the start of the session, the researcher set up the relevant antecedent (i.e., withheld tangible items or presented demands) and placed the communication modality (if applicable) in front of the participant. The researcher provided the relevant consequence for 30 s following the communicative response. The researcher or trained graduate students collected data on the target communicative response and challenging behavior. There were no programmed consequences for challenging behavior during FCT. However, if challenging behavior occurred immediately preceding the delivery of the consequence for engaging in the communicative response, the researcher implemented a 5 s change over delay. Sessions lasted 5 min and continued until three consecutive sessions with stable responding occurred.

Exposure trials. Prior to each phase in the following comparison, the researcher conducted three exposure trials. The researcher placed the laminated colored paper with the relevant consequence microswitch in front of the participant with the relevant consequence behind the item and prompted the participant to press the button. Once the participant pressed the button with enough force to produce the pre-recorded output, the researcher provided the consequence associated with the button. An additional six exposure trials were conducted for Aliyah. For these trials, the researcher placed both buttons with relevant consequences in front of Aliyah and prompted her to press the button associated with a specific consequence while simultaneously pointing to the correct button. Once Aliyah pressed the correct button, the researcher provided the relevant consequence associated with the button.

Magnitude versus Immediacy. This comparison evaluated if the participants' behavior was more sensitive to magnitude or immediacy of reinforcement in the context of clinically relevant behavior. There were two consequences evaluated: immediate, low magnitude reinforcement and high magnitude, delayed reinforcement. The magnitude and immediacy values were identical to the values used in parameter sensitivity screener. In phase A, the consequence for engaging in challenging behavior was immediate, low magnitude reinforcement and the consequence for engaging in the communicative response was high magnitude, delayed reinforcement. In phase B, these consequences were switched. A high-quality item was used to keep the quality constant.

Data Analysis

The researcher used visual analysis to analyze the data. Specifically, they analyzed the level, trend, and variability of the data paths. The data were depicted in a cumulative record. The researcher analyzed changes in the level, trend, and variability across phases. Changes in these variables of one schedule of reinforcement from phase A to phase B provide evidence of experimental control.

Alignment of Results. The researcher compared results from the parameter hierarchy assessment with arbitrary behaviors to the parameter hierarchy assessment with clinically relevant behavior. Alignment of results was scored as alignment, partial alignment, or no alignment. *Alignment* was scored if both assessments demonstrated the same parameter hierarchy with no larger than a 20% difference in cumulative response allocation. *Partial alignment* was scored if both assessments demonstrated the same parameter hierarchy with a greater than 20% difference in cumulative response allocation. *No alignment* was scored if there was no correspondence between the identified hierarchies.

Table 5

Summary of IOA and Procedural Fidelity for Pre-Experimental Assessments

Participants	Functional Analysis		Stimulus Preference Assessment		Reinforcer Assessment	
	IOA %	Fidelity %	IOA %	Fidelity %	IOA %	Fidelity %
Travis	97.2	100	100	100	95.8	100
Aliyah	99.7	100	100	100	95.9	100
Carson	N/A	N/A	N/A	N/A	98.9	100

Note. IOA = Interobserver Agreement

Table 6

Summary of IOA and Procedural Fidelity for Experimental Assessments

Participants	Parameter Sensitivity Screener		PHA with Arbitrary Behaviors		PHA with Clinically Relevant Behaviors	
	IOA %	Fidelity %	IOA %	Fidelity %	IOA %	Fidelity %
Travis	98.6	99	98.6	100	97.9	100
Aliyah	99.2	100	96.1	100	98.6	100
Carson	97.7	98.7	96.6	99	99.7	99

Note. IOA = Interobserver Agreement; PHA = Parameter Hierarchy Assessment

Table 7

Summary of Individual Parameter Sensitivity Screener

Reinforcement Parameter	Definition	Consequence 1	Consequence 2	Constant Parameters
Magnitude	Duration of reinforcement	High magnitude (90 s)	Low magnitude (15 s)	Immediate, high quality
Quality	Stimulus preference	High quality	Low quality	Immediate, 30 s magnitude
Immediacy	Time between behavior and reinforcement	Immediate	Delay (90 s)	30 s magnitude, high quality

CHAPTER 4

RESULTS

Preference Assessment

Figures 2 and 3 display the results of the preference assessments. Figure 2 displays the results of the response-restriction analysis conducted with Travis. During this assessment, the two items with the highest average percent of interaction were considered high-preferred and the two items with the lowest average percent of interaction were considered low-preferred. Travis interacted with the bubbles during 100% of the sessions in which it was presented and the cars during 50% of the sessions in which it was presented. Thus, the bubbles and cars were identified as high-preferred items. He interacted with Legos during 22.4% of the sessions in which it was presented and the ball during 17.6% of sessions in which it was presented. The Legos were identified as the low-preferred item. Due to the occurrence of challenging behavior when only the ball and coloring was available, both items were excluded.

Figure 3 displays the results of the paired-choice preference assessment conducted with Aliyah. During the paired-choice preference assessment, an item selected in more than 80% of trials was considered high-preferred and an item selected in 40% or less of trials was considered low-preferred. Aliyah selected the cars in 100% of the trials, so this item was identified as high-preferred. She selected the magnetic letters in 40% of trials, so this item was identified as low-preferred.

Reinforcer Assessment

Figures 4-6 display the results of the reinforcer assessment conducted with Travis, Aliyah, and Carson. In each figure, cumulative responses across 10 trials are displayed. The top left panel displays the results in low magnitude/high preferred phase, the top right panel displays the results in the high magnitude/high preferred phase, the bottom left panel displays the results in the low magnitude/low preferred phase, and the bottom right panel displays the results in the high magnitude/low preferred phase. Travis selected the task with a consequence 10 times each across all four phases (Figure 4). This result indicated that the high and low preferred stimuli functioned as reinforcers at both high and low magnitudes of reinforcement. Aliyah selected the task with a consequence seven times in the low magnitude/high preferred phase, 19 times in the high magnitude/high preferred phase, six times in the low magnitude/low preferred phase, and six times in the low magnitude/low preferred phase (Figure 5). The slope of the line depicting selection of the task with a consequence was steeper than the slope of the line depicting selection of the task without a consequence in all four phases, indicating that both the high and low preferred item functioned as reinforcers at high and low magnitudes of reinforcement. Carson selected the task with a consequence 10 times in the low magnitude/high preferred and high magnitude/high preferred phases and eight times in the low magnitude/high magnitude and high magnitude/low preferred phase (Figure 6). These results indicated that the high and low preferred stimuli functioned as reinforcers at both high and low magnitudes of reinforcement.

Functional Analysis

Figure 7 displays the results of the pairwise functional analysis conducted with Travis. The level of aggression was elevated in four of the six tangible test sessions. Travis did not exhibit aggression during five of the six control sessions. These results indicated that Travis' aggression was maintained by access to tangible items.

Figure 8 displays the results of the latency-based functional analysis conducted with Aliyah. The latency to aggression was between 103 and 130 s during three of the five tangible sessions. Aggression never occurred during sessions from the other test conditions or during control sessions. These results indicated that Aliyah's aggression was maintained by access to tangible items.

Parameter Sensitivity Screener

Figures 9-11 display the results of the parameter sensitivity screener conducted with Travis, Aliyah, and Carson. In each figure, the top panel displays cumulative responses during the magnitude sensitivity screener, the middle panel displays cumulative responses during the quality sensitivity screener, and the bottom panel displays cumulative responses during the immediacy sensitivity screener.

Figure 9 displays the results of the screener for Travis. During the magnitude screener, Travis' response allocation favored high magnitude reinforcement. In the first A phase, Travis selected high magnitude reinforcement in all three sessions. In the first B phase, he selected low magnitude reinforcement in two sessions and high magnitude reinforcement in four sessions. In the subsequent A and B phases, Travis selected high magnitude reinforcement in all three sessions. In the quality screener, Travis' response allocation favored low quality reinforcement across all sessions. In the immediacy screener, his response allocation favored immediate reinforcement across all sessions. These data suggested Travis' behavior was sensitive to changes in magnitude and immediacy of reinforcement. His behavior was not sensitive to the quality of reinforcement in the way quality was defined (i.e., based on the results of the preference assessment).

Figure 10 displays the results of the screener for Aliyah. During the magnitude screener, Aliyah's response allocation favored high magnitude reinforcement. In the first A phase, Aliyah initially selected low magnitude reinforcement. However, her response allocation shifted to high magnitude reinforcement following additional exposure trials. In the subsequent phases, response allocation favored high magnitude reinforcement. During the quality screener, Aliyah's response allocation was undifferentiated between high and low quality reinforcement. In the first A phase and first B phase, Aliyah's response allocation favored high quality reinforcement. However, in the second A phase, response allocation was undifferentiated. Following additional exposure trials, Aliyah's response allocation favored low quality reinforcement. During the immediacy screener, Aliyah's response allocation favored immediate reinforcement. In the first A phase, Aliyah selected immediate reinforcement in all three sessions. In the first B phase, Aliyah selected immediate reinforcement in four sessions and delayed reinforcement in one session. Aggression occurred during the delay in the session that she selected delayed reinforcement. This response allocation replicated in subsequence phases. These results indicated that Aliyah's behavior was sensitive to changes in the magnitude and immediacy of reinforcement. However, her behavior was not sensitive to changes in the quality of reinforcement.

Figure 11 displays the results of the screener for Carson. During the magnitude screener, Carson's response allocation favored high magnitude reinforcement across all sessions. During the quality screener, his response allocation was undifferentiated. In the first A phase, Carson selected high quality reinforcement in all three sessions. However, in the first B phase, response allocation was undifferentiated. During the immediacy screener, Carson's response allocation favored immediate reinforcement across all sessions. These results suggested that Carson's behavior was sensitive to changes in the magnitude and immediacy of reinforcement. However, his behavior was not sensitive to changes in the quality of reinforcement.

Parameter Hierarchy Assessments

Figures 12-14 displays the results of the parameter hierarchy assessments conducted with Travis, Aliyah, and Carson. In each figure, the top panel displays cumulative responses during the parameter hierarchy assessment conducted with arbitrary behaviors and the bottom panel displays cumulative responses during the parameter hierarchy assessment conducted with clinically relevant behaviors.

Figure 12 displays the results of the parameter hierarchy assessments for Travis. There was no functional relation between response allocation and change in the reinforcement parameters in both parameter hierarchy assessments. During the assessment with arbitrary behaviors, Travis' initially allocated responding to the button associated with immediate, low magnitude reinforcement in the first A and B phases. This finding was not replicated in subsequent phases. In the second A and B phases, Travis allocated responding exclusively to the button associated with the high magnitude, delayed reinforcement. No challenging behavior occurred during the parameter hierarchy assessment with arbitrary behaviors. During the parameter hierarchy assessment with clinically relevant behaviors, Travis' response allocation favored aggression, regardless of the change in reinforcement parameters. This assessment ended with only three sessions in the second phase B due to the increase in intensity and frequency of aggression. Based on the results of these assessments, a parameter hierarchy was not identified for arbitrary or clinically relevant behaviors.

Figure 13 displays the results of the parameter hierarchy assessments for Aliyah. During the assessment with arbitrary behaviors, Aliyah's response allocation favored the microswitch associated with immediate, low magnitude reinforcement across all phases. In the first A phase, Aliyah allocated responding to the microswitch associated with immediate, low magnitude reinforcement on 23 trials and the microswitch associated with high magnitude, delayed on two trials. In the first B phase, she selected the microswitch associated with immediate, low magnitude reinforcement on 19 trials and the microswitch associated with high magnitude, delayed on six trials. In the second A phase, she selected the microswitch associated with the immediate, low magnitude on all five trials. In the second B phase, she selected the microswitch associated with immediate, low magnitude reinforcement on six trials and the microswitch associated with high magnitude, delayed on two trials. Aliyah engaged in low rates of aggression (0.17 and 0.29 responses per minute) in two sessions during the assessment with arbitrary behaviors. During the assessment with clinically relevant behaviors, there was no functional relation between response allocation and the change in reinforcement parameter. Aliyah's response allocation favored the functional communicative response regardless of changes in the reinforcement parameter. In the first A phase, she engaged in aggression (associated with immediate, low magnitude reinforcement) on two trials and engaged in the functional communicative response (associated with high magnitude, delayed reinforcement) on 22 trials. Aliyah engaged in the functional communicative response on the remaining trials in the subsequent phases. Based on the results of these assessments, the parameter hierarchy identified for arbitrary behaviors was immediacy then magnitude. There was no alignment between the results of the assessments with arbitrary and clinically relevant behaviors.

Figure 14 displays the results of the parameter hierarchy assessments for Carson. During the assessment with arbitrary behaviors, Carson's response allocation favored the microswitch associated with immediate, low magnitude reinforcement. In the first A phase, Carson selected the microswitch associated with immediate, low magnitude reinforcement on 24 trials and the microswitch associated with high magnitude, delayed on one trial. In the first B phase, he selected the microswitch associated with immediate, low magnitude reinforcement on 24 trials and the microswitch associated with high magnitude, delayed on five trials. In the second A phase, he selected the microswitch associated with immediate, low magnitude reinforcement on all five trials. In the second B phase, he selected the microswitch associated with immediate, low magnitude reinforcement on six trials and the microswitch associated with high magnitude, delayed on one trial. No challenging behavior occurred during the assessment with arbitrary behaviors. During the assessment with clinically relevant behaviors, there was no functional relation between response allocation and the change in reinforcement parameter. Carson's response allocation favored the appropriate communicative response regardless of changes in the reinforcement parameter. In the first A phase, he engaged in screaming (associated with immediate, low magnitude reinforcement) on four trials and engaged in the functional communicative response (associated with high magnitude, delayed reinforcement) on 21 trials. Carson engaged in the functional communicative response on the remaining trials in the subsequent phases. Based on the results of these assessments, the parameter hierarchy identified for arbitrary behaviors was immediacy, followed by magnitude. There was no alignment between the results of the assessments with arbitrary and clinically relevant behaviors.

Figure 2

Travis's Response Restriction Analysis Results

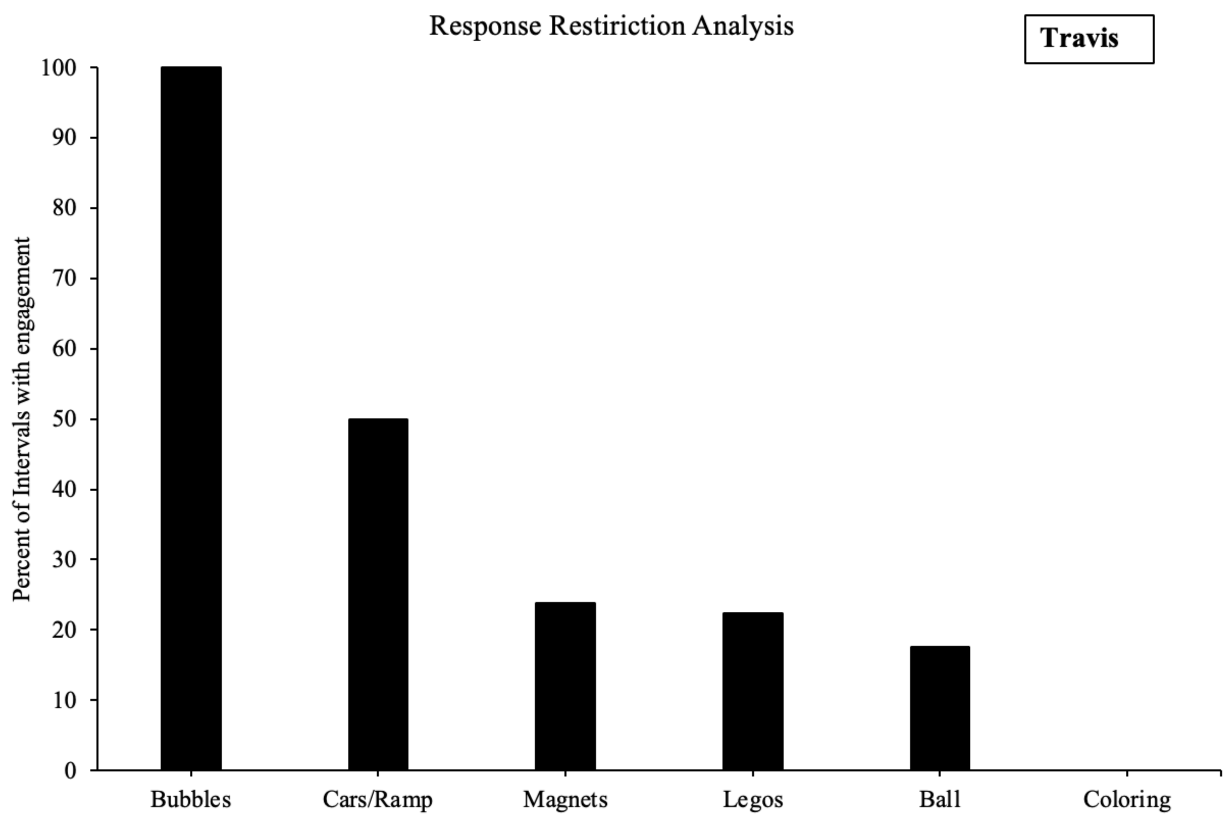


Figure 3

Aliyah's Paired-Choice Preference Assessment Results

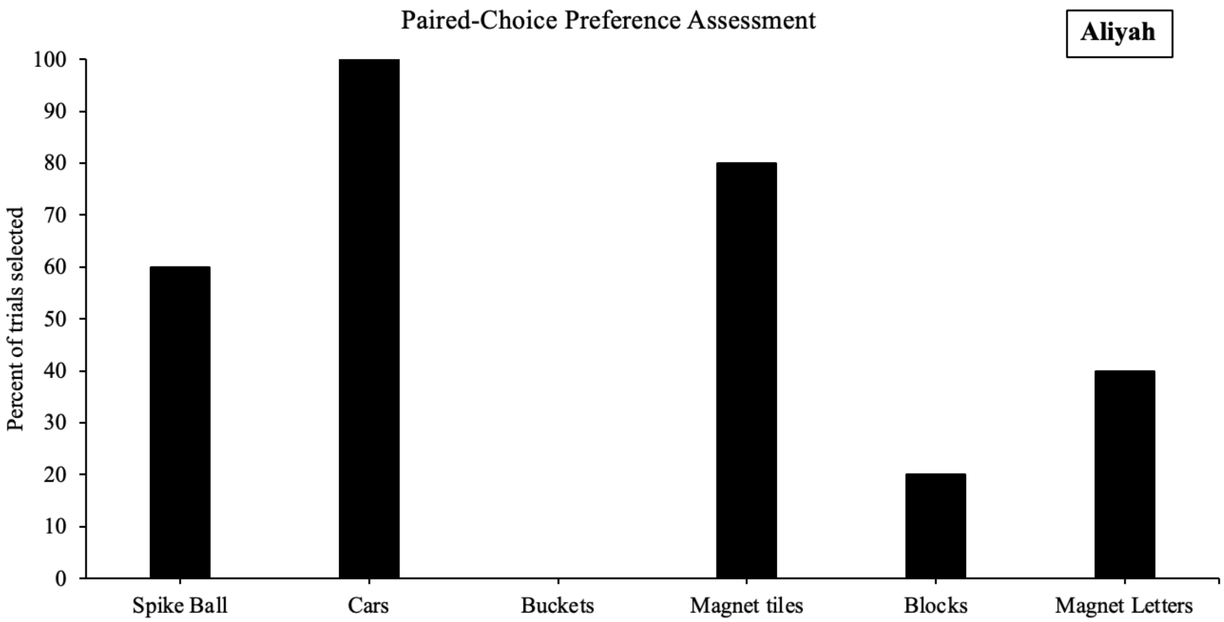
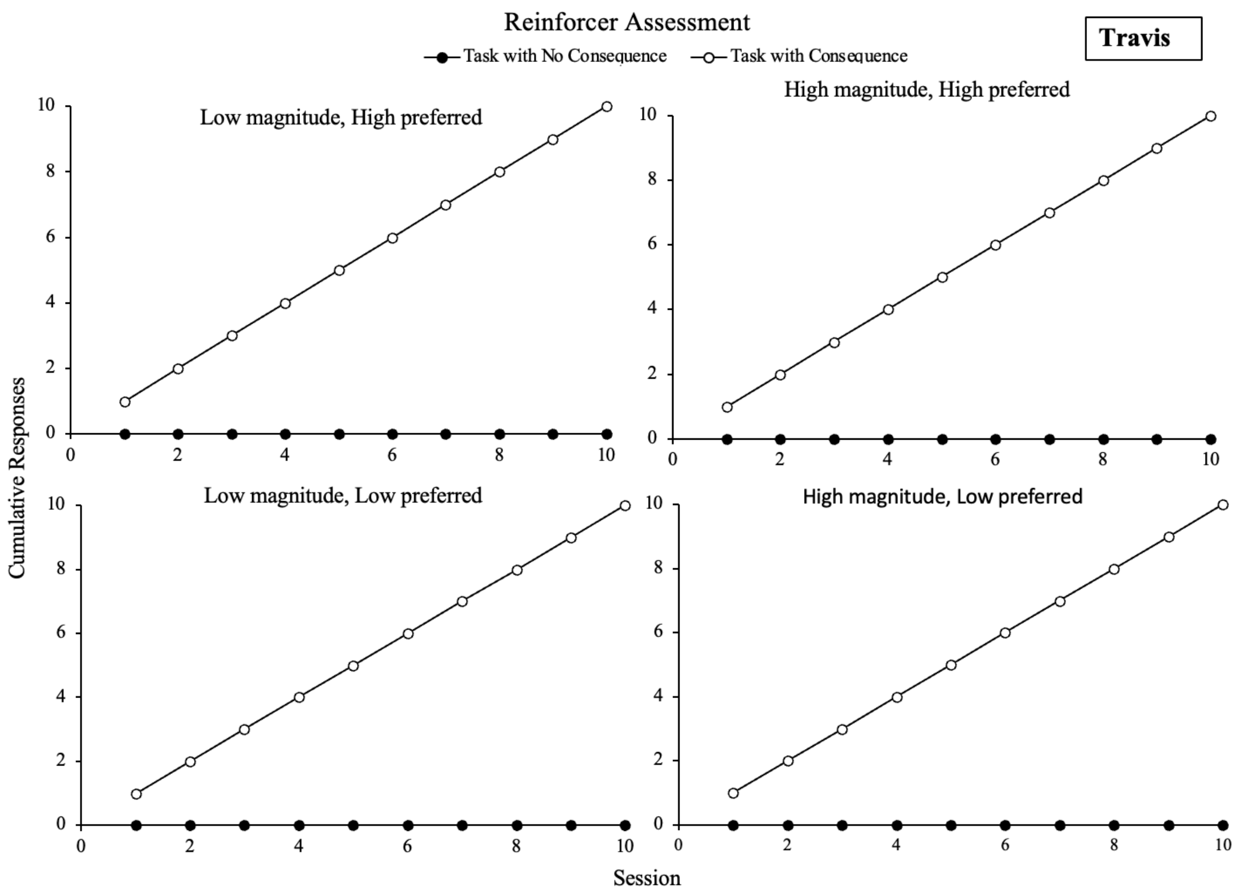


Figure 4

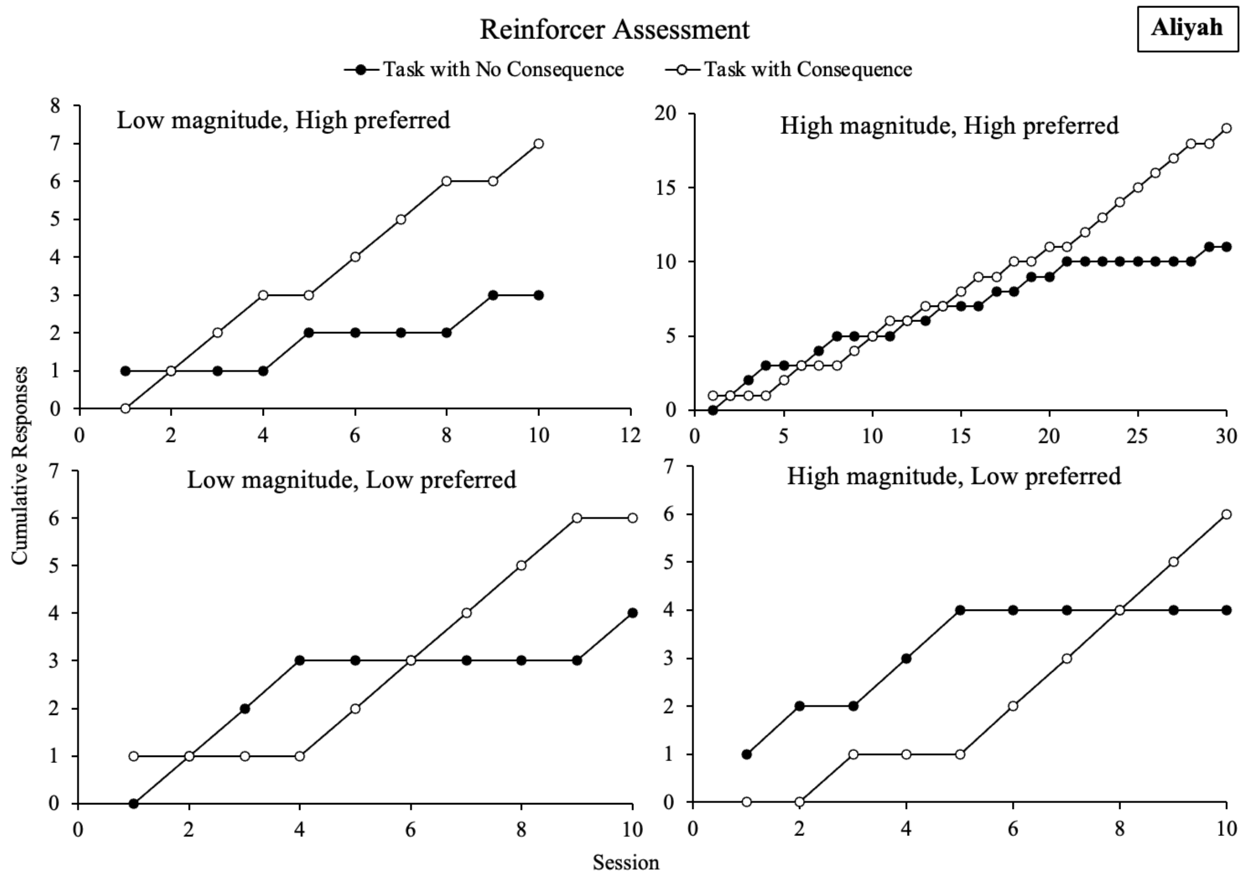
Travis's Reinforcer Assessment Results



Note. The top left panel displays the results of the low magnitude/high preferred assessment; the top right panel displays the results of the high magnitude/high preferred assessment; the bottom left panel displays the results of the low magnitude/low preferred assessment; and the bottom right panel displays the results of the high magnitude/low preferred assessment.

Figure 5

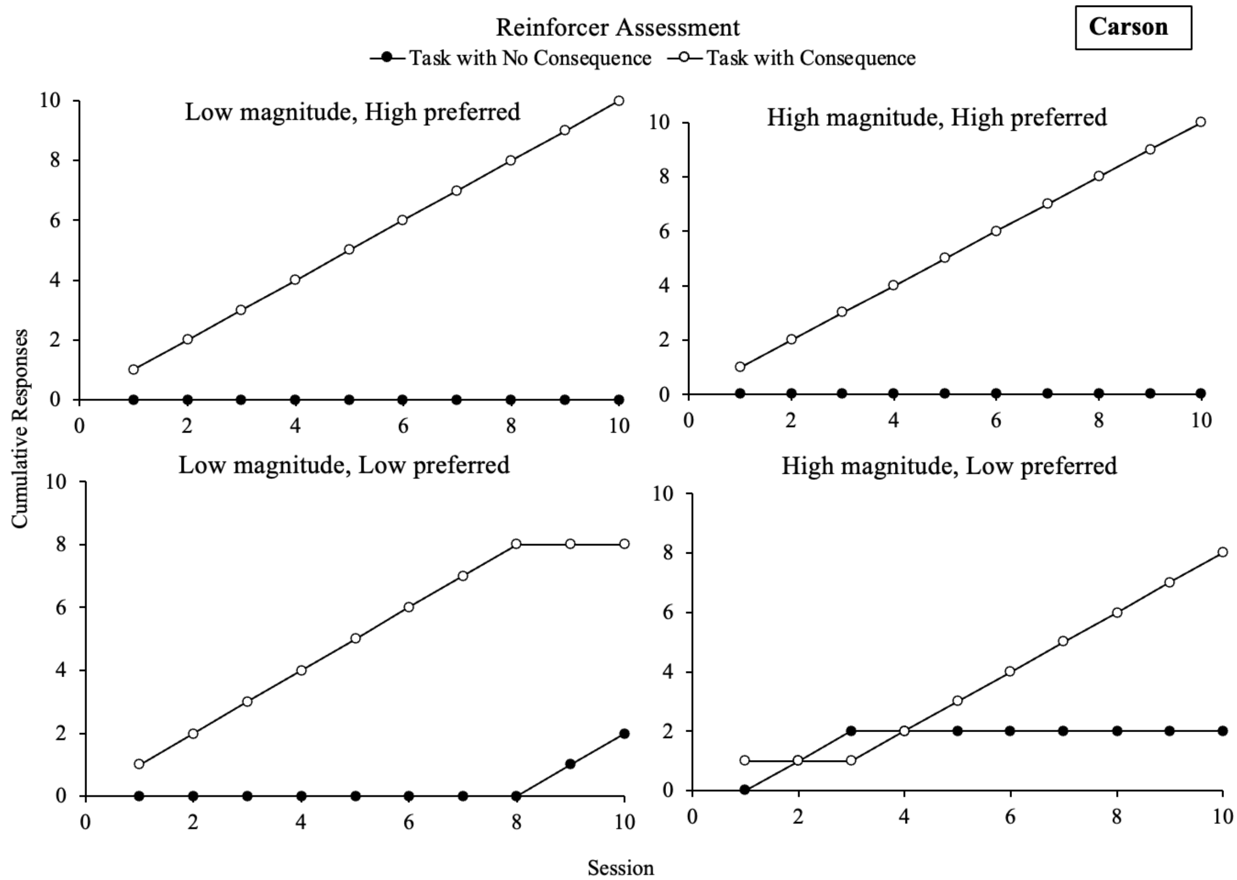
Aliyah's Reinforcer Assessment Result



Note. The top left panel displays the results of the low magnitude/high preferred assessment; the top right panel displays the results of the high magnitude/high preferred assessment; the bottom left panel displays the results of the low magnitude/low preferred assessment; and the bottom right panel displays the results of the high magnitude/low preferred assessment.

Figure 6

Carson's Reinforcer Assessment Results



Note. The top left panel displays the results of the low magnitude/high preferred assessment; the top right panel displays the results of the high magnitude/high preferred assessment; the bottom left panel displays the results of the low magnitude/low preferred assessment; and the bottom right panel displays the results of the high magnitude/low preferred assessment.

Figure 7

Travis's Functional Analysis Results

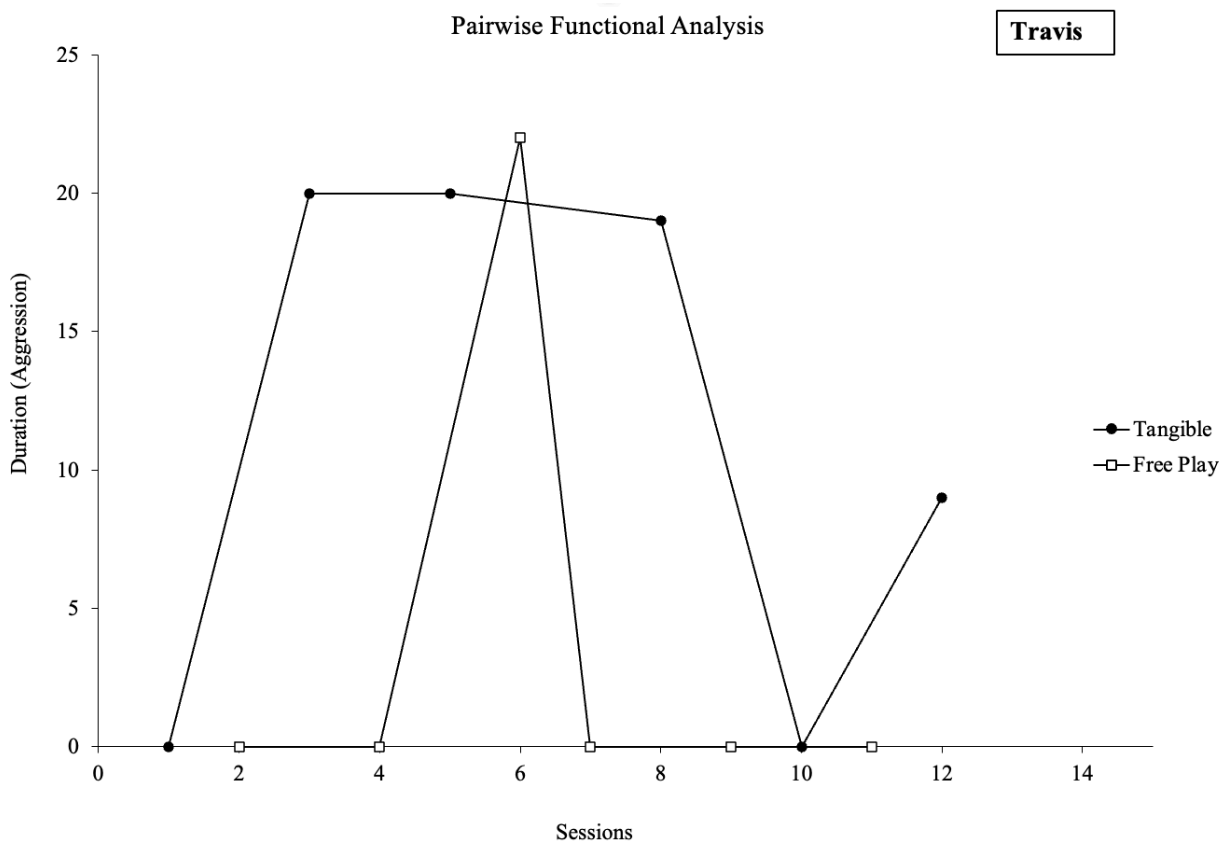


Figure 8

Aliyah's Functional Analysis Results

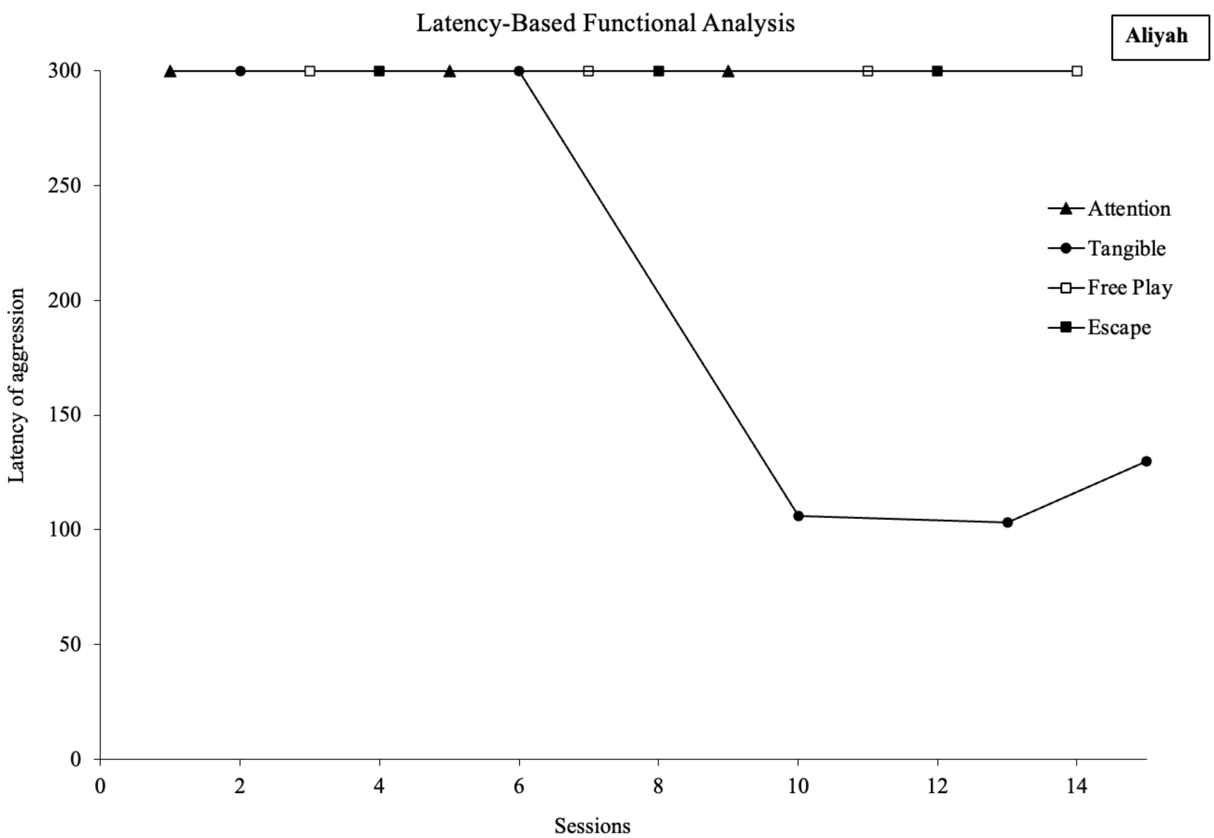
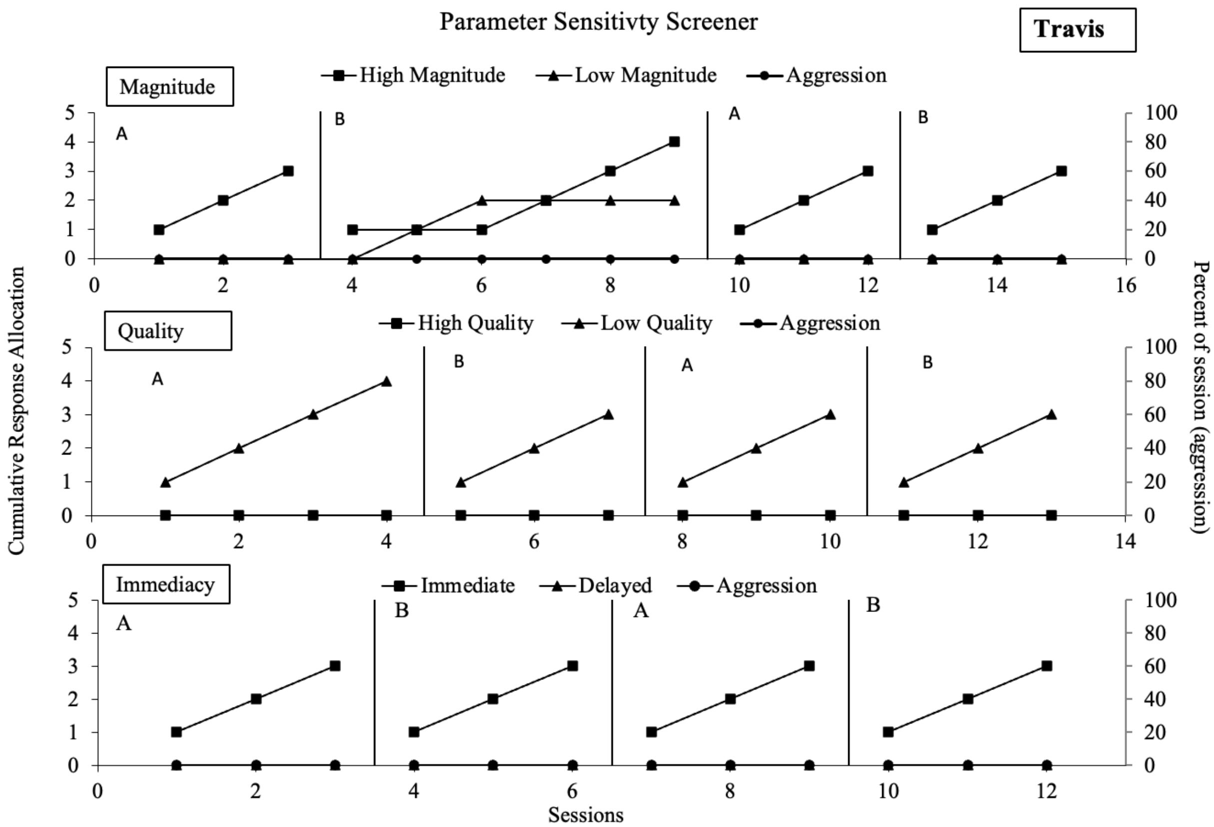


Figure 9

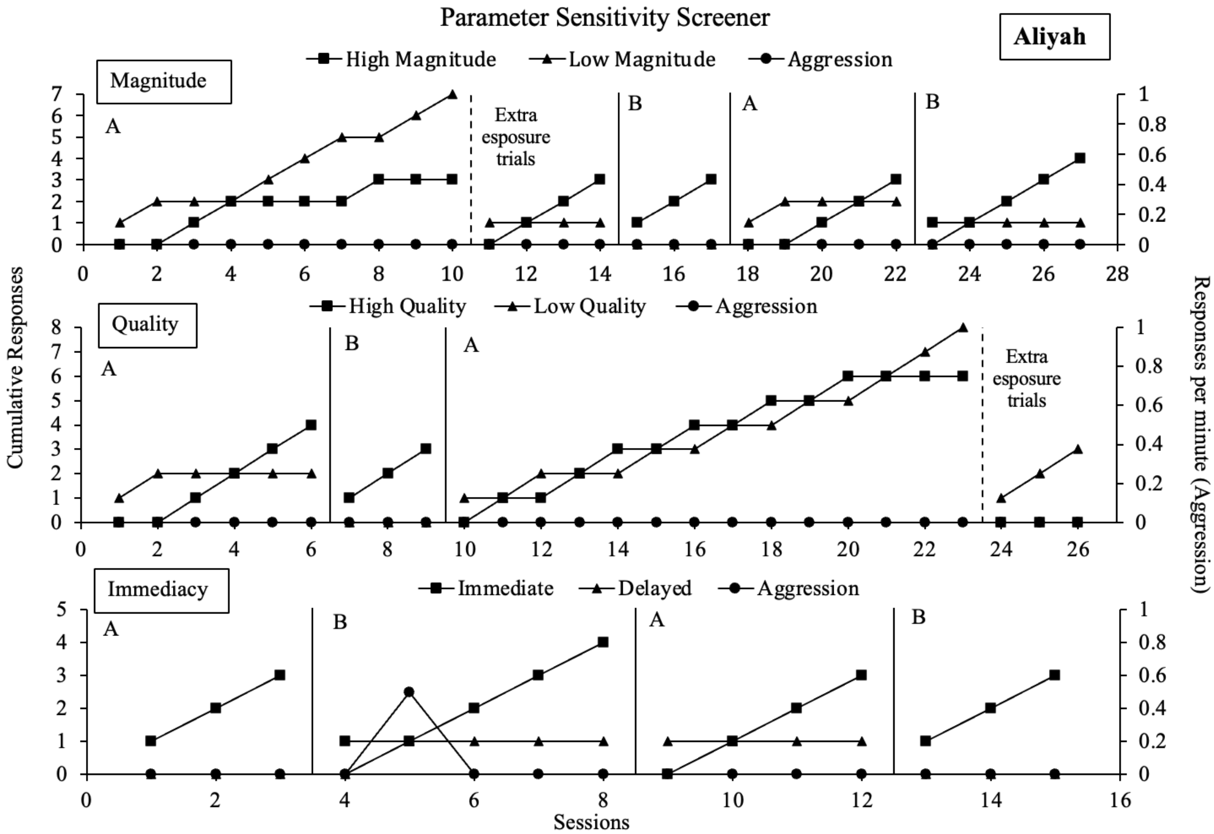
Travis's Parameter Sensitivity Screener Results



Note. The top panel displays the results of the magnitude screener; the middle panel displays the results of the quality screener; and the bottom panel displays the results of the immediacy screener.

Figure 10

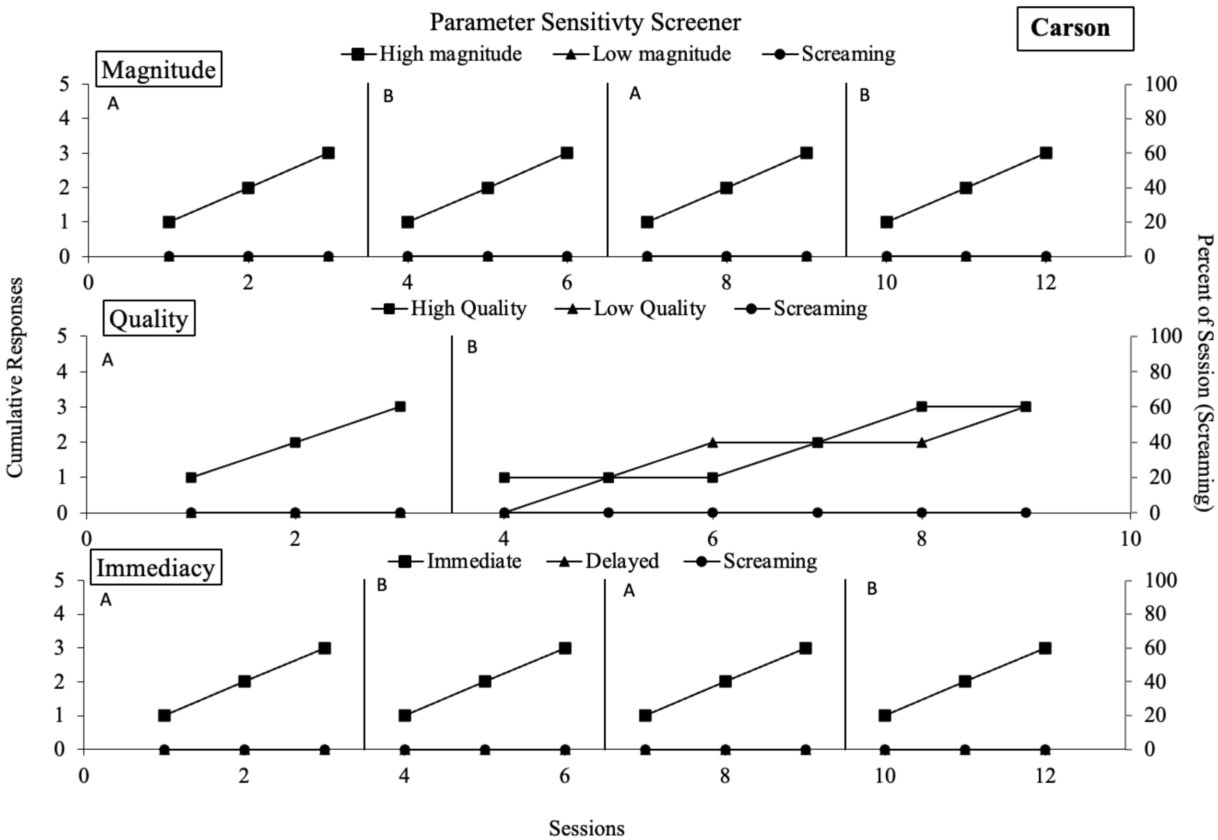
Aliyah's Parameter Sensitivity Screener Results



Note. The top panel displays the results of the magnitude screener; the middle panel displays the results of the quality screener; and the bottom panel displays the results of the immediacy screener.

Figure 11

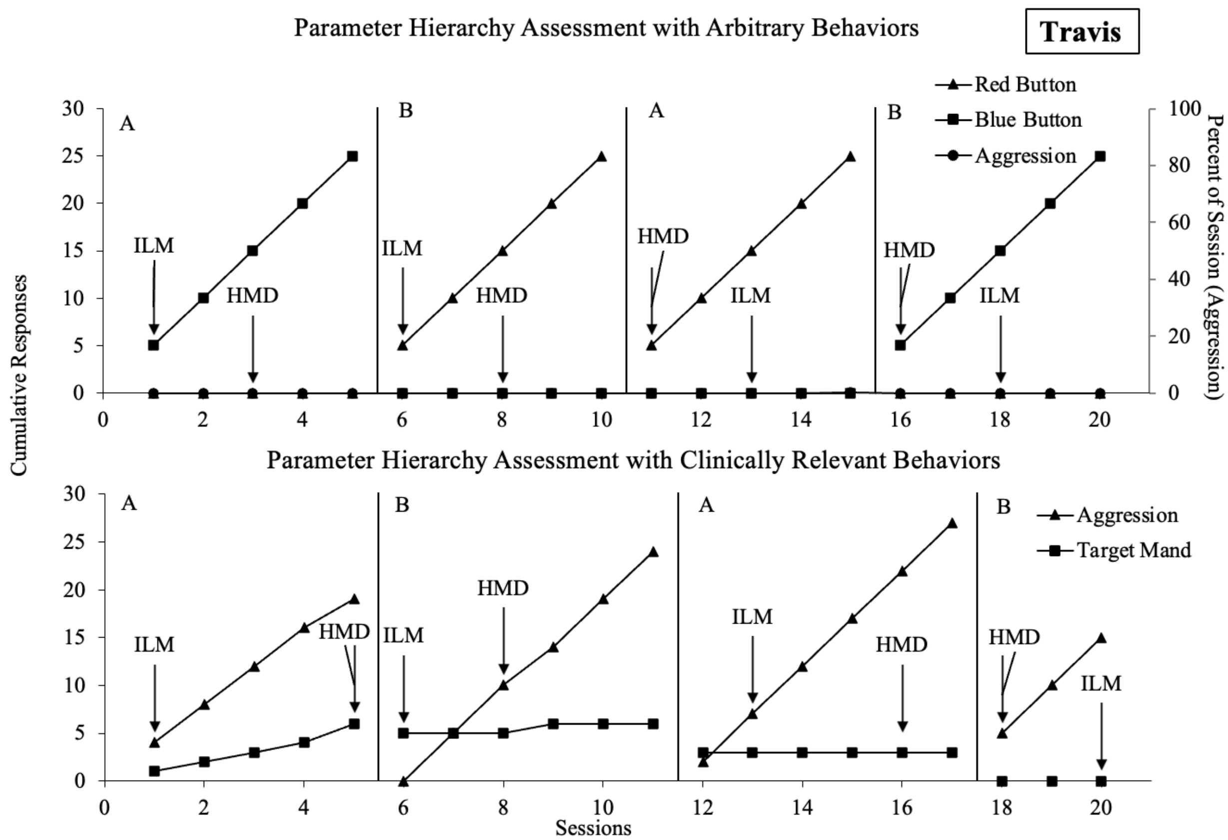
Carson's Parameter Sensitivity Screener Results



Note. The top panel displays the results of the magnitude screener; the middle panel displays the results of the quality screener; and the bottom panel displays the results of the immediacy screener.

Figure 12

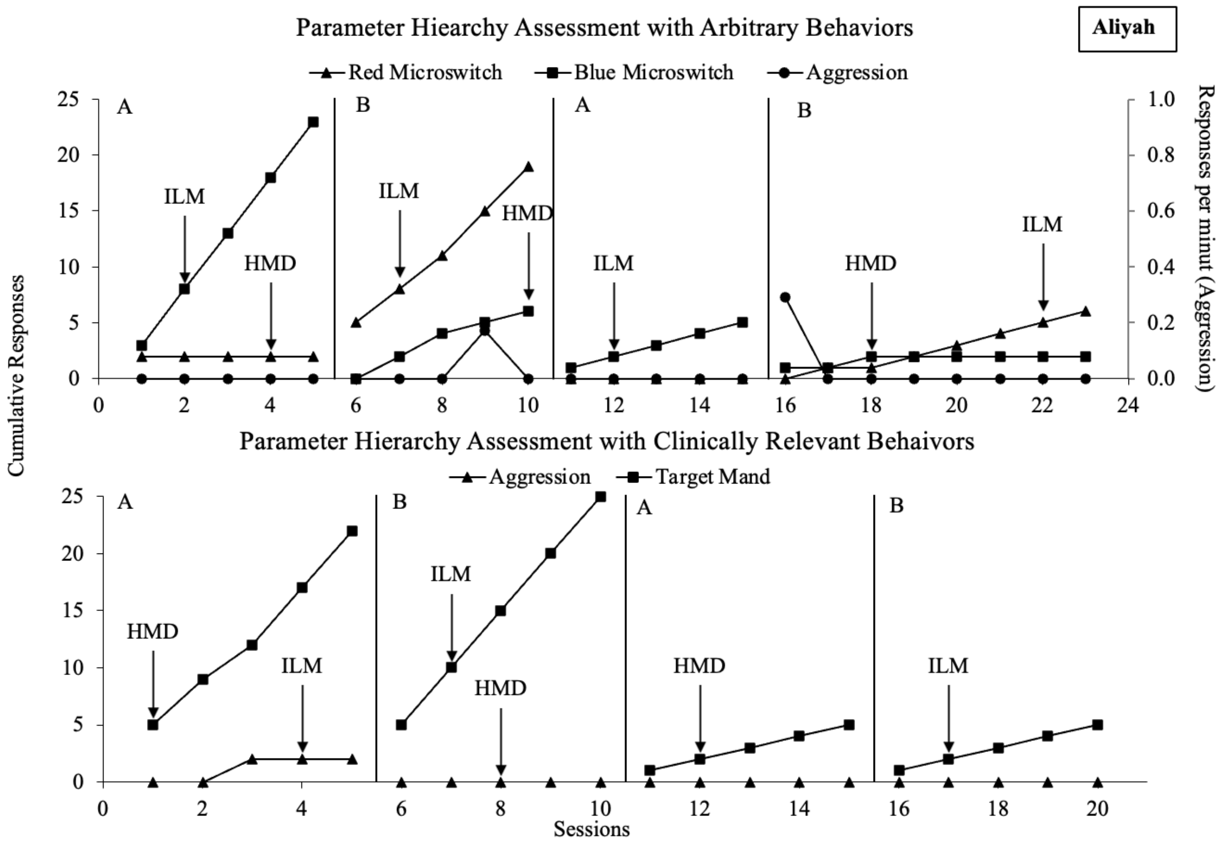
Travis's Parameter Hierarchy Assessment Results



Note. The top panel displays the results of the parameter hierarchy assessment with arbitrary behaviors; and the bottom panel displays the results of the parameter hierarchy assessment with clinically relevant behaviors. ILM = Immediate, low magnitude reinforcement; HMD = High magnitude, delayed reinforcement.

Figure 13

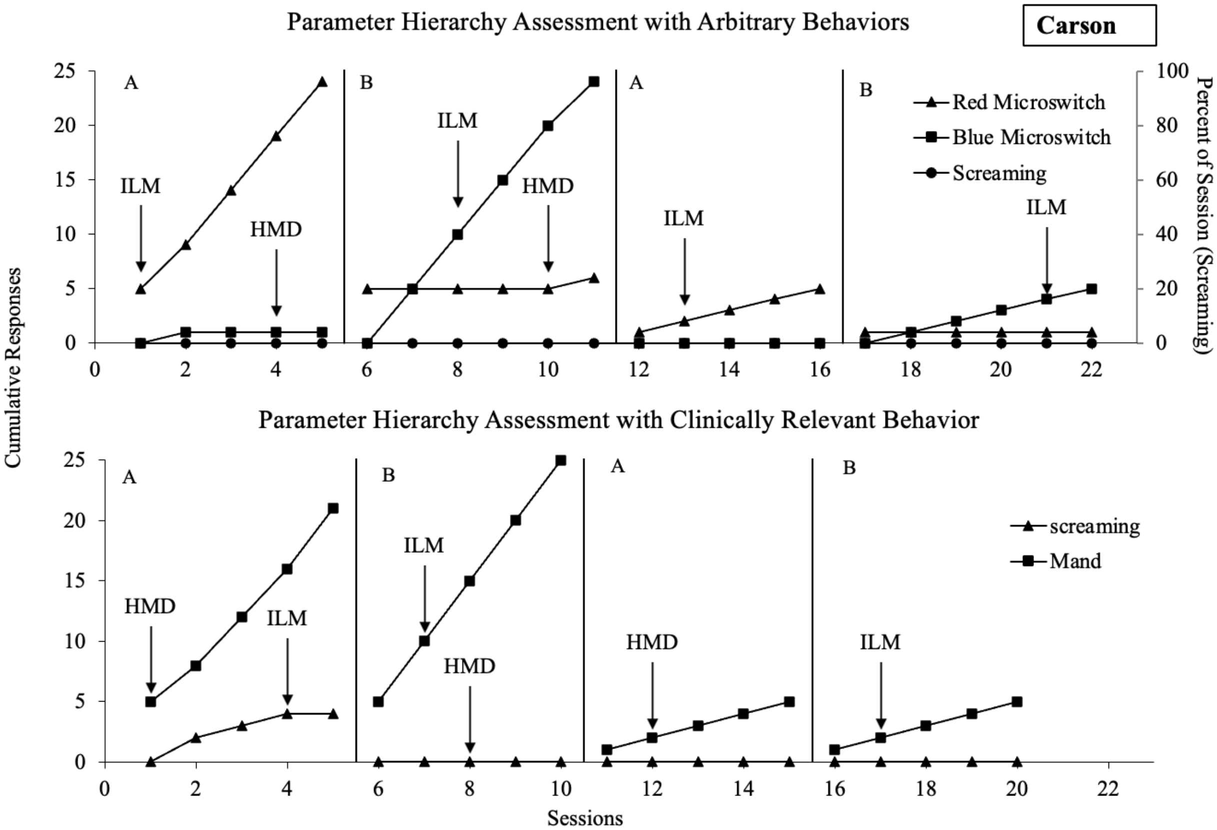
Aliyah's Parameter Hierarchy Assessment Results



Note. The top panel displays the results of the parameter hierarchy assessment with arbitrary behaviors; and the bottom panel displays the results of the parameter hierarchy assessment with clinically relevant behaviors. ILM = Immediate, low magnitude reinforcement; HMD = High magnitude, delayed reinforcement.

Figure 14

Carson's Parameter Hierarchy Assessment Results



Note. The top panel displays the results of the parameter hierarchy assessment with arbitrary behaviors; and the bottom panel displays the results of the parameter hierarchy assessment with clinically relevant behaviors. ILM = Immediate, low magnitude reinforcement; HMD = High magnitude, delayed reinforcement.

CHAPTER 5

DISCUSSION

The current study evaluated a) if a parameter hierarchy can be identified for both arbitrary and clinically relevant behaviors and b) if results from a parameter hierarchy assessment with arbitrary behaviors and a parameter hierarchy assessment with clinically relevant behaviors aligned. Alignment of results would provide evidence supporting the use of an assessment with arbitrary behaviors to inform concurrent schedule-based interventions to reduce challenging behaviors.

Prior to conducting the parameter hierarchy assessments, the researcher conducted a parameter sensitivity screener to determine which parameters to include in the hierarchy assessment. This screener was a replication of Experiment 1 in the study conducted by Kunnavatana et al. (2018). In their study, all three participants' behavior demonstrated sensitivity to reinforcement quality. Two of the three participants' behavior demonstrated sensitivity to magnitude and immediacy of reinforcement. The results from the current screener study partially replicated the results from Kunnavatana et al. (2018) Experiment 1. In the current study, all three participants' behavior was sensitive to changes in the magnitude and immediacy of reinforcement. However, none of the participants' behavior was sensitive to reinforcement quality in the way "sensitivity to quality" was defined. Sensitivity to quality was defined as response allocation favoring the response associated with the high preferred item when the low and high preferred item were concurrently available contingent on similar responses. For two participants (Aliyah and Carson), there was no difference in response allocation across the response options associated with the high and low preferred items. One participant (Travis) allocated responses exclusively to the response option associated with the low preferred (i.e, low quality) item. These results relative to reinforcement quality did not align with the outcomes of the preference assessment and suggested that relative item preference was transient and/or not able to be determined by preference assessments for these participants. Due to the results of the parameter sensitivity screener, the manipulation of quality was omitted from the parameter hierarchy assessments for all participants during the parameter hierarchy assessments.

The parameter hierarchy assessment with arbitrary behaviors showed a clear parameter hierarchy for two of the three participants. Aliyah and Carson allocated more responding to the response options associated with immediate reinforcement, relative to high magnitude reinforcement. These results suggested that an effective intervention to increase appropriate behavior and decrease challenging behavior for Aliyah and Carson should include assigning immediate reinforcement for the appropriate behavior. For Travis, response allocation initially favored immediate reinforcement. However, this finding was not replicated in subsequent phases. A potential explanation for Travis's outcome was that his sensitivity to parameters was transient. It was possible that the parameter that impacted response allocation one day did not impact response allocation on subsequent days. This result may have been due to uncontrolled environmental variables that affected the reinforcement parameters. For example, if there was a strong motivating operation to escape demands, an individual might have selected the high magnitude, delayed reinforcement because there was a longer intertrial reinforcement interval than the immediate, low magnitude reinforcement. On the other hand, if there was a strong motivating operation to access the tangible item, an individual might have selected the immediate, low magnitude reinforcement to access the item quicker. Berg et al. (2000)

demonstrated pre-session exposure to attention impacted the outcomes of functional analyses of challenging behavior. Specifically, they demonstrated that less challenging behavior occurred during an attention test condition of a functional analysis when the session was preceded by a noncontingent attention free play session than when the session was preceded by an escape or alone test session. These results suggest that variables other than the variables manipulated in the session impact responding. Future research should look at how extra-experimental variables that may impact motivating operations change reinforcement parameter sensitivities.

The parameter hierarchy assessment with clinically relevant behaviors did not identify a hierarchy for any participant. During this assessment, Aliyah and Carson engaged in the appropriate communicative response regardless of the reinforcement parameter manipulation associated with communication or challenging behavior. This finding suggested that the results of the parameter sensitivity screener could be used to design successful treatment for Aliyah and Carson. During the screener, both participants' behavior was sensitive to changes in magnitude and immediacy, and both parameters were effective at maintaining the appropriate communicative response in the assessment with clinically relevant behaviors. However, the results of the screener were not useful in designing a successful treatment for Travis. In the parameter hierarchy assessment with clinically relevant behaviors, Travis engaged in aggression regardless of the reinforcement parameter manipulation associated with communication or challenging behavior. This result demonstrated that changes in magnitude or immediacy did not shift Travis's response allocation toward a functional communicative response. One explanation for this finding may be the reinforcement history for challenging behavior. Travis was referred to the clinic for the treatment of aggression in the school and home environment. The occurrence of aggression across multiple contexts suggested there was an extensive reinforcement history for aggression. This history may have set the occasion for Travis to bias responding toward aggression, even if it produced delays to reinforcement because it was the strongest response in the response class. Future research should look into the impact of reinforcement history on response allocation when reinforcement parameters are manipulated.

Based on these results, the results of the parameter hierarchy assessments with arbitrary behaviors and clinically relevant behaviors did not align for any of the participants. For two of the three participants (Aliyah and Carson), a parameter hierarchy was identified for arbitrary behaviors but not for clinically relevant behaviors. One explanation for this difference in outcomes across assessments may be the impact of reinforcement history and response effort on response allocation. In the assessment with arbitrary behaviors, reinforcement history and response effort were held constant. Prior to the experiment, the same number of exposure trials were conducted with each response, and it was unlikely that engagement in these responses had been reinforced prior to or were reinforced outside of the study context, resulting in relatively similar reinforcement histories for both responses. Additionally, both responses required the participant to press a microswitch with the same amount of force to produce the pre recorded output. This similarity kept the response effort constant for each response. In the assessment with clinically relevant behaviors, the reinforcement histories and the response effort differences were unknown. Both participants engaged in challenging behavior and similar communicative responses outside of the study context. It was likely that both behaviors received reinforcement outside of the study context. For this reason, the reinforcement history for each response was unknown and uncontrolled. The response effort required to engage in challenging behavior and the communicative response may not be equal and was difficult to control for. The lack of control for these variables may be the reason a hierarchy for clinically relevant behavior was not identified for Aliyah and Carson. Despite the lack of alignment, the results of the current study indicated that manipulations in reinforcement schedules that aligned with sensitivity assessments resulted in response allocation towards appropriate behavior for Aliyah and Carson. For one participant (Travis), a parameter hierarchy was not identified for arbitrary or clinically relevant behavior. As discussed previously, this may be due to other environmental variables that were uncontrolled. For this reason, an alignment of outcomes was not possible to assess.

Practical Implications

As mentioned in previous chapters, identifying successful interventions that do not include an extinction component is an important area of research given (a) the difficulties with implementing extinction components with high fidelity, and (b) the potential for increases in severity and frequency of challenging behavior when extinction is encountered. From a practical standpoint, extinction may be dangerous to implement with certain behaviors and may be difficult to implement with fidelity (St. Peter Pipkin et al., 2010; Vollmer et al., 2020). When fidelity decreases below 50%, especially at the beginning of the intervention, treatment success declines (St. Peter Pipkin et al., 2010). Lerman et al. (1999) noted that extinction bursts and aggressive behavior was a side effect for nearly one half of the published cases of interventions with an extinction component they reviewed. While the presence of other treatment components reduced the likelihood of such outcomes, extinction bursts and other side effects were still noted in some instances.

Inclusion of extinction may also be problematic from an acceptability standpoint. School personnel, caregivers, and the individuals we serve may not find the inclusion of an extinction component acceptable. For this reason, more research is needed to determine effective non-extinction-based interventions. The review of the literature provided in Chapter 2 of this document identified studies whose results suggested manipulating reinforcement parameters may have produced desirable decreases in challenging behavior without the use of extinction. However, the idiosyncratic effects of concurrent schedule-based interventions in the literature suggest more research is needed to examine ways to identify effective parameter manipulations. Kunnavatana et al. (2018) identified one way to evaluate individual sensitivity to reinforcement parameter manipulations. Researchers assessed individual and relative parameter sensitivity in assessments with arbitrary behaviors (i.e., button pressing) and designed an intervention for challenging behavior based on the results of the assessment. Although the intervention was effective for all three participants, the assessment does not follow the logic of other assessments used to design interventions to reduce challenging behavior. Usually, assessments conducted prior to intervention for challenging behavior include contingency arrangements that provide consequences for the clinically relevant challenging behavior. However, Kunnavatana et al. (2018) conducted assessments using arbitrary behaviors as a proxy for challenging behavior.

The current study compared the results of a parameter hierarchy assessment with arbitrary behaviors to the same assessment with clinically relevant behaviors. The results of the assessments indicate no alignment between the two assessments. This suggests that an assessment with arbitrary behaviors might not inform effective interventions for clinically relevant behavior. Due to this, more research is necessary to evaluate effect ways to arrange non-extinction-based interventions.

Limitations

Several limitations need to be considered when evaluating the results of the current study. First, the parameter sensitivity screener conducted prior to the experiments was conducted with arbitrary behaviors. The purpose of the screener was used to determine if participants could identify changes in the reinforcement parameters prior to conducting the parameter hierarchy assessments. However, as mentioned previously, and as demonstrated by the comparison of the two parameter hierarchy assessments, it is unknown whether the parameters that impact arbitrary behavior will impact clinically relevant behavior in the same way. The results of the screener may have differed if it had been conducted with clinically relevant behaviors.

A second limitation relates to the definition of the quality parameter. In the current study and in the literature, quality is defined relative to the preference of the stimulus/stimuli. Although preference assessments are often used to identify potential reinforcers, the results of the preference assessment did not predict the relative reinforcing value of each stimulus. Specifically, it is possible that an item identified as high preferred in a preference assessment has a lower reinforcing value than an item identified as low preferred. This potentially may explain the lack of sensitivity to reinforcement quality demonstrated across participants in the current study. This limitation may be broader than just the impact on the current study. It may be the case that preference assessment outcomes can be arranged in a hierarchy, but the hierarchy is not relevant to (a) the relative potency of the stimuli as reinforcers or (b) the stability of that hierarchy over time.

Third, only two reinforcer parameters were assessed in the parameter hierarchy assessments. The quality manipulation was excluded from the parameter hierarchy assessment due to the lack of sensitivity in the screener conducted with each participant. Although it was necessary to exclude the quality manipulation in the current study, a parameter hierarchy assessment with only two parameters limits the ability to identify a meaningful parameter hierarchy. For Travis specifically, it is possible that a successful arrangement might have been identified if more reinforcement parameters were included in the assessments.

A fourth limitation is that implementer error resulted in a deviation from procedural protocol for one participant. Specifically, the FCT phase prior to the parameter hierarchy assessment with clinically relevant behaviors for Carson did not include the number of intended sessions with independent communicative responding. This oversight may have impacted the results of the parameter hierarchy assessment for clinically relevant behavior. However, Carson's response allocation during the assessment favored the functional communicative response, even though the FCT phase was truncated. These results demonstrated Carson's proficiency with the functional communicative response despite the error in FCT implementation.

A fifth limitation relates to the exposure trials conducted in the parameter hierarchy assessment. During the clinically relevant parameter hierarchy assessment, participants were only exposed to the contingencies for FCT prior to the experiment. This decision was made due to the difficulty and impracticality for researchers and practitioners to conduct exposure trials with challenging behavior. Only exposing participants to the contingency for FCT might explain Aliyah's and Carson's results. Both participants' response allocation favored the functional communicative response in the assessment with clinically relevant behaviors regardless of reinforcement schedule. This outcome may be an example of a carryover effect from the exposure trials prior to each condition.

Future Research

Although there are several limitations in the current study, the results may have implications for future conceptual and practical research. From a conceptual standpoint, this direction for future research cuts across research domains related to preference assessment as well as assessment designed to identify relevant reinforcer parameters to include in concurrent schedule-based interventions. For example, preference assessment outcomes could be evaluated

(a) to determine if hierarchies predict potency and (b) for stability of outcomes. Roane et al. (2001) used a progressive ratio assessment to assess the reinforcer potency of stimuli identified as high preferred in a paired-choice preference assessment for four individuals with developmental disabilities. They found that the stimuli had different potency levels despite being equally preferred in the preference assessment. These results suggest that hierarchies may not predict potency for all individuals. MacNaul et al. (2021) conducted a systematic literature review on the stability of preference assessment outcomes. After reviewing 20 studies, they found that preference assessment outcomes are most stable when conducted no more than 30 days apart. These results suggest that individual preference for stimuli should be reassessed at least monthly. The lack of reinforcer potency prediction and stability of preference assessments have implications for assessing quality of reinforcement. The current definition of quality relates to the relative preference of stimuli identified using a preference assessment. If the preference hierarchy identified in a stimulus preference assessment does not predict reinforcer potency (Carter & Zonneveld, 2020; Graff & Larson, 2011; Roane et al., 2001), researchers should examine whether changes in sensitivity to quality can be explained by reinforcer potency. Additionally, if preference for stimuli is not stable longer than a month, manipulating quality of reinforcement in a concurrent schedule arrangement may not be effective long term. Future research should examine ways to increase the long-term effectiveness of such interventions given the lack of preference stability. This might look like reevaluating the quality of reinforcement at multiple points throughout the intervention (i.e., once a month).

Future research on parameter hierarchy assessments should also evaluate the potential transient properties of parameter sensitivity. In the current study, Travis's response allocation in the assessment with arbitrary behaviors initially favored the immediate, low magnitude reinforcement. However, this response allocation shifted to high magnitude, delayed reinforcement following a month-long break from the study. This shift in response allocation might be because sensitivity is transient. Future research should assess the stability of parameter hierarchies over time with more participants to determine if parameter sensitivity is truly transient. If sensitivity is transient, future research should examine the effects of motivating operations on parameter sensitivity. This might look like researchers systematically manipulating environmental variables immediately prior to a parameter sensitivity assessment.

From a practical standpoint, the assessments in the current study were time consuming to conduct. The parameter sensitivity screener took one to three hours per parameter (total of three to nine hours) to conduct and the parameter hierarchy assessment both took five to seven hours to conduct. When developing interventions for challenging behavior, the goal is to identify an effective intervention as quickly as possible. Lengthy assessments will delay the implementation of interventions and may result in loss of services or community access for the individual engaging in challenging behavior. For this reason, future research should focus on ways to refine and develop streamlined assessments.

Summary and Conclusion

The goal of the current study was to compare the result of an assessment conducted with arbitrary behaviors (replicating Experiment 2 in Kunnavatana, et al., 2018) to the results of the same type of assessment conducted with clinically relevant behavior. In the current study, there was no alignment of outcomes between the two assessments. This outcome may have occurred because the parameters that impacted arbitrary behavior did not impact clinically relevant behaviors in the same way. Despite the lack of alignment, a successful intervention may have been identified for two of the three participants. Aliyah's and Carson's response allocation favored the appropriate behavior regardless of reinforcement schedule. This result suggested that an intervention for challenging behavior based on the assessment with arbitrary behaviors would likely be successful for Aliyah and Carson. However, for Travis, an intervention based on the results of the assessment with arbitrary behaviors would likely not be successful. In the assessment with clinically relevant behavior, his response allocation favored challenging behavior. These results indicated that the parameters identified by the parameter sensitivity screener did not impact clinically relevant behavior in a meaningful way. Given the noted limitations related to practicality, side effects, and social validity of interventions for challenging behavior that include an extinction component and the lack of uniformity of the effect of various parameters, both singly and in combination, for any given individual's behavior, research on identifying successful parameter arrangements is needed to enhance the likelihood that non-extinction based interventions will be maximally effective and specific to individual behavior sensitivities.

References

References marked with an asterisk (*) were included in the chapter 2 literature review.

*Adelinis, J. D., Piazza, C. C., & Goh, H.-L. (2001). Treatment of multiply controlled destructive behavior with food reinforcement. Journal of Applied Behavior Analysis, 34(1), 97–100. https://doi.org/10.1901/jaba.2001.34-97

*Athens, E. S., & Vollmer, T. R. (2010). An investigation of differential reinforcement of alternative behavior without extinction. *Journal of Applied Behavior Analysis, 43*(4), 569–589. https://doi.org/10.1901/jaba.2010.43-569

Baum W. M. (1974). On two types of deviation from the matching law: bias and undermatching. *Journal of the Experimental Analysis of Behavior*, *22*(1), 231–242. https://doi.org/10.1901/jeab.1974.22-231

Berg, W. K., Peck, S., Wacker, D. P., Harding, J., McComas, J., Richman, D., & Brown, K. (2000). The effects of presession exposure to attention on the results of assessments of attention as a reinforcer. *Journal of Applied Behavior Analysis*, *33*(4), 463–477. https://doi.org/10.1901/jaba.2000.33-463

*Berth, D. P., Bachmeyer, M. H., Kirkwood, C. A., Mauzy, C. R., IV, Retzlaff, B. J., & Gibson, A. L. (2019). Noncontingent and differential reinforcement in the treatment of pediatric feeding problems. *Journal of Applied Behavior Analysis*, 52(3), 622–641. https://doi.org/10.1002/jaba.562

*Bishop, M. R., Kenzer, A. L., Coffman, C. M., Tarbox, C. M., Tarbox, J., & Lanagan, T. M. (2013). Using stimulus fading without escape extinction to increase compliance with toothbrushing in children with autism. Research in Autism Spectrum Disorders, 7, 680-686. https://doi.org/10.1016/j.rasd.2013.02.004

*Borrero, C. S., Vollmer, T. R., Borrero, J. C., Bourret, J. C., Sloman, K. N., Samaha, A. L., & Dallery, J. (2010). Concurrent reinforcement schedules for challenging behavior and appropriate behavior: experimental applications of the matching law. *Journal of the Experimental Analysis of Behavior, 93*(3), 455–469. https://doi.org/10.1901/jeab.2010.93-455

*Boyle, M. A., Bacon, M. T., Brewer, E. M., Carton, S. M., & Gaskill, L. A. (2020). Evaluating a treatment without extinction for elopement maintained by access to stereotypy. Journal of Applied Behavior Analysis, 53(3), 1531–1541. https://doi.org/10.1002/jaba.682

*Briggs, A. M., Dozier, C. L., Lessor, A. N., Kamana, B. U., & Jess, R. L. (2019). Further investigation of differential reinforcement of alternative behavior without extinction for escape-maintained destructive behavior. *Journal of Applied Behavior Analysis, 52*(4), 956–973. https://doi.org/10.1002/jaba.648

*Brown, K. R., Greer, B. D., Craig, A. R., Sullivan, W. E., Fisher, W. W., & Roane, H. S. (2020). Resurgence following differential reinforcement of alternative behavior implemented with and without extinction. Journal of the Experimental Analysis of Behavior, 113(2), 449–467. https://doi.org/10.1002/jeab.588

Carr, E. G., and Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18*(2), 11-126. https://doi.org/10.1901/jaba.1985.18-111

Carter, A. B. & Zonneveld, K. L. M. (2020). A comparison of displacement and reinforcer potency for typically developing children. *Journal of Applied Behavior Analysis, 53*(2), 1130-1144. https://doi.org/10.1002/jaba.636

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

Corte, H. E., Wolf, M. M., Locke, B. J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis, 4*(3), 201-213. https://doi.org/10.1901/jaba.1971.4-201

Cuvo, A. J., Lerch, L. J., Leurquin, D. A., Gaffaney, T. J., & Poppen, R. L. (1998) Response allocation to concurrent fixed-ratio reinforcement schedules with work requirements by adults with mental retardation and typical preschool children. *Journal of Applied Behavior Analysis, 31*(1), 43-63. https://doi.org/10.1901/jaba.1998.31-43

*Davis, D. H., Fredrick, L. D., Alberto, P. A., & Gama, R. (2012). Functional communication training without extinction using concurrent schedules of differing magnitudes of reinforcement in classrooms. *Journal of Positive Behavior Interventions, 14*(3), 162–172. https://doi.org/10.1177/1098300711429597

*Davis, T. N., Weston, R., Hodges, A., Uptegrove, L., Williams, K., & Schieltz, K. M. (2018). Functional communication training and demand fading using concurrent schedules of reinforcement. *Journal of Behavioral Education, 27*(3), 343–357. https://doi.org/10.1007/s10864-017-9289-0

DeLeon, I. G., Fisher, W. W., Rodriguez-Catter, V., Maglieri, K., Herman, K., & Marhefka, J. M. (2001). Examination of relative reinforcement effects of stimuli identified through pretreatment and daily brief preference assessments. Journal of Applied Behavior Analysis, 34(4), 463-473. https://doi.org/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*(4), 519–533. https://doi.org/10.1901/jaba.1996.29-519

*Dowdy, A., & Tincani, M. (2020). Assessment and treatment of high-risk challenging behavior of adolescents with autism in an aquatic setting. Journal of Applied Behavior Analysis, 53(1), 305–314. https://doi.org/10.1002/jaba.590

*Dowdy, A., Tincani, M., Nipe, T., & Weiss, M. J. (2018). Effects of reinforcement without extinction on increasing compliance with nail cutting: A systematic replication. Journal of Applied Behavior Analysis, 51, 924-930. https://doi.org/10.1002/jaba.484

Emerson, E. (2001). *Challenging behaviour: Analysis and intervention in people with learning difficulties*. Cambridge: University Press

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*(2), 491–498. https://doi.org/10.1901/jaba.1992.25-491

Graff, R.B., & Larsen, J. (2011). The relation between obtained preference value and reinforcer efficacy. *Behavioral Interventions*, *26*(2), 125-133. https://doi.org/10.1002/bin.325

Hagopian, L.P., Long, E.S., & Rush, K.S. (2004) Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28*(5), 668-677. https://doi.org/10.1177/0145445503259836

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis, 30*(3), 459–473. https://doi.org/10.1901/jaba.1997.30-459

Herrnstein R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*(3), 267–272. https://doi.org/10.1901/jeab.1961.4-267

Hoch, H., McComas, J. J., Johnson, L., Faranda, N., & Guenther, S. L. (2002a). The effects of magnitude and quality of reinforcement on choice responding during play activities. *Journal of Applied Behavior Analysis*, *35*(2), 171–181. https://doi.org/10.1901/jaba.2002.35-171

*Hoch, H., McComas, J. J., Thompson, A. L., & Paone, D. (2002b). Concurrent reinforcement schedules: Behavior change and maintenance without extinction. *Journal of Applied Behavior Analysis, 35*(2), 155–169. https://doi.org/10.1901/jaba.2002.35-155

Horner, R. H., & Day, H. M. (1991). The effects of response efficiency on functionally equivalent competing behaviors. Journal of Applied Behavior Analysis, 24(4), 719-732. https://doi.org/10.1901/jaba.1991.24-719

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994a). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197–209. https://doi.org/10.1901/jaba.1994.27-197 (Reprinted from Analysis and Intervention in Developmental Disabilities, 2, 3–20, 1982)

Iwata, B. A., Duncan, B. A., Zarcone, J. R., Lerman, D. C., & Shore, B. A. (1994b). A sequential, test-control methodology for conducting functional analyses of self-injurious behavior. *Behavior Modification, 18*(3), 289-306. https://doi.org/10.1177/01454455940183003

Iwata, B. A., Pace, G. M., Kalsher, M. J., Cowder, G. E., & Cataldo, M. F. (1990). Experimental analysis and extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis, 23*(1), 11-27, https://doi.org/10.1901/jaba.1990.23-11

*Kelley, M. E., Lerman, D. C., & Van Camp, C. M. (2002). The effects of competing reinforcement schedules on the acquisition of functional communication. *Journal of Applied Behavior Analysis, 35*(1), 59–63. https://doi.org/10.1901/jaba.2002.35-59

*Kunnavatana, S. S., Bloom, S. E., Samaha, A. L., Slocum, T. A., & Clay, C. J. (2018). Manipulating parameters of reinforcement to reduce challenging behavior without extinction. *Journal of Applied Behavior Analysis, 51*, 283–302. https://doi.org/10.1002/jaba.443.

*Lalli, J. S., Vollmer, T. R., Progar, P. R., Wright, C., Borrero, J., Daniel, D., Barthold, C. H., Tocco, K., & May, W. (1999). Competition between positive and negative reinforcement in the treatment of escape behavior. *Journal of Applied Behavior Analysis, 32*(3), 285–296. https://doi.org/10.1901/jaba.1999.32-285

Ledford, J. R. & Gast, D. L. (Eds.). (2018). *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed.). New York, NY: Routledge

Ledford, J. R., Chazin, K. T., Lane, J. D., Zimmerman, K. N., & Ayres, K. A. (2020). Single case analysis and review framework (SCARF). Retrieved from: http://ebip.vkcsites.org/scarfv2

Lerman, D. C. & Iwata, B. A. (1995). Prevalence of the extinction burst and its attenuation during treatment. *Journal of Applied Behavior Analysis, 28*(1), 93-94. https://doi.org/10.1901/jaba.1995.28-93

Lerman, D. C., Iwata, B. A., & Wallace, M. D. (1999). Side effects of extinction: Prevalence of bursting and aggression during the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis, 32,* 1–8. http://dx.doi.org/10.1901/jaba.1999.32-1

MacNaul, H., Cividini-Motta, C., Wilson, S., & DiPaola, H. (2020). A systematic review of research on stability of preference assessment outcomes across repeated administrations. *Behavior Interventions, 36*(4), 962-983. https://doi.org/10.1002/bin.1797

*Martens, B. K., Lambert, T. L., Sullivan, W. E., Magnuson, J. D., Morley, A. J., Sallade, S. J., & Baxter, E. L. (2016). Choice in transition: Replication and extension to preschool children in a naturalistic setting. *Journal of the Experimental Analysis of Behavior, 105*(2), 307–321. https://doi.org/10.1002/jeab.201

Neef, N. A., & Lutz, M. N. (2001). A brief computer-based assessment of reinforcer dimensions affecting choice. *Journal of Applied Behavior Analysis, 34*(1), 57-60. https://doi.org/10.1901/jaba.2001.34-57

Neef, N.A., Shade, D., & Miller, M. S. (1994). Assessing influential dimensions of reinforcers on choice in students with serious emotional disturbance. *Journal of Applied Behavior Analysis, 27*(4), 575-583. https://doi.org/10.1901/jaba.1994.27-575

*Peterson, S. M., Frieder, J. E., Smith, S. L., Quigley, S. P., & Van Norman, R. K. (2009). The effects of varying quality and duration of reinforcement on mands to work, mands for break, and challenging behavior. *Education & Treatment of Children, 32*(4), 605–630. https://doi.org/10.1353/etc.0.0075

Petscher, E. S., Rey, C., & Bailey, J. S. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities, 30*(3), 409-425, https://doi.org/10.1016/j.ridd.2008.08.008

*Piazza, C. C., Fisher, W. W., Hanley, G. P., Remick, M. L., Contrucci, S. A., & Aitken, T. L. (1997). The use of positive and negative reinforcement in the treatment of escape-maintained destructive behavior. *Journal of Applied Behavior Analysis, 30*(2), 279–298. https://doi.org/10.1901/jaba.1997.30-279

Peic, D., & Hernández, V. (2016). Countee- data collection system for behavioral data (Version 2.2.1) [Mobile application software]. Retrieved from http://itunes.apple.com

Poppes, P., van der Putten, A. J. J., & Vlaskamp, C. (2010). Frequency and severity of challenging behavior in people with profound intellectual disabilities. Research in Developmental Disabilities, 31, 1269– 1275. https://doi.org/10.1016/j.ridd.2010.07.017

Ringdahl, J. E., Berg, W. K., Wacker, D. P., Ryan, S., Ryan, A., Crook, K., & Molony, M. (2016). Further demonstrations of individual preference among mand modalities during functional communication training. *Journal of Developmental and Physical Disabilities, 28*(6), 905–917. https://doi.org/10.1007/s10882-016-9518-9

Roane, H. S., Lerman, D. C., & Vorndran, C. M. (2001). Assessing reinforcers under progressive schedule requirements. *Journal of applied behavior analysis*, *34*(2), 145–166. https://doi.org/10.1901/jaba.2001.34-145

*Rogalski, J. P., Roscoe, E. M., Fredericks, D. W., & Mezhoudi, N. (2020). Negative reinforcer magnitude manipulations for treating escape-maintained challenging behavior. *Journal of Applied Behavior Analysis, 53*(3), 1514–1530. https://doi.org/10.1002/jaba.683

St. Peter Pipkin, C, Vollmer, T. R., & Sloman, K. N. (2010). Effects of treatment integrity failures during differential reinforcement of alternative behavior: A translational model. *Journal of Applied Behavior Analysis, 43*(1), 47-70. https://doi.org/10.1901/jaba.2010.43-47

*Stuesser, H. A., & Roscoe, E. M. (2020). An evaluation of differential reinforcement with stimulus fading as an intervention for medical compliance. Journal of Applied Behavior Analysis, 53(3), 1606–1621. https://doi.org/10.1002/jaba.685

Task Force Promoting Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *Clinical Psychology*, 48, 3–23.

Thomason-Sassi, J. L., Iwata, B. A., Neidert, P. L., Roscoe, E. M. (2011). Response latency as an index of response strength during functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 44*(1), 51-67. https://doi.org/10.1901/jaba.2011.44-51

Tincani, M. & Travers, J. (2019). Replication research, publication bias, and applied behavior analysis. *Perspectives on Behavior Science, 42*(1), 59-75. https://doi.org/10.1007/s40614-019-00191-5

Trump, C. E., Ayres, K. M., Quinland, K. K., & Zabala, K. A. (2020). Differential reinforcement without extinction: A review of the literature. *Behavior Analysis: Research and Practice, 20*(2), 94–107. https://doi.org/10.1037/bar0000169

Vollmer, T. R., Peters, K. P., Kronfli, F. R., Lloveras, L. A., & Ibañez, V. F. (2020). On the definition of differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis, 53*(3), 1299-1303. https://doi.org/10.1002/jaba.701

*Worsdell, A. S., Iwata, B. A., Hanley, G. P., Thompson, R. H., & Kahng, S. W. (2000). Effects of continuous and intermittent reinforcement for challenging behavior during functional communication training. *Journal of Applied behavior Analysis, 33*(2), 167–179. https://doi.org/10.1901/jaba.2000.33-167