MAINTAINING REDUCTIONS IN CHALLENGING BEHAVIOR FOLLOWING REINFORCEMENT-BASED INTERVENTION WITH SCHEDULE THINNING AND DELAY-TO-REINFORCEMENT

by

Emily Gregori

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Educational Studies

West Lafayette, Indiana

August 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Mandy Rispoli, Chair

Department of Educational Studies

Dr. Rose Mason

Department of Educational Studies

Dr. Ronald Martella

Department of Educational Studies

Dr. Blake Allan

Department of Educational Studies

Approved by:

Dr. Richard Olenchak

Head of the Graduate Program

To all of the consumers at the Wabash Day Program Center who I have had the privilege of knowing for the last three years.

ACKNOWLEDGMENTS

I would like to acknowledge and thank my dissertation committee members, Drs. Rispoli, Mason, Martella, and Allan for their mentorship and support over the last year. It has been a privilege to work with such an accomplished and talented group of faculty. I am particularly thankful to Dr. Rispoli, who has been an outstanding advisor over the last five years. I would also like to thank my colleagues in the Special Education Program. I would like to offer specific thanks to Catharine Lory and So Yeon Kim, who have not only been outstanding colleagues but incredible friends.

I would also like to acknowledge my community partners at the Wabash Center. I am sincerely grateful for the support of Britteny, Stacia, Katie, and Jamie, who assisted me in conducting all of my research. I am particularly thankful for Summer and Kaila, who graciously allowed me to conduct my research studies in their classroom. It has been a pleasure to work with such a passionate group of people who are committed to improving the lives of individuals with developmental disabilities.

Finally, I would like to acknowledge my family and thank them for their unwavering support over the last five years. I am particularly grateful for my sisters, who have given me unconditional encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF DATA COLLECTION SHEETS

ABSTRACT

Author: Gregori, Emily V. PhD
Institution: Purdue University
Degree Received: August 2019
Title: Maintaining Reductions in Challenging Behavior Following Intervention with Schedule Thinning and Delay-to-Reinforcement
Committee Chair: Mandy Rispoli

Individuals with developmental disabilities often engage in challenging behavior. Reinforcement-based behavioral interventions are the most common treatment for such behaviors. During the initial stages of implementation, reinforcement is delivered at a high intensity, to weaken the relation between challenging behavior and reinforcement. Although this is a critical component in the reduction of challenging behavior, these dense schedules of reinforcement are not feasible in most applied settings. To address these issues, schedule thinning and delay-to-reinforcement are often added post-treatment as a systematic method to reduce the intensity of reinforcement while maintaining low levels of challenging behavior.

Despite professional recommendations to incorporate schedule thinning and delay-to-reinforcement post-treatment, limited research has examined the efficacy of these procedures. Thus, the current dissertation conducted two evaluations of schedule thinning and delay-to-reinforcement following reinforcement-based intervention. Study 1 synthesized the existing literature on the use of schedule thinning following reinforcement-based intervention and evaluated the effects of schedule thinning using a new set of evidence standards. Study 2 evaluated a novel method for preventing the resurgence of challenging behavior following reinforcement-based intervention that involved incorporating activity choice into a delay-to-reinforcement procedure.

Results of study 1 showed that schedule thinning is an effective method for maintaining reductions in challenging behavior following a variety of reinforcement-based interventions. Study

2 demonstrated positive effects for embedding activity choice into delays-to-reinforcement. Across all participants, levels of challenging behavior remained low during the activity choice condition. Taken together, these findings provide additional support for the efficacy of schedule thinning and delay-to-reinforcement. Major findings, limitations, and implications for practice for each study are discussed.

CHAPTER I: GENERAL INTRODUCTION

Challenging behaviors are defined as behaviors that are disruptive to the environment or put the individual or others at risk of injury (Kiernan & Qureshi, 1993; Lundqvist, 2013). Many individuals with developmental disabilities (DD) engage in one or more topographies of challenging behavior. Engagement in challenging behaviors is particularly common among individuals diagnosed with an intellectual disability (ID) and autism spectrum disorder (ASD; Emerson et al., 2001; Lowe et al., 2007). Challenging behaviors emerge in various forms, including self-injury, aggression, property destruction, and stereotypy (Emerson et al., 2001; Holden & Gitlesen, 2006; Lowe et al., 2007). Engagement in challenging behaviors often leads to poor outcomes for individuals diagnosed with DDs including social isolation, poor academic outcomes, and placement in congregate facilities (Emerson et al., 2001; Murphy et al., 2005). Without treatment, challenging behaviors can persist over time, often worsening in adolescence and adulthood (Holden & Gitlesen, 2006).

However, research has shown that treatments based on the principles of operant conditioning can lead to significant reductions in challenging behavior and improvements in quality of life (Gerber et al., 2011). Operant conditioning suggests that an organism's behavior is controlled by its consequences (Cooper, Heron, & Heward, 2007; Skinner, 1938). In other words, behaviors develop and are strengthened by contacting reinforcement following their occurrence. Interventions based on operant conditioning, also known as reinforcement-based interventions, systematically deliver reinforcement to increase prosocial behaviors and/or decrease challenging behaviors (Mayer, Sulzer-Azaroff, & Wallace, 2014).

Decades of research have demonstrated the efficacy of behavioral interventions for the reduction of challenging behavior. Interventions such as differential reinforcement,

noncontingent reinforcement (NCR), and functional communication training (FCT) are among the most common treatments for these behaviors (Khang, Lewin, & Iwata, 2002; Lloyd & Kennedy, 2014; Martinez, Werch, & Conroy, 2016). FCT is the most common reinforcement-based intervention used to treat challenging behavior (Gerow, Davis, Radhakrishnan, Gregori, & Rivera, 2018; Tiger, Hanley, Bruzek, 2008). FCT involves teaching a socially appropriate communicative response to replace challenging behavior (Carr & Durand, 1985; Gerow et al., 2018). The communicative response is functionally equivalent to the challenging behavior, meaning that it produces the same reinforcer as the challenging behavior. The efficacy of FCT has been demonstrated in numerous empirical studies and meets standards to be considered an evidence-based practice (EBP; Wong et al., 2014).

When first implementing reinforcement-based interventions, including FCT, reinforcement is delivered at high intensity to establish the contingency between newly acquired appropriate behaviors and reinforcement, and to weaken the contingency between challenging behavior and reinforcement. However, this level of reinforcement is often impractical for natural change agents (e.g., teachers, parents, direct service personnel) to maintain in typical settings. When natural change agents are unable to maintain high rates of reinforcement, challenging behavior is likely to resurge (Briggs, Fisher, Greer, & Kimball, 2018; LeBlanc, Hagopian, Maglieri, & Poling, 2002).

Resurgence refers to the recurrence of behavior that previously received reinforcement, when an alternative response is placed on extinction or when the rate of reinforcement decreases (Briggs et al., 2018; Fuhrman, Fisher, & Greer, 2018). Periods when reinforcement is unavailable or delayed can resemble extinction-like conditions and lead to the reemergence of challenging behaviors. For example, if a young child requests a cookie and his caregiver is not

able to provide access to that cookie (i.e., the request cannot be reinforced), the child may engage in behaviors such as hitting or screaming that previously were effective in obtaining a cookie.

Resurgence is a critical issue and represents a serious threat to the long-term sustainability of many reinforcement-based interventions, including FCT. Reemergence of challenging behavior is particularly problematic for natural change agents who are unlikely to continue implementing an intervention if challenging behaviors begin to reemerge (LeBlanc et al., 2002). A number of studies have demonstrated resurgence of challenging behavior following initial treatment evaluations of various reinforcement-based procedures, including FCT (Volkert, Lerman, Call, and Trosclair-Lasserre, 2009; Wacker et al., 2011; Wacker et al., 2013).

For example, Volkert et al. (2009), evaluated if resurgence would occur following the implementation of FCT among five children diagnosed with DDs. Results demonstrated that periods of extinction, or lean schedules of reinforcement, resulted in resurgence of challenging behavior in 80% of participants, with some participants engaging in higher levels of challenging behavior than observed during baseline. The results of Volkert et al. and others, illustrate many of the issues associated with resurgence, and indicate the importance of identifying methods to mitigate the effects of resurgence.

Two common approaches exist for decreasing the intensity of reinforcement and preventing resurgence of challenging behavior following intervention. The first approach, schedule thinning, involves a systematic decrease in the rate of reinforcement until a final schedule is met that resembles what is appropriate in a natural setting (Davis et al., 2018; Hagopian, Boelter, & Jarmolowicz, 2011; LeBlanc, Hagopian, Maglierei, & Poling, 2002). Schedule thinning consists of a variety of procedures that have been shown to prevent the resurgence of challenging behavior in a number of studies (Austin & Tiger, 2015;

Ghaemmaghami, Hanley, & Jessel, 2016; Hagopian, Kuhn, Long, & Rush, 2005; Hanley, Iwata, & Thompson, 2001).

For example, Fuhrman et al. (2018) evaluated the effects of multiple schedules on the resurgence of challenging behavior following FCT with two children diagnosed with multiple DDs. The initial FCT evaluation resulted in low levels of challenging behavior for both participants. During schedule thinning, two schedules of reinforcement were in effect and each was signaled with a specific colored index card. The first schedule was signaled with a green card and represented the availability of reinforcement. The second schedule was signaled with a red card and represented that reinforcement was unavailable (i.e., extinction). Periods of non-reinforcement (schedule two) were gradually increased across sessions. Following schedule thinning, each participant was exposed to an extinction challenge to test for resurgence. During the extinction challenge neither appropriate communication nor challenging behavior produced reinforcement. For both participants, schedule thinning using multiple schedules, reduced resurgence of challenging behavior.

The second approach, delay-to-reinforcement, involves maintaining a dense schedule of reinforcement, but inserting a delay between the emission of an alternative response and the delivery of reinforcement (Hagopian et al., 2011). Delay-to-reinforcement is an ideal procedure to use because it reflects what happens in most applied settings, and teaches individuals to tolerate periods of non-reinforcement. In a recent investigation, Ghaemmaghami et al. (2016) implemented a delay-to-reinforcement with four individuals diagnosed with DDs. During the delay, the participants were prompted to engage with a variety of leisure activities or work tasks. Results indicated that gradually increasing the delay-to-reinforcement plus the use of alternative activities prevented the resurgence of challenging behavior.

The findings of the studies described above suggest that schedule thinning and delay-to-reinforcement may be necessary for maintaining reductions in challenging behavior following many reinforcement-based interventions, including FCT. However, compared to the abundance of literature on reinforcement-based interventions, significantly less research exists on the use of schedule thinning and delay-to-reinforcement. Given the importance of reducing reinforcement intensity, more research in this area is warranted. Thus, two studies were conducted to address gaps in the current literature base on schedule thinning and delay-to-reinforcement.

The first study synthesized the extant literature on schedule thinning following reinforcement-based interventions. The review included a summary and narrative description of commonly utilized schedule thinning procedures, as well as an analysis of the efficacy of schedule thinning using a novel set of evidence standards. Specifically, the following three research questions were addressed:

a) What are the common schedule thinning procedures used following reinforcement-based interventions?
b) What are the characteristics of schedule thinning procedures used following reinforcement-based interventions?
c) What is the effect of schedule thinning on challenging behavior following the implementation of reinforcement-based interventions?

The second study used an experimental single-case design to evaluate the effects of delay-to-reinforcement on challenging behavior following FCT. Three schedule thinning conditions were evaluated using a multielement design (Kennedy, 2005). During the first condition, participants were given access to researcher selected stimuli during delays-to-reinforcement. In the second condition, participants were given a choice of alternative stimuli to use during the delay. In the

last condition, participant's were given no alternate stimuli to interact with during the delay. Levels of challenging behavior, appropriate communication, and item engagement were measured in each condition to evaluate the unique effects of each condition on challenging behavior. The following two research questions were addressed:

a) What is the effect of providing alternative activities on challenging behavior during delays-to-reinforcement following functional communication training?
b) How does the provision of choice of alternative activities impact challenging behavior during delays-to-reinforcement?

CHAPTER II: SCHEDULE THINNING FOLLOWING REINFORCEMENT-BASED INTERVENTIONS: A SYSTEMATIC REVIEW

Introduction

Challenging Behavior and Disability

Developmental disabilities encompass several conditions that impact a variety of domains including physical, behavioral, and cognitive (National Centers for Disease Control and Prevention, 2010). DDs manifest during the developmental period (i.e., ages 0-18) and impact an individual's daily functioning (CDC, 2010). The most common DDs affecting individuals in the United States are ID, ASD, attention deficit hyperactivity disorder (ADHD), Down syndrome, fetal alcohol syndrome, and Fragile X syndrome (CDC, 2010). Although not a core feature of many DDs, many individuals with IDD, particularly ID and ASD, engage in challenging behavior (Bowring, Totskika, Hastings, Toogood, Griffith, 2017; Emerson et al., 2001; Holden & Gitlesen, 2006; Lowe et al., 2007; Totsika, Toogood, Hastings, & Lewis, 2008).

Challenging behaviors are defined as behaviors that are disruptive to the environment or put the individual or others at risk of injury (Kiernan & Qureshi, 1993; Lundqvist, 2013). Challenging behaviors appear in many forms such as self-injurious behavior (SIB), aggression, property destruction, and stereotypy (Bowring, Totsika, Hastings, Toogood, & Griffith, 2017; Emerson et al., 2001; Lowe et al., 2007; Lundqvist, 2013). Engagement in these behaviors can have detrimental consequences that severely and negatively impact an individual's life (Llyod & Kennedy, 2014). Engagement in challenging behaviors often leads to poor academic outcomes, social isolation, inability to obtain employment, and placement in restrictive facilities (Emerson et al., 2001; Murphy et al., 2005). Additionally, individuals who engage in severe topographies

of challenging behavior are often exposed to intrusive treatments such as restraint and psychotropic drugs (Emerson et al., 2001).

Prevalence estimates suggest that 10-20% of individuals with DDs engage in one or more topographies of challenging behavior (Bowring et al., 2017; Emerson et al., 2001; Lowe et al., 2007; Lundqvist, 2013). Individuals diagnosed with an ID, ASD, or multiple disabilities (e.g., ID plus a visual impairment) are more likely to engage in challenging behavior than individuals diagnosed with other DDs (Emerson et al., 2001; Lowe et al., 2007). Several other correlates of challenging behavior have been identified and include poor receptive and expressive communication, urinary incontinence, social deficits, and an overall lower skill level (Bowring et al., 2017; Jones et al., 2008; Lundqvist, 2013). These correlates suggest that individuals with the most significant needs are at an increased risk of engaging in challenging behavior, and are more likely to engage in high-risk challenging behaviors such as aggression and SIB (Kahng et al., 2002; Poppes, van der Putten, & Vlaskamp, 2014).

Individuals who engage in challenging behavior as young children are likely to experience increases in these behaviors as they age (Davies & Oliver, 2013). Davies & Oliver found that many high-risk topographies of challenging behavior, including aggression and SIB, increase as individual's progress from childhood to young adulthood. Murphy et al. (2005) also found that challenging behaviors can persist over time in a longitudinal total population study that included 166 children with DDs. The authors found that children who displayed the most significant levels of challenging behavior at time 1, also engaged in the highest levels of challenging behavior 12 years later, at time 2. The chronicity, severity, and consequences associated challenging behaviors have resulted in an abundance of research examining the causes of challenging behavior, and effective treatments based on those causes.

Etiology of Challenging Behavior

The development of challenging behavior among individuals with DDs has been investigated by medical and psychological researchers for decades. Although some researchers have indicated a biological origin of challenging behavior, early research in the area of applied behavior analysis has suggested an operant function (Carr, 1977; Iwata, Dorsey, Slifer, Buaman, & Richman, 1982; 1994). According to the operant perspective, all human behavior, including challenging behaviors, develops and are maintained as a result of exposure to specific environmental variables (Cooper et al., 2007; Skinner, 1938). Operant behavior (i.e., behavior controlled by its consequences; Cooper et al., 2007) develops as a result of interactions between the target behavior and the environmental events that follow that behavior (i.e., consequences; Cooper et al., 2007; Michael, 2000).

In 1994, Iwata et al. demonstrated the operant nature of SIB in a series of experiments conducted with individuals diagnosed with ID. Each participant was exposed to four experimental conditions designed to test if SIB was sensitive specific environmental contingencies. Two of the conditions were designed to test if SIB was sensitive to social positive and negative reinforcement contingencies. During each condition, relevant antecedents and consequences were arranged and presented in a mass trial format. Engagement in SIB during the social positive and negative reinforcement conditions resulted in access to the putative reinforcer. For example, during the academic demand condition, the experimenter presented a series of nonpreferred tasks (e.g., sorting) and prompted the participant to complete the demands using a series of least-to-most prompts. If at any point during the work session the participant engaged in SIB, the experimenter removed the demand and turned away from the participant for 30 s.

The third condition tested whether SIB persisted in the absence of socially mediated contingencies. During this condition, the participant was left in an empty therapy room without preferred items, academic demands, or attention and engagement in SIB produced no programmed consequences. The final condition, unstructured play, was designed as a control. During the play condition the participant had access to preferred items, moderate levels of attention, and no work demands were made. Similar to the alone condition, all instances of SIB were ignored.

Results of the experiments, known as a functional analysis (FA), demonstrated a functional relation between SIB and distinct environmental contingencies for six of the nine participants. The findings of this study provided empirical evidence of the operant nature of SIB. Since the development of the FA researchers and clinicians have been able to treat challenging behavior by identifying and manipulating the contingencies maintaining those behaviors. Currently, interventions based on the results of FA, or function-based interventions, are some of the most common treatments for challenging behavior among individuals with IDD (Wong et al., 2014).

Reinforcement-Based Interventions

When using reinforcement-based interventions, clinicians systematically deliver reinforcement on predetermined schedules (Mayer et al., 2014). The specific reinforcers delivered within a reinforcement-based intervention differ based on the function of the individuals challenging behavior. Reinforcers can be categorized as positive or negative. Positive reinforcement is the application of a stimulus contingent on the emission of a specific behavior while negative reinforcement refers to the removal of an aversive stimulus as the consequence of a given response (Cooper et al., 2007; Mayer et al., 2014).

Reinforcement-based interventions can be defined as response-based or time-based. Within response-based procedures, reinforcement is delivered contingent on the emission of a socially appropriate behavior that often matches the function of the challenging behavior. Time-based procedures differ in that reinforcement is delivered contingent on the passage of a specific period of time in which the target challenging behavior may or may not have occurred. Many reinforcement-based interventions are combined with an extinction component (Cooper et al., 2007; MacNaul & Neely, 2018). When using an extinction procedure, challenging behavior no longer results in the delivery of reinforcement (Cooper et al., 2007). Several reviews and meta-analyses have identified common reinforcement-based interventions for the treatment of challenging behavior. These include differential and noncontingent reinforcement (Gregori, Rispoli, Gerow, & Lory, 2018; Kahng et al., 2002; Lloyd & Kennedy, 2014; Martinez et al., 2016).

Response-Based Reinforcement Procedures

Differential reinforcement of alternative/incompatible behavior.

Differential reinforcement of alternative behavior (DRA) is a procedure used to decrease challenging behavior by teaching an appropriate response that serves as an alternative to the behavior of concern (MacNaul & Neely, 2018). When implementing DRA, reinforcement is provided contingent on the emission of the alternative behavior and is often withheld following the occurrence of challenging behavior (MacNaul & Neely, 2018). For example, teachers often teach students to raise their hands instead of calling out answers during a lesson. In most cases, the teacher only responds to students whose hands are raised. Differential reinforcement of incompatible behavior (DRI) is a variation of DRA. DRI differs only slightly from DRA in that the alternative behavior cannot occur concurrently with the challenging behavior. For example,

teachers often train children who run through the hallway to walk in instead of run. This constitutes a DRI procedure because the children cannot walk and run at the same time.

DRA and DRI are often preceded by a pretreatment functional behavior assessment (FBA) to determine the variables maintaining the challenging behavior. The results of the FBA are then used to inform selection of the alternative response. Alternative responses are often selected based on one or more of the following criteria: the behavior (a) is functionally equivalent to the challenging behavior, (b) requires less response effort to emit than the challenging behavior, (c) is socially appropriate, and (d) is in the individuals current repertoire (Cooper et al., 2007). Reinforcement for the alternative behavior is typically provided on a fixed-ratio 1 (FR-1) schedule of reinforcement when DRA or DRI are first implemented. This means that each emission of the alternative behavior produces reinforcement. This schedule is used to establish the new behavior in the individual's repertoire.

In 2018, Slocum, Mehrkam, Peters, and Vollmer used DRA to treat automatically maintained pica for a 13-year-old girl diagnosed with ASD using a withdrawal design. During the baseline sessions, all attempts to ingest sand and dirt were blocked but resulted in no other consequences. During the intervention phase, the participant was prompted to discard nonedible items in a wastebasket. Discarding of items resulted in brief praise and a preferred edible item. Prior to the implementation of DRA, attempts to ingest sand and dirt were moderately high (M= 2.6 per session). Additionally, the participant made no attempts to discard nonedible items. Immediately following the implementation of DRA, pica attempts reduced (M= 0.44 per session) and instances of discarding increased (M= 4.06 per session).

To date, DRA and DRI have been used to treat challenging behavior among children and adults with DDs including ID, ASD, ADHD, and emotional behavioral disorders (Gregori et al.,

2018; Ivy, Meindl, Overley, & Robson, 2017; MacNaul & Neely, 2018; Matson & Boisjoli, 2009; Soares, Harrison, Vannest, & McClelland, 2016). Several systematic reviews have been conducted to determine the state and quality of the extant research on DRA (including DRI). Petscher, Rey, and Bailey (2009) reviewed 116 empirical studies on DRA against The Division 12 Task Force criteria for empirically supported treatments. Results of the review indicated that DRA was a well-established practice for the treatment of destructive behavior and food refusal for children and adults with DDs.

Functional communication training.

FCT is a variation of DRA and is one of the most commonly implemented interventions for the treatment of challenging behavior (Tiger et al., 2008). FCT differs from a traditional DRA procedure in two distinct ways. First, a pretreatment FBA is conducted to identify the variables reinforcing the challenging behavior. Second, the alternative response is a socially appropriate communicative response that produces the same reinforcer as the challenging behavior (Carr & Durand, 1985). FCT is typically combined with an extinction procedure, which has been shown to be a critical component of the procedure (Hagopian, Fisher, Sullivan, Acquisto, & LeBlanc, 1998; Shukla & Albin, 1996). In the early stages of FCT, reinforcement is typically provided immediately following each occurrence of the alternative communicative response. This continuous schedule of reinforcement establishes the contingency between the appropriate communicative response and the delivery of reinforcement (Cooper et al., 2007).

In a recent evaluation, Torres-Viso, Stohmeier, and Zarcone (2018) evaluated the effects of FCT plus extinction on the challenging behavior and appropriate communication of a 12-year-old female diagnosed with multiple DDs. A FA determined that challenging behavior was maintained by adult compliance with mands for rearrangement of materials and body positions.

For example, the participant would often aggress toward her father if he uncrossed his legs while sitting. During baseline, challenging behavior resulted in adult compliance with mands, and levels of challenging behavior were high. During intervention, compliance was reinforced contingent on emission of one of three appropriate mands (e.g., “Move your leg please”), while challenging behavior produced no programmed consequences. FCT was shown to be effective, resulting in near zero levels of challenging behavior and increases in appropriate communication.

FCT is one of the most effective interventions for the treatment of challenging behavior (Gerow et al., 2018; Tiger et al., 2008; Wong et al., 2014). FCT has been evaluated in a number of reviews and has been shown to be an effective treatment for challenging behavior among children and adolescents with ASD (Heath et al., 2015; Wong et al., 2014). In the most recent review of FCT, Gerow et al. (2018) reviewed the quality and evidence of the extant literature on FCT across disability categories using the WWC criteria for EBP determination. The authors reviewed the quality and strength of the evidence of 493 experiments across 215 studies. Findings indicated that FCT resulted in reductions in challenging behavior for 136 participants. The authors concluded that FCT meets standards to be considered an EBP for individuals diagnosed with ASD, ID, other health impairments, and multiple disabilities.

Time-Based Reinforcement Procedures

Differential reinforcement of other behavior.

Differential reinforcement of other behavior (DRO) is a procedure in which reinforcement is provided contingent on the emission of any behavior other than the challenging behavior during a predetermined period of time (i.e., the DRO interval; Cooper et al., 2007; Reynolds, 1961; Wong et al., 2014). Unlike other differential reinforcement procedures, DRO does not reinforce a socially acceptable alternative behavior. Instead, any behavior other than the

target behavior is reinforced. DRO intervals are initially very brief (e.g., 5 s seconds). The interval is then gradually increased on a fixed or variable schedule contingent on stable decreases in the target behavior (Cooper et al., 2007). To implement DRO, an interventionist sets a timer for a preselected period of time (e.g., 5s). If during the 5 s interval the individual engages in any behavior other than the target behavior, the interventionist provides access to a functional or nonfunctional reinforcer. Like other differential reinforcement procedures, the target challenging behavior is often placed on extinction (Mazaleski, Iwata, Vollmer, Zarcone, & Smith, 1993). Additionally, if during the interval the individual engages in the target behavior, the interventionist may reset the interval. When using resetting DRO, the interventionist restarts the interval immediately after the individual engages in challenging behavior. The individual must then refrain from engaging in the target behavior for another interval before reinforcement is available (Gehram, Wilder, Forton, & Albert, 2017).

Tiger, Fisher, and Bouxsein (2009) implemented DRO to treat a 19-year-old man's automatically maintained SIB (i.e., skin picking). At the time of the study, self-injury was observed at high levels leaving the man with permanent scars and wounds. A 5-min DRO schedule was implemented in which the man was given tickets for each 5 min interval in which he did not self-injure. DRO resulted in immediate reductions in self-injury, and the DRO interval was extended to 15 min. Additionally, the participant was taught to self-monitor his skin picking behavior in several novel settings.

DRO has been evaluated in a number of experimental studies and has been shown to be effective in treating numerous topographies of challenging behaviors including SIB, stereotypy, and aggression (Figueroa, Thyer, & Thyer, 1992; Lustig et al., 2014; Matson, Dixon, & Matson, 2005; Tiger et al., 2009; Wong et al., 2014). Weston, Hodges, and Davis (2018) summarized the

available literature on DRO to treat challenging behavior for children and adults with ASD. A single-case effect size index, the percent of nonoverlapping data (PND; Scruggs, Mastropierei, & Casto, 1987), was calculated for studies that met inclusion for the review. Of the 17 studies included, nearly half yielded high effect sizes (PND of .90 or higher). The omnibus effect size across all included studies was .78, which according to the authors, suggests that DRO is an effective treatment for challenging behavior among children and adults with ASD (Rakap, 2015).

Noncontingent reinforcement.

NCR is a procedure in which reinforcement is delivered after a predetermined period of time, independent of the target challenging behavior (Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993). Reinforcement is given on either a fixed or variable-time schedule. The initial schedule of reinforcement is initially very dense so that the delivery of reinforcement precedes the occurrence of challenging behavior. The dense schedule of reinforcement creates an enriched environment that functions as an abolishing operation and reduces an individual's motivation to engage in challenging behavior (Cooper et al., 2007; Richman, Barnard-Brak, Grubb, Bosch, & Abby, 2015).

Rispoli, Brodhead, Wolfe, and Gregori (2018) evaluated the effects of noncontingent attention (NCA) on vocal scripting for three boys diagnosed with ASD. Results of a pretreatment trial-based functional analysis (TBFA) indicated that scripting was maintained by social positive reinforcement in the form of adult attention. Prior to NCA, moderate to high levels of vocal scripting were observed across all three boys. For all three boys, the introduction of NCA was associated with immediate reductions in vocal scripting.

NCR is a common treatment for challenging behavior for both children and adults with various DDs (DiGennaro Reed, Hirst, & Hyman, 2012; Gregori et al., 2018; Llyod & Kennedy,

2014; Rapp & Vollmer, 2005), and has been shown to be effective in the treatment of various challenging behaviors (Carr et al., 2000; Carr, Severtson, & Lepper, 2009; Richman et al., 2015; Tucker, Sigafoos, & Bushell, 1998). Richman et al. conducted a meta-analysis on the use of NCR to treat challenging behavior among individuals with DDs. Hierarchical linear modeling was used to evaluate 55 studies (including 91 participants) using single-case experimental designs. Results of the analyses produced a strong effect size for NCR (d=-1.58), suggesting that NCR is effective in treating problem behavior among individuals with various DDs.

Sustainability of Reinforcement-Based Interventions

The initial implementation of reinforcement-based interventions requires a dense schedule of reinforcement. The initial dense schedule is critical because it establishes the contingency between an alternative response and reinforcement or weakens the contingency between the challenging behavior and reinforcement. However, these dense schedules are impractical outside of therapeutic settings and can lead to problems for both the individual engaging in challenging behavior and their caretakers (Call et al., 2018; LeBlanc et al., 2002; Lundqvist, 2013).

For example, dense schedules of reinforcement do not reflect the contingencies in place in natural environments such as schools and community employment settings, where individuals with DDs are expected to participate. Dense schedules of reinforcement in these settings may interfere with the individual's skill to perform as expected. To illustrate, if a young child with autism is given a break from instructional demands every time he or she asks, this will result in a significant loss of instructional time and likely poor academic performance. For older individuals with DDs, taking frequent breaks in an employment setting may lead to a decline in performance and result in termination.

Natural change agents are also typically not prepared to deliver reinforcement at such high rates. Teachers, parents, and direct service personnel have numerous responsibilities that make it impractical, if not impossible, to deliver frequent and immediate reinforcement. When caregivers are asked to implement reinforcement at these rates, errors in the fidelity of intervention implementation often occur (LeBlanc et al., 2002). One such error is a decrease in the frequency of reinforcement, or an error of omission (LeBlanc et al., 2002; St. Peter Pipkin, Vollmer, & Sloman, 2010). Rapid or unplanned changes in the schedule of reinforcement often lead to resurgence of challenging behavior (Cooper et al., 2007; LeBlanc et al., 2002; Mayer et al., 2014).

Resurgence is the recovery of behaviors that were previously reinforced when an alternative behavior is no longer receiving reinforcement (Briggs et al., 2018; Mayer et al., 2014). When caregivers are unable to deliver reinforcement at the rate defined by the interventionist, the individual with DD is likely to engage in behaviors that previously produced reinforcement. For example, if a teacher is unable to call on a child who is raising his hand, the child may return to screaming or hitting his desk to obtain his teachers' attention. Resurgence has been observed in both response and time-based reinforcement interventions and often leads to the discontinuation of the intervention (Briggs et al., 2018; LeBlanc et al., 2002; Saini, Fisher, & Pisman, 2017). To address the issues associated with reinforcement density within reinforcement-based interventions, researchers have developed methods for decreasing the intensity of reinforcement during reinforcement-based interventions.

Schedule Thinning and Delay-to-Reinforcement

Several approaches have been developed to decrease the intensity of reinforcement-based interventions. Two of the most common methods include schedule thinning and delay-to-

reinforcement. Schedule thinning is a process of decreasing the frequency or density of reinforcement following reinforcement-based interventions (Davis et al., 2018; Hagopian et al., 2011; LeBlanc et al., 2002). Delay-to-reinforcement involves keeping the initial schedule of reinforcement intact but inserting a delay between the emission of an alternative response and the delivery of reinforcement, or increasing the amount of time the individual is expected to wait before they can request access to a functional reinforcer (Hagopian et al., 2011; LeBlanc et al., 2002). Although conceptually different, in the applied literature, delay-to-reinforcement is typically described as an approach to schedule thinning (Hagopian et al., 2011). Thus for this review, we include it under the list of schedule thinning procedures described below. Several procedures have been developed to facilitate schedule thinning following the initial implementation of reinforcement-based interventions. Among the most common procedures are: (a) delay-to-reinforcement (a) demand fading, (b) increasing intervals, (c) multiple schedules, and (d) response restriction (Hagopian et al., 2011).

Delay-to-reinforcement.

Delay-to-reinforcement (also known as delay fading or delay schedules) involves inserting a delay between the emission of the alternative behavior and the delivery of reinforcement, or increasing the amount of time an individual has to wait before requesting access to a functional reinforcer (Hagopian et al., 2011; LeBlanc et al., 2002). To illustrate, a parent may tell his or her child to wait immediately after the child asks for a cookie. The parent would then allow the child to have a cookie a few seconds later (Fisher, Greer, Fuhrman, & Querim, 2015; Hagopian et al., 2011). The delay is initially very brief (e.g., 10 s) and gradually increases following steady reductions in challenging behavior across multiple sessions. Delay cues, or signaled delays, are often used during delays-to-reinforcement. Delay cues are used to

signal the contingencies in place during the delay (e.g., “You have to wait 10 s before you can ask for tablet again;” Reichle, Johnson, Monn, & Harris, 2010). Signaled delay cues can be verbal, visual, explicit or general (Hong et al., 2014; Reichle et al., 2010). Research has shown that signaled delays can be used to facilitate greater tolerance for delays-to-reinforcement (Reichle et al., 2010; Hong et al., 2014). Fisher et al. (2000) implemented a delay-to-reinforcement following a FCT evaluation for a man who exhibited destructive behavior. During the delay-to-reinforcement, the man was told to wait for increasing periods of time for access to his preferred tangible item. During the delay, the man was given alternative work tasks to complete. The delay-to-reinforcement procedure was effective in facilitating schedule thinning and resulted in low levels of challenging behavior until the terminal delay was reached.

Demand fading.

Demand fading is a commonly used schedule thinning procedure for escape maintained challenging behavior (Hagopian et al., 2011; Lalli, Casey, & Kates, 1995). Demand fading involves systematically increasing the number of task demands an individual must complete before he or she is given access to reinforcement (Hagopian et al., 2011). Knox, Rue, Wildenger, Lamb, & Luiselli (2012) used differential reinforcement, prompting, and demand fading to treat food selectivity for a girl diagnosed with ASD. During the intervention, the girl was given three foods that previously evoked food refusal. Contingent on consumption of each food, the girl was given a sticker that could be exchanged for a preferred item. The amount of food presented during each session increased by 20% until 100% of the meal was consumed. The girl was still consuming 100% of her meal seven months following the termination of the intervention.

Increasing intervals.

Increasing intervals refer to the gradual increase in the duration of the DRO or NCR interval. For example, Bergstrom, Tarbox, & Gutshall (2011) used increasing intervals to increase the amount of time a young boy went without injuring his dog. During the initial DRO training, the boy received access to a preferred tangible contingent on the absence of aggression towards the dog (i.e., touching the dog's rear) for 10 s. Following an initial reduction in challenging behavior, the interval increased progressively until a terminal interval of 10 min was reached. Levels of aggression towards the dog remained near zero throughout the fading procedure.

Multiple schedules.

When implementing a multiple schedule, two schedules of reinforcement are presented independently in an alternating fashion (Cooper et al., 2007). One schedule indicates the availability of reinforcement, while the other schedule indicates that reinforcement is unavailable (i.e., extinction). A specific discriminative stimulus (e.g., different colored cards) is correlated with each schedule and is present when each schedule is in effect (Fuhrman et al., 2018). Multiple schedules are commonly used to facilitate schedule thinning following the implementation of response-based reinforcement procedures such as FCT.

Following the successful implementation of FCT, Fisher et al. (2015) used multiple schedules to facilitate schedule thinning for three boys with DDs who engaged in tangibly maintained challenging behavior. During the multiple schedule sessions, two 60 s schedules (60/60 schedule) were alternated in a quasi-random fashion. During one schedule, access to preferred tangible items was given continuously (i.e., reinforcement available), while access was denied during the other schedule (i.e., extinction). During the reinforcement available sessions,

the therapist wore a yellow wristband and described the reinforcement contingency in place. For all three children, low levels of challenging behavior maintained with the use of the 60/60 schedule. The schedule was subsequently thinned using a 60/300 schedule in which periods of extinction increased to 300 s. All other procedures were the same as the 60/60 schedule sessions. Like the 60/60 schedule, the 60/300 schedule was associated with low levels of challenging behavior relative to baseline.

Response restriction.

Response restriction, also known as RR FCT, involves removing access to an individual's communication materials (e.g., communication cards, speech-generating device) for gradually increasing periods of time (Fisher et al., 2014; Roane, Fisher, Sgro, Falcomata, & Pabico, 2004). Initially, access is restricted only briefly (i.e., a few seconds). Periods of restriction are then increased following stable decreases in challenging behavior across multiple sessions (Fisher et al., 2014; Roane et al., 2004). Fyffe, Kahng, Fittro, and Russell (2004) used RR FCT following an initial FCT evaluation for a boy with multiple disabilities who displayed inappropriate sexual behavior. During the initial FCT evaluation, the boy had continuous access to his communication card. When RR FCT was implemented, periods in which the card was unavailable gradually increased from 5 s to the terminal schedule of 5 min. Throughout the schedule thinning evaluation, instances of challenging behavior remained low.

Gaps in the Research, Study Rationale, and Research Questions

Schedule thinning is an important component in the long-term sustainability of reinforcement-based interventions. Numerous studies have shown that schedule thinning can lead to decreases in reinforcement frequency while maintaining low levels of challenging

behavior (Fisher et al., 2000; Hagopian et al., 2011; Richman et al., 2015). Despite the importance of schedule thinning, few reviews have synthesized the existing literature on these procedures. Additionally, the reviews that have been conducted on schedule thinning have been intervention-specific and have not used objective measures to classify the evidence of schedule thinning.

For example, Hagopian et al. (2011) synthesized the literature on schedule thinning procedures following FCT. The purpose of their review was to describe the common schedule thinning procedures used following FCT (i.e., delay-to-reinforcement, demand fading, multiple schedules, and response restriction) and to summarize studies that utilized these procedures. The authors found that schedule thinning was effective in maintaining low levels of challenging behavior across schedules with or without the use of additional treatment components. However, the effects of schedule thinning were classified based on visual analysis and author report. No other metrics were used to objectively evaluate the effects of the schedule thinning procedures.

In another investigation, Richman et al. (2015) conducted a quantitative synthesis of NCR for the treatment of challenging behavior. Overall, NCR resulted in a strong effect size and suggested that NCR is an effective treatment for challenging behavior. Additional analyses were conducted to determine whether schedule thinning impacted the effectiveness of NCR. Results showed that schedule thinning resulted in a minor decrease in effect size, but the results were not statistically significant. The authors concluded that decreases in effect size were not unexpected, given that minor increases in challenging behavior during schedule thinning are common.

To date, only one review has addressed schedule thinning for all reinforcement-based interventions. LeBlanc et al. (2002) summarized schedule thinning procedures following response and time-based reinforcement interventions. The authors provided a summary of

common schedule thinning procedures and guidelines for each schedule. Although LeBlanc et al. provided the most comprehensive summary of schedule thinning, the review was broad and not intended to synthesize all of the extant literature on schedule thinning. In fact, no review has systematically evaluated the existing literature on schedule thinning following reinforcement-based interventions. Given that schedule thinning is considered "best practice" (LeBlanc et al., 2002), there is a need for a comprehensive review that objectively evaluates the efficacy of this practice. Therefore, the purposes of the current review were to (a) summarize the extant literature on schedule thinning following reinforcement-based interventions, (b) describe the characteristics of commonly used schedule thinning procedures, and (c) determine the effects of schedule thinning using a new set of evidence standards. The following research questions were addressed:

a) What are the common schedule thinning procedures used following reinforcement-based interventions?
b) What are the characteristics of schedule thinning procedures used following reinforcement-based interventions?
c) What is the effect of schedule thinning on challenging behavior following the implementation of reinforcement-based interventions?

Method

Search Procedures

Article identification.

A systematic search was developed in partnership with a reference librarian specializing in systematic reviews. The entire search consisted of four phases including: (a) database search,

(b) inclusion review, (c) ancestral search, and (d) hand search. The first author conducted all four phases of the search. Three databases including PscyhINFO, ERIC, and Education Source, were searched via the Ebsco interface, using 28 search terms across two categories (see Table A1 for a full list of search terms). The first category consisted of intervention terms (e.g., differential reinforcement, noncontingent reinforcement, functional communication training, and their variations) and the second category included terms related to schedule thinning (e.g., schedule thinning, delay-to-reinforcement, signaled delay, etc.). Terms in each category were combined using the term "OR." Each category was searched independently and then combined with the term "AND." All articles were uploaded to an electronic database manager and duplicates were removed. The remaining articles were evaluated against nine inclusion criteria.

Inclusion evaluation.

To be included in this review each study had to: (a) implement FCT, NCR, or any variation of differential reinforcement, (b) implement schedule thinning following the initial treatment evaluation, (c) include challenging behavior as a dependent variable, (d) display initial treatment data on a line graph, (e) display schedule thinning data on a line graph (f) include human participants, (g) include at least one participant diagnosed with a DD, (h) be published in a peer-reviewed journal, and (i) be published in English. The raters responsible for reviewing the included articles spoke English as their primary language, and were not able to review articles published in other languages. Thus, studies published in languages other than English were excluded. Each article was given a rating of include or exclude based on the criteria. To be given a rating of include, the study had to meet all nine criteria. If a study failed to meet one or more of the criteria it was given a rating of exclude.

Ancestral and hand searches.

An ancestral search was conducted to obtain additional articles not identified via the database search. Articles from the reference lists of included studies were uploaded to an electronic database manager and evaluated against the inclusion criteria. A hand search of the *Journal of Applied Behavior Analysis, Research in Developmental Disabilities, Behavior Modification, Behavioral Interventions, Journal of Behavioral Education,* and *Education and Training in Autism and Developmental Disabilities* was conducted to identify supplemental articles. Finally, a hand search of three recent systematic reviews on reinforcement-based interventions and schedule thinning were reviewed for additional articles.

Quality Review

Quality standards and coding procedures.

The methodological quality of each of the included studies was evaluated using the What Works Clearinghouse (WWC) Basic Design Standards. For the purposes of this review, only the initial treatment evaluation was considered in the design evaluation. Each article was evaluated using the scoring protocol described by the WWC and adapted by Maggin, Briesch, and Chafouleas (2013) and Hong et al. (2016). The Basic Design Standards include six criteria that are scored as 2 (meets design standards), 1 (meets design standards with reservations), or 0 (does not meet design standards). Each study was given an overall score based on the results of the Design Standards evaluation. Design Standard 1 requires that the researcher systematically manipulate the independent variable. Design Standard 2 consists of three criteria. Design Standard 2A, evaluates whether inter-assessor (IAA) data were collected. Design Standard 2A is rated as meets (IAA collected) or does not meet standards (IAA not collected). To meet Design

Standard 2B IAA must be collected on a minimum of 20% of data points per study phase (Hong et al., 2016). To meet Design Standard 2B with reservations, the study must collect IAA on a minimum of 20% of data points across study phases (e.g., baseline and intervention combined; Hong et al., 2016). A study was rated as not meeting Design Standard 2B if IAA was collected on less than 20% of data points across study phases. Design Standard 2C indicates that IAA scores should be above minimal thresholds (i.e., ≥80% or 0.6 Kappa). For Design Standard 2C studies are given a rating of meets or does not meet standard. To meet Design Standard 3, a study must make at least three attempts to demonstrate a functional relation at three different points in time. If a study failed to make three attempts, it was given a rating of does not meet standards. To meet Design Standard 4, a study must include a minimum of five data points in each study phase, including baseline. Studies that included three to four data points per phase were rated as meets standards with reservations. If any phase (baseline or intervention) included less than three data points, it was rated as not meeting standards. Each study was given an overall score based on the results of the Design Standards evaluation. If a study met all six Design Standards it was given a score of 2 (meets design standards). If any of the standards was given a rating of meets standards with reservations, and did not receive any ratings of does not meet standards, the study was given a score of 1 (meets standards with reservations). If any of the standards were not met, the study was given a score of 0 (does not meet standards). See Table A2 for a description of the WWC Design Standards.

Evidence evaluation.

Studies that met the Basic Design Standards with or without reservations were evaluated against the WWC Evidence Standards. The purpose of the evidence evaluation is to use visual analysis techniques to determine if a functional relation exists between the independent and

dependent variables. The Evidence Standards include 22 criteria that are used to evaluate individual experiments based on six features including (a) level, (b) trend, (c) variability, (d) immediacy of effect, (e) overlap, and (f) consistency of data patterns in similar phases (See Table A3 for a list of all Evidence Standards). Level, trend, and variability are examined within each study phase, while immediacy of effect, overlap, and consistency of data patterns are evaluated across adjacent phases (i.e., baseline and intervention). Following the application of each criterion, experiments are given an overall score. According to the Evidence Standards, overall effect is based on three criteria. Criteria 1 is the number of data points per phase. Experiments that included five or more data points within each study phase (including baseline) are rated as having a strong effect (2). Experiments that include 3-4 data points per phase are rated as having evidence of moderate effectiveness (1), and experiments that include less than three data points per phase are rated as having no evidence of effectiveness (0). Criteria 2 describes the number of demonstrations of treatment effect between baseline and intervention. Experiments that provide a minimum of three demonstrations are rated as having a strong effect. Experiments that present fewer than three demonstrations are rated as having no evidence of effect. Criteria 3 examines the ratio of effects to non-effects. An experiment is rated as having strong evidence, if no instances of non-effect are observed. If the ratio of effects to non-effects is less than or equal to 3:1, the experiment demonstrates evidence of moderate effectiveness. If the ratio of effects to non-effects exceeds 3:1, the experiment is deemed as having no evidence of effectiveness. The overall effect of the experiment is determined by evaluating the ratings of the previous three criteria. If all three criteria receive a score of 2, the experiment is rated as strong effect. If any of the three criteria receives a score of 1 and no scores of 0, it is given a rating of moderate effect. If one or more of the three criteria is given a score of 0 the experiment is rated as no effect. For this

study, only experiments found to have evidence of a strong or moderate functional relation were included in the descriptive and schedule thinning evaluations.

Schedule Thinning Evidence Evaluation

Evidence standards based on the format developed by the WWC were developed to evaluate the efficacy of schedule thinning. To date, there are no standards that evaluate the strength of the evidence specifically for schedule thinning. Thus, the WWC Evidence Standards were selected as a template for the proposed standards, as they are widely used to determine intervention effect across the special education and behavior analytic literature. The current standards also incorporated features from the design standards developed by Neely, Garcia, Bankson, and Green (2018). In their review of FCT studies that reported maintenance and generalization data, Neely et al. proposed a set of design standards, based on the WWC standards that were modified to evaluate maintenance and generalization data. The authors also proposed a method for categorizing study outcomes based on specific features of visual analysis which have been incorporated in some of the evidence standards described below.

The standards created for this review consist of four standards and were applied to all experiments that demonstrated evidence of strong or moderate effectiveness during the initial treatment evaluation. Each standard was then applied across three categories. Categories included overall effect, initial effect, and terminal effect. The overall effect considered all data points within the schedule thinning phase. Standards within the initial and terminal effect categories were applied to the first and last 30% of data points of the schedule thinning phase respectively.

Standard 1(Between Phase Effect-Treatment) was used to determine the presence of an effect between the last phase of the initial treatment evaluation and the schedule thinning phase

(Kratochwill et al., 2010; Maggin et al., 2013). The last treatment phase was selected to ensure that comparisons were made across adjacent phases, which is recommended for visual and statistical analysis of single-case data (Kratochwill et al., 2010; Maggin et al., 2013; Parker, Vannest, Davis, & Sauber, 2011; Rakap, 2015). To meet criteria 1 with evidence of a strong effect (2), 80-100% of all data points in the schedule thinning phase had to overlap with the data in the last treatment phase. In cases where 79-51% of data points in the schedule thinning phase overlapped with the data in the last treatment phase, the experiment was given a score of 1. If less than 50% of data points in the schedule thinning phase overlapped with the data in the last treatment phase, the experiment was given a score of 0 (Parker, Vannest, & Brown, 2009). The WWC Evidence Standards, the Neely et al. standards, and current effect sizes in single-case research all consider the percentage of overlapping or non-overlapping when determining the strength of an intervention's effect. Thus, overlap was selected as the primary indicator of effect for schedule thinning in the current review.

Standard 2 (Between Phase Effect-Baseline) was used to determine the presence of an effect between the last baseline phase and the schedule thinning phase. To receive a rating of strong effect, no data in the schedule thinning phase could overlap with data in the last baseline phase (Neely et al., 2018). Experiments were given a rating of moderate effect, if between 1% and 30% of data points in the schedule thinning phase overlapped with data in the last baseline phase. If more than 30% of data points in the schedule thinning phase overlapped with data in the last baseline phase, the experiment was given a rating of no effect.

Standard 3 evaluated the number of data points in the schedule thinning phase and was only applied to category 1 (overall effect). Given that, there are no standards to guide the number of data points needed for schedule thinning, Standard 3 was developed based on existing

standards. Experiments were given a score of 2 if the schedule thinning phase contained five or more data points (Kratochwill et al., 2010; Maggin et al., 2013). If the schedule thinning phase included 3-4 data points, the experiment was given a score of 1 (Kratochwill et al., 2010; Maggin et al., 2013; Neely et al., 2018). If less than three data points were included in the schedule thinning phase, Standard 3 was given a score of 0 (Kratochwill et al., 2010; Maggin et al., 2013).

Standard 4 was used to determine the overall effect of schedule thinning within each category. Each category was given an overall rating of strong effect (2), moderate effect (1), or no effect (0) (Kratochwill et al., 2010; Maggin et al., 2013). To obtain a rating of strong effect, all evidence standards had to receive a score of 2 (Kratochwill et al., 2010; Maggin et al., 2013). To receive a rating of moderate effect, all evidence standards had to be given a score of 2 or 1, and no ratings of 0 (Kratochwill et al., 2010; Maggin et al., 2013). If any standard was given a score of 0, the experiment was given an overall rating of no effect (Kratochwill et al., 2010; Maggin et al., 2013). A detailed description of the schedule thinning evidence standards can be found in Table A4.

Descriptive Evaluation

Coding variables and procedures.

Studies that demonstrated moderate or strong evidence of effectiveness during the schedule thinning evaluation were coded for specific descriptive information. Descriptive data were collected across nine categories. Categories included: (a) participant demographics, (b) function of challenging behavior, (c) intervention type, (d) reinforcement schedule, (e) schedule thinning procedures, (f) procedural modifications, (g) initial schedule, (h) terminal schedule, and (i) treatment fidelity.

Participant gender was coded as male or female. Participant age was coded as young child (ages 0-5), child (ages 6-10), adolescent (ages 11-17), or adult (ages 18 and older). Disability status was coded as ID only, ASD only, other DD only, or multiple. To be coded as multiple, the participant had to be diagnosed with two or more disabilities (e.g., ID and ASD, ID and Down syndrome, etc.). Intervention type was coded as DRA/I only, DRO only, other differential reinforcement procedure only, FCT only, NCR only, or treatment package. DRA/I was defined as a reinforcement-based procedure in which consequences were provided contingent on the emission of an appropriate alternative or incompatible response.

DRO was defined as a reinforcement-based procedure in which consequences were provided contingent on the emission of any behavior other than the target challenging behavior, after a predetermined period of time (Weston et al., 2018; Wong et al., 2014). Other differential reinforcement procedures included differential reinforcement of low rates, differential reinforcement of diminishing rates, or differential reinforcement of high rates of behavior (Cooper et al., 2007). FCT was defined as a procedure in which reinforcement was provided contingent on an appropriate communicative response that matched the function of challenging behavior, determined via a pretreatment FBA (Carr & Durand, 1985). NCR was defined as a procedure in which reinforcement was provided on a predetermined schedule independent of the target behavior (Vollmer et al., 1993). Treatment packages were defined as a combination of any of the above-listed procedures or any of the listed procedures combined with additional treatment components (e.g., choice making, visual supports, self-monitoring).

Function of challenging behavior was coded as attention, automatic, escape, tangible, or other. Reinforcement schedule was coded as response or time-based. Response schedules were defined as an arrangement in which reinforcement was provided contingent on the emission of a

specific number of responses. Time schedules were defied as the provision of reinforcement following the passage of a predetermined period of time.

Thinning procedures were coded as: (a) delay-to-reinforcement, (b) demand fading, (c) interval increase, (d) multiple schedule, (e) response restriction, (f) combination, or (g) other. Delay-to-reinforcement was defined as an increase in the time between the emission of the alternative behavior and the onset of reinforcement (Hagopian et al., 2011; LeBlanc et al., 2002). Demand fading was defined as an increase in the work requirement necessary to produce reinforcement (Hagopian et al., 2011). Interval increase was defined as an increase in the duration of time before reinforcement was provided. Multiple schedules were defined as alternating between two reinforcement schedules each signaled with a different discriminative stimulus, one representing the availability of reinforcement and the other representing that reinforcement was unavailable (Hagopian et al., 2011). Response restriction was defined as a period in which access to communication materials (e.g., picture cards, speech-generating devices, etc.) were unavailable (Hagopian et al., 2011). Procedures that did not align with the descriptions listed above (e.g., prompt fading, manipulating parameters of reinforcement) were coded as other.

Procedural modifications were defined as any variations made to the schedule thinning procedure based on the definitions listed above. Procedural modifications were coded descriptively based on author report. Initial schedule was defined as the first response requirement or interval necessary to obtain reinforcement. Terminal schedule was defined as the final response requirement or interval necessary to obtain reinforcement. Both initial and terminal schedule was defined based on author report. Treatment fidelity was rated as yes (fidelity data reported) or no (fidelity data not reported).

Inter-rater Agreement

To obtain agreement data, secondary raters replicated the inclusion, design, evidence, descriptive, and the schedule thinning evidence evaluations. Raters were doctoral students in special education or Assistant professors in special education with experience conducting systematic reviews. All raters were trained in the scoring procedures by the primary rater using sample articles. Training was considered complete when the primary and secondary raters reached a minimum of 80% agreement across three training articles. Percent agreement for all phases of the review was calculated by multiplying the number of variables scored as agreements divided by the total number of variables and multiplying by 100.

Inclusion evaluation.

Inter-rater agreement (IRA) was collected for 30% of articles obtained via the database and ancestral searches. Agreements were defined as both raters assigning the same rating to the article (i.e., include or exclude). Percent agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100.

Design, evidence, and descriptive analyses.

IRA was collected for 30% of the included articles. Agreements were defined as both raters assigning the same rating to each criterion or variable (e.g., “meets standards” for the design evaluation, or “young child” for the descriptive evaluation). Percent agreement was calculated by dividing the total number of agreements by the total number of coding variables and multiplying by 100.

Schedule thinning evidence evaluation.

IRA was collected for 30% of experiments included in the schedule thinning evidence evaluation. Agreements were defined as both raters assigning the same score to each criterion. Percent agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100.

Results

Interrater Agreement

IRA scores across all phases of the study were above minimum thresholds. IRA scores are displayed for each phase of the review in Table A5.

Article Identification and Inclusion Criteria

A total of 1,929 articles were obtained via the electronic database search. Of those articles, 147 were duplicates and were excluded from further review. Thus, a total of 1,782 articles were retained and included for final review. An additional 27 articles were identified from the ancestral search of articles that met inclusion criteria. Thirty-one additional articles were identified via the hand search of previous literature reviews, and two articles were obtained from the hand search of educational journals. In total, 1,842 articles were identified for the current review. Of those articles, 173 met all nine inclusion criteria and were evaluated against the WWC Basic Design Standards.

Quality Review

Results of the quality evaluation are displayed in Table A6. All 173 (100%) of the included studies met Design Standard 1, and experimentally manipulated the independent

variable. The majority of studies (*n*=166; 96%) reported IAA data and met Design Standard 2A. Twenty-four studies (14%) reported IAA data across baseline and intervention phases independently and met Design Standard 2B without reservations. Most studies (*n*=120; 69%) reported IAA data as an aggregate across phases and met Standard 2B with reservations. Twenty-nine studies (17%) failed to report IAA data and were given a rating of does not meet standards. Of the 145 studies that reported IAA data, 135 (93%) reported IAA scores that met or exceeded minimum thresholds (≥80% or 0.6 Kappa) and met Design Standard 2C. The remaining 10 studies (7%) reported IAA scores below minimum thresholds across one or more study phases and were rated as not meeting standards for standard 2C. Seventy-two studies (42%) demonstrated a minimum of three attempts to demonstrate an experimental effect during the initial treatment evaluation and were rated as meets standards without reservations for Design Standard 3. However, the majority of studies (*n*=101; 58%) provided less than three attempts and did not meet Design Standard 3. Over a quarter of included studies (*n*= 63; 36%) met Design Standard 4 without reservations, and included a minimum of five data points in each study phase. Most studies (*n*= 85; 49%) met Design Standard 4 with reservations and included 3-4 data points within each study phase. Twenty-three studies (13%) included less than three data points in at least one study phase, and received a rating of does not meet standard. Overall, three studies (2%) met standards without reservations, 43 studies (25%) met standards with reservations, and 125 studies (72%) were rated as not meeting standards. Overall, 46 studies met the Basic Design Standards and were evaluated against the WWC Evidence Standards.

Evidence Evaluation

A total of 76 unique experiments were included across the 46 studies that met the Basic Design Standards with or without reservations. Each unique experiment was evaluated against

the 22 Evidence Standards to determine the strength of the functional relation between the reinforcement-based intervention and challenging behavior. Results of the evidence evaluation can be found in Table A7. Sixteen experiments (21%) included five or more data points per study phase and were rated as having strong evidence of effect for Criteria 1. Most experiments (*n*= 60; 79%) provided 3-4 data points per study phase and were rated as having moderate evidence of effect. No studies included fewer than three data points per phase. For Criteria 2, 86% of studies (*n*= 65) demonstrated at least three instances of effect and were rated as having strong evidence. Eleven experiments (14%) provided less than three demonstrations and were found to have no evidence of effect. More than 80% of the included experiments (*n*= 64) demonstrated no instances of non-effect between baseline and treatment phases and were rated as having strong evidence for Criteria 3. For only one study (1%), the ratio of effects to non-effects was less than or equal to 3:1, and was scored as having moderate evidence of effectiveness. For 11 experiments (14%), the ratio of effects to non-effects was greater than 3:1. Overall, 12 experiments (16%) were found to have strong evidence of effectiveness, 53 experiments (70%) were found to have moderate evidence of effectiveness, and 11 experiments (14%) demonstrated no evidence of effect. One study was rated as strong evidence, but the treatment and schedule thinning data were presented on separate graphs, and the units of measurement did not align between experiments. Thus, this article was excluded for further review. The remaining experiments that demonstrated evidence of strong or moderate effectiveness were included in the schedule thinning evidence evaluation.

Schedule Thinning Evidence Evaluation

Eighty-two unique evaluations of schedule thinning were included across the 76 experiments that met the WWC Evidence Standards. Each experiment was evaluated against four

standards to determine the presence and strength of the functional relation between schedule thinning and challenging behavior. The overall, initial, and terminal effects of each schedule thinning evaluation are described below and displayed in Table A8.

Overall effect.

For evidence standard (ES) 1, 49 experiments (60%) demonstrated strong evidence, with 80-100% of data points in the schedule thinning phase overlapping with the data in the last treatment phase. Fourteen experiments (17%) were found to have moderate evidence of effectiveness, with 79-51% of data points in the schedule thinning phase overlapping with data in the last treatment phase. Nearly a quarter of included experiments (*n*=19; 23%) showed less than 50% overlap between phases and were rated as demonstrating no evidence of effectiveness. The majority of experiments (*n*= 53; 65%) were found to have strong evidence for ES 2 and showed no overlap between the schedule thinning and baseline phases. Twenty experiments (24%) were found to have moderate effectiveness, with no more than 30% of data points in the schedule thinning phase overlapping with data in the last baseline phase. Only 19 experiments (11%) were rated as having no evidence of effectiveness. Nearly all of the experiments (*n*= 77; 94%) demonstrated strong evidence for ES 3 and included five or more data points in the schedule thinning phase. Four experiments (5%) included 3-4 data points, and only one experiment (1%) included fewer than three data points. Overall, most of the included experiments (*n*= 59; 72%) provided strong or moderate evidence of effectiveness for schedule thinning. In total, 33 experiments (40%) provided strong evidence of effectiveness, 26 studies (32%) demonstrated moderate effectiveness, and 23 studies (28%) provided no evidence of efficacy for schedule thinning.

Initial effect.

The initial effect of schedule thinning was determined by evaluating the first 30% of data points in the schedule thinning phase. Evidence standards one and two were applied to each experiment to determine the strength of the functional relation between schedule thinning and challenging behavior. Most of the included experiments (*n*= 52; 63%) provided strong evidence of effect for ES 1. Thirteen experiments (16%) showed moderate evidence of effectiveness and 17 experiments (21%) demonstrated no evidence of effectiveness. Fifty-eight experiments (71%) demonstrated strong evidence of effectiveness with no overlap between the first 30% of data points and data in the baseline phase. Eight experiments (10%) were rated as having moderate evidence of effectiveness, and 16 studies were rated as having no evidence of effectiveness. Of the 82 included experiments, just over half (*n*= 45; 55%) provided strong evidence for the efficacy of schedule thinning during the initial stages. Eleven experiments (13%) were rated as having moderate evidence of effectiveness and just over one quarter (*n*= 26; 32%) of experiments provided no evidence for the efficacy of schedule thinning during the initial stages.

Terminal effect.

Overall, 48 experiments (59%) were found to have strong evidence of efficacy when the terminal schedule of reinforcement was in effect. Fourteen experiments (17%) were found to have moderate evidence of effectiveness, while 20 experiments (24%) demonstrated no evidence of effectiveness. Most experiments (*n*= 56; 68%) demonstrated strong evidence of effectiveness for ES 1 with 80-100% of data points in the schedule thinning phase overlapping with data in the initial treatment phase. Of the remaining 26 experiments, 13 (16%) were rated as having moderate evidence of effectiveness, and 13 (16%) were found to have no evidence of effectiveness. Nearly all of the included experiments were rated as having strong (*n*= 64; 78%) or

moderate (*n*= 8; 10%) evidence of effectiveness for ES 2. Only 10 experiments were found to have no evidence of effectiveness due to high levels (≥30%) of overlap between the schedule thinning and baseline phases.

Descriptive Evaluation

The 59 experiments that demonstrated moderate or strong evidence of effectiveness for schedule thinning were coded for specific descriptive information across nine categories. Results of the descriptive evaluation are presented in Table A9.

Participant demographics.

A total of 51 unique participants were included across the 59 experiments that met the schedule thinning ES with moderate or strong evidence of effectiveness. Of those participants, thirty-three participants (65%) were classified as male, and 15 (29%) were categorized as female. Gender was not reported for three of the participants (6%). Most of the participants (*n*= 22; 43%) were categorized as a child (i.e., between the ages of 6-10 years). Adults and adolescents were each included in 20% of experiments (*n*= 10). Young children (i.e. age's birth to five years) were included least frequently across studies, with only nine participants (18%) included in this category. Nearly half of the participants (*n*= 24; 47%) were diagnosed with multiple developmental disorders (e.g., ASD and ID). Just over one quarter (*n*= 14; 27%) of the participants were described as having a diagnosis of ASD only. Only five participants (10%) were described only as having an ID only; however, the severity of ID (e.g., mild, moderate, or profound) varied across participants. Specific diagnostic information was not reported for four (8%) participants.

Function of challenging behavior.

Social positive reinforcement was the most commonly reported operant function of challenging behavior across participants (*n*= 32; 63%). A tangible function was reported for 18 participants (35%), and an attention function was reported for 14 participants (34%). Social negative reinforcement in the form of escape from aversive task demands was reported as the operant function of challenging behavior for four participants (8%). An automatic function was reported for three participants (6%), and multiple functions were reported for seven participants (14%). Other operant functions of challenging behavior including activity interruptions (Falcomata, Roane, Muething, Stephenson, & Ing, 2012; Hagopian, Bruzek, Bowman, & Jennett, 2007) and access to stereotypy (Falcomata, Roane, Feeney, & Stephenson, 2010), were reported for five participants (10%).

Reinforcement schedule.

Interventions using interval, or time-based schedules of reinforcement were implemented most frequently across participants (*n*= 26; 51%). Interventions in this category included NCR and DRO. Ratio or response-based interventions were implemented for 24 participants (47%). Interventions in this category consisted of FCT and DRA. Two studies, including two participants (4%), implemented multiple interventions that were classified both as response and time-based.

Intervention type

Time-based interventions.

Twenty-three participants were exposed to interventions that utilized interval or time-based schedules of reinforcement. Approximately 40% of the participants (*n*= 19) participated in

a NCR intervention. For the majority of these participants, NCR was combined with an extinction component. Four participants (8%) participated in a DRO evaluation with or without the use of extinction.

Response-based interventions.

The most commonly implemented response-based intervention was FCT. For 23 participants (45%), FCT with or without extinction was implemented as the primary intervention. Other DRA interventions were only implemented for one participant (2%). Treatment packages, which were defined as the application of two or more response or time-based interventions, were implemented for four participants (8%). For example, Lambert et al. (2018) utilized a treatment package consisting of DRA, extinction, and a token economy to increase latency to challenging behavior for a seven-year-old female diagnosed with Prader-Willi syndrome. During baseline, latency to food stealing averaged 3 s. Immediately following the introduction of the treatment package, latency to challenging behavior increased to an average of 17.6 s.

Schedule thinning procedures

Time-based procedures.

Increasing intervals (i.e., increasing the duration of the NCR or DRO interval) was the most commonly utilized schedule thinning procedure across participants. This procedure was used to facilitate schedule thinning during NCR and DRO for a total of 24 participants (47%).

Response-based procedures.

Multiple schedules and demand fading were utilized to facilitate schedule thinning following FCT for nine (18%) and four participants (8%) respectively. Response restriction was

used most infrequently across experiments. Only two participants (4%) were exposed to response restriction following FCT. Delay-to-reinforcement was used for nine participants (18%) following response-based interventions (i.e. DRA, FCT). A combination of schedule thinning and or delay-to-reinforcement procedures were used to reduce the intensity of reinforcement for three participants (6%). For example, Falcomata et al. (2012) utilized a delay-to-reinforcement procedure with demand fading, to facilitate tolerance for delays to reinforcement following FCT for two children diagnosed with ASD. During the delay, the children were required to complete a series of academic demands. As the delay interval increased, the students were required to complete an increasing number of task demands. For two participants (4%), other procedures were used to facilitate schedule thinning including, mixed schedules.

Procedural modification.

For fourteen participants (27%), the initial schedule thinning procedure was modified. Modifications were either (a) embedded in the initial schedule thinning evaluation, (b) added contingent on the persistence of challenging behavior following treatment, or (c) included at the end of the schedule thinning phase to enhance the feasibility of the procedure in the participant's typical environment. Procedural modifications were made most frequently when a delay-to-reinforcement was implemented following FCT (*n*= 9; 64%). Modifications included providing alternative activities or work tasks during the delay. For the remaining five participants (36%), procedural modifications were made during NCR or DRO schedule thinning. Modifications for these participants included the use of punishment (e.g., contingent timeout), NCR, noncontingent attention, and adjusting the magnitude of the reinforcer. Combinations of the above-listed procedures were used for two (14%) participants.

Initial schedule

Time-based interventions.

The initial reinforcement schedule varied considerably across participants and interventions. For NCR, reinforcement was delivered on an FT-10 s schedule most commonly during the initial stage of schedule thinning. However, the initial reinforcement schedule ranged from continuous to FT-100 s across participants (*M*= 33 s). Considerable variation in the initial reinforcement schedule was also observed across the four experiments that implemented DRO. The initial DRO interval ranged from 0 s to 300 s (*M*= 98 s), with no two experiments implementing the same schedule.

Response-based interventions.

For experiments that evaluated multiple schedules, the most common initial duration of the extinction or S^{Δ} component was 60 s. However, the duration of the S^{Δ} component ranged from 0 s to 60 s across experiments (*M*= 35 s). For demand fading, the initial schedule across experiments was an FR-1. A 2 s restriction period was implemented as the initial schedule for experiments that implemented a response restriction procedure. The initial duration of the delay for experiments that implemented a delay-to-reinforcement was 5 s. Similar to the other procedures described above, the duration of the initial delay varied markedly across experiments ranging from 0 s to 300 s (*M*= 14 s).

Terminal schedule

Time-based interventions.

For studies that implemented NCR, the most common terminal reinforcement schedule was FT- 5 min. However, the terminal schedule of reinforcement varied across experiments

ranging from FT-30 s to FT-10 min (*M*= 3 min 45 s). Different terminal schedules were implemented across the four experiments that implemented DRO. The average terminal schedule for DRO was 7 min, with a range of 3 min to 15 min.

Response-based interventions.

For studies that evaluated the effects of multiple schedules, the most common durations of the S^{Δ} component were 4 min, 5 min, and 15min. Duration of the S^{Δ} component ranged from 4 min to 15 min, with an average duration of 8 min 40 s across experiments. A terminal FR-20 schedule of reinforcement for demand fading was implemented most commonly across experiments. However, final schedules ranged from FR-20 to VR-7. The duration of restriction for studies that implemented response restriction averaged 7 min 30 s, with a range of 5 min to 10 min. A 5 min delay was implemented most frequently as the terminal delay across experiments that implemented a delay-to-reinforcement procedure. Across experiments, the terminal delay ranged from 15 s to 9 min, with an average delay of 4 min.

Treatment fidelity.

Data on the experimenter's adherence to study procedures were only reported for seven of the 51 participants (14%). It is important to note that for most of these studies, it was unclear whether treatment fidelity was collected on experimenter behavior during the initial treatment evaluation or the schedule thinning evaluation.

Discussion

The purpose of the current review was to evaluate the available literature on schedule thinning following reinforcement-based intervention and to summarize the characteristics of studies that demonstrated strong or moderate evidence of effectiveness for schedule thinning. The current

review extends the literature by providing the first comprehensive review of schedule thinning, and by piloting a set of standards to evaluate the evidence for these procedures. Major findings from the evidence and descriptive evaluations, as well as limitations and directions for future research, are described below.

Evaluation of Schedule Thinning Procedures

Five common procedures were used to decrease the intensity of reinforcement across studies. The most commonly used procedure to facilitate schedule thinning for interventions using time-based reinforcement schedules was increasing the duration of the NCR or DRO interval. For response-based interventions, multiple schedules and demand fading were reported most frequently. Delay-to-reinforcement was also consistently reported across studies that implemented FCT. Response restriction was implemented infrequently across studies. Although this approach has been shown to be effective, there are ethical considerations that should be made before implementing this procedure. Although schedule thinning (including delay-to-reinforcement) is commonly used to reduce the intensity of reinforcement, other empirically supported procedures, such as manipulating parameters of reinforcement, were reported infrequently across studies. In fact, only one experiment evaluated the effects of manipulating the magnitude of reinforcement during schedule thinning. Compared to the considerable research on the procedures described above, there is a dearth on knowledge of how parameters of reinforcement impact the efficacy of schedule thinning.

Schedule Thinning Evidence Evaluation

Of the 82 experiments included in the schedule thinning evidence evaluation, 72% demonstrated strong or moderate evidence of effectiveness. This finding is promising given that

schedule thinning is often recommended following reinforcement-based intervention (Cooper et al., 2007; Hagopian et al., 2011; LeBlanc et al., 2002; Mayer et al., 2014). Previous reviews on schedule thinning have been intervention-specific (Hagopian et al., 2011), or have been descriptive (LeBlanc et al., 2002) limiting the generalizability of findings to other reinforcement-based procedures. Additionally, previous reviews did not use objective measures to classify the effects of schedule thinning. Thus, the results of the current review provide initial objective evidence that schedule thinning can be used to maintain low levels of challenging behavior following reinforcement-based intervention.

The evidence standards developed for this review are unique in that they were designed to measure treatment efficacy during different points in intervention (i.e., initial and terminal phases of schedule thinning/delay-to-reinforcement). To the author's knowledge, these are the first set of evidence standards that allow for such an evaluation. These standards mark an important contribution to the literature, given that many researchers have noted that undesirable changes in participant behavior may occur at one or more points during schedule thinning or delays-to-reinforcement. Thus, evaluating changes in participant behavior during different phases of schedule thinning or delay-to-reinforcement is critically important. Such evaluations may assist researchers in understanding the extent to which schedule thinning is effective, and help determine at what point during schedule thinning procedural modifications are warranted.

Nearly three-quarters of the included experiments (68%) included in this review demonstrated strong or moderate evidence of efficacy during the initial phases (i.e., first 30% of data points) of schedule thinning. This finding was not surprising given that when schedule thinning is first implemented, the schedule of reinforcement is typically relatively dense (e.g., FR-2) or the duration of the delay is brief (e.g., 5 s). Thus during the initial phases, schedule

thinning often resembles the initial reinforcement schedule used during intervention. Strong evidence was also found for schedule thinning when the final reinforcement schedule was in effect. Most of the included experiments (76%) demonstrated moderate or strong evidence of effectiveness for schedule thinning during the terminal phases (i.e., last 30% of data points). These findings provide support that schedule thinning can facilitate maintenance of treatment effects over time.

Although the majority of experiments provided moderate or strong evidence of efficacy for schedule thinning, approximately one-quarter of experiments were found to have no evidence of effectiveness. The most common reason experiments were rated as having no evidence of effectiveness during the initial and terminal phases of schedule thinning, was overlap between the schedule thinning and baseline phases. More than 30% overlap was reported for 19 experiments (11%) during the initial schedule thinning phase and 10 experiments (12%) when the terminal schedule was in effect. Overlapping data with baseline during the schedule thinning phase indicate resurgence of challenging behavior, which is detrimental to the sustainability of the intervention (Hagopian et al., 2011; LeBlanc et al., 2002).

There are several reasons why challenging behavior returns to baseline levels during schedule thinning. Sharp increases in challenging behavior during schedule thinning may indicate that the schedule of reinforcement was thinned too quickly. Although gradual decreases in the density of reinforcement are recommended, there is limited empirical evidence available to guide on how quickly, or in what manner, changes to reinforcement schedules should be made (LeBlanc et al., 2002). In fact, decisions to alter the rate or intensity of reinforcement are often made based solely on professional judgement (LeBlanc et al., 2002). Although professional judgement and individual client factors are important when determining at what level of intensity

reinforcement should be delivered, additional research examining methods for systematically decreasing the intensity of reinforcement is warranted.

The procedures used to decrease the intensity of reinforcement may also influence if resurgence of challenging behavior occurs. For example, considerable research has shown that challenging behavior is likely to reoccur when a delay-to-reinforcement is implemented (Fisher et al., 2000; Hagopian et al., 2011; Hanley et al., 2001). The absence of reinforcement during the delay may resemble an extinction condition, which results in increases in previously reinforced challenging behaviors. To reduce the risk of resurgence, researchers have recommended providing alternative stimuli during the delay (Austin & Tiger, 2015; Fisher et al., 2000; Hanley et al., 2001; Hagopian et al., 2005), using contingency statements before the delay (Logue, 1988), and increasing the delay-to-reinforcement gradually (Logue, 1988). Although some comparative studies have been conducted, few studies have been conducted that systematically compare two or more schedule thinning procedures. As such, conclusive statements regarding which schedule thinning procedures are most effective cannot be made at this time.

Given that nearly one-quarter of studies were found to have no evidence of effectiveness during the initial or terminal stages of schedule thinning, the available evidence demonstrates that in some cases, modifications to schedule thinning or delay-to-reinforcement should be made. However, few studies reported procedural modifications. Thus, additional research exploring (a) when procedural modifications should be made and (b) the efficacy of various procedural adaptations should be conducted.

Descriptive Evaluation

Results of the descriptive analysis revealed several important themes and gaps in the schedule thinning literature. First, schedule thinning was evaluated most frequently for children

with DDs. Only 20% of the participants included across experiments were classified as adults. This finding aligns with previous reviews indicating that adults with DDs are significantly underrepresented in the behavior analytic literature (Gerow et al., 2018; Heath, Ganz, Parker, Burke, & Ninci, 2015). Additional research evaluating the impact of schedule thinning on challenging behavior for adults is needed given that adults with DDs have longer learning histories and the mechanisms that influence the efficacy of schedule thinning may differ between children and adults.

Results of the descriptive evaluation indicated that schedule thinning is used relatively infrequently following DRO and DRA interventions. Only four of the included experiments evaluated DRO schedule thinning. Additionally, only one experiment evaluated schedule thinning following an evaluation of DRA. Results also showed that schedule thinning is implemented less often for individuals whose challenging behavior is escape or automatically maintained, compared to individuals with challenging behavior maintained by social positive reinforcement. It is important to note that many studies utilizing DRO and DRA or studies that treated escape and automatically maintained challenging behavior were initially included in this review but were excluded following the quality evaluation, after failing to meet one or more of the WWC Basic Design Standards.

Although not a primary purpose of this study, of the 173 studies included in this review, only 46 met the minimum thresholds to be considered methodologically sound by the WWCs Basic Design Standards. The most common reason studies failed to meet the Basic Design Standards was for providing less than three attempts to demonstrate an experimental effect between baseline and intervention. Although reinforcement intensity is altered during schedule thinning, the primary intervention (e.g., FCT, NCR) is still in place. Thus, evidence of strong

internal validity during the initial treatment phase increases the believability that a functional relation was demonstrated between schedule thinning and the outcome variable.

One of the primary objectives for decreasing the intensity of reinforcement is to reach a schedule of reinforcement that can be maintained in natural settings (e.g., classrooms, community-based programs). One noteworthy finding from this review is that the terminal schedule or delay-to-reinforcement was relatively thin. For example, the duration of the terminal delay in some studies that evaluated delays-to-reinforcement reached 15 min. This is significant given that reinforcement is often delivered continuously, or almost continuously, during the initial treatment evaluation. However, there are contexts in which extended delays or denials to reinforcement will occur and far exceed 15 min (e.g., parent preparing dinner, sibling doing homework on the only family computer). None of the included studies examined the impact of schedule thinning or delays-to-reinforcement during situations in which an individual was expected to tolerate long periods of non-reinforcement. This is an important area for future research given that delays or denials to reinforcement often occur multiple times a day.

Finally, data on treatment fidelity were scarcely reported across the included studies. This finding was not surprising given that previous systematic and quality reviews have reported similar findings (Gregori et al., 2018). However, this concerning given that treatment fidelity is critical in determining the presence of a functional relation between an independent and dependent variable (Smith, Daunic, & Taylor, 2007). Without data on treatment fidelity, it is impossible to ascertain whether changes in participant behavior were the result of the target intervention, or unrelated factors (Bellg et al., 2004). Data on treatment fidelity is especially important in evaluations of schedule thinning, given that errors in fidelity are more likely to occur when the intensity of reinforcement is decreased (LeBlanc et al., 2002).

Limitation and Future Directions

The current review is not without limitations. Results of the evidence evaluation should be interpreted with caution, given that (a) this was the first application of these standards and (b) there were limitations with the standards themselves. The standards used visual analysis of overlap as the primary indicator of treatment efficacy. Existing standards used to determine the efficacy of maintenance and generalization data, as well as most single-case effect size measures (e.g., Tau-U), consider overlapping or non-overlapping data as a primary measure of treatment effect, which is why analysis of overlap was a primary consideration for the standards used in this review. However, other features, including changes in level, trend, and variability that are hallmarks of visual analysis, were not incorporated into these standards. Additionally, because overlap was the primary metric, it is possible that outcomes could be influenced by one or more outlying data points in the baseline or intervention phase. Refinements to the current standards should be made to address such issues. Despite the limitations, the evidence standards developed for this review are novel and can assist in providing important information regarding the efficacy of schedule thinning and delay-to-reinforcement.

Although the current review provided an evaluation of the overall efficacy of schedule thinning, an analysis of variables that influence the efficacy of schedule thinning was not conducted. An analysis of potential moderating variables may provide important information regarding the contexts and conditions in which schedule thinning is most effective. Thus, a meta-analysis of schedule thinning and delay-to-reinforcement following reinforcement-based intervention is an important next step in research on this topic. Additionally, meta-analysis utilizes quantitative metrics to determine treatment efficacy. These metrics provide more

objective measures of treatment efficacy, compared to measures based on visual analysis, such as the ones used in the current review.

The current review only evaluated the impact of schedule thinning and delay-to-reinforcement on challenging behavior. A common side effect of these procedures following response-based interventions is that as the delivery of reinforcement becomes less intense, increases in the alternative response can occur at excessively high rates (Hanley et al., 2001; LeBlanc et al., 2002). This is problematic because if emitted at high rates, the alternative behavior can itself become a challenging behavior (LeBlanc et al., 2002). Thus, future research should examine the impact of schedule thinning on appropriate alternative and communicative behavior. Last, this review only included peer-reviewed articles. Therefore, it is possible that other studies, including dissertations, which conducted experimental evaluations of schedule thinning, were not identified.

CHAPTER III: EFFECTS OF ACTIVITY CHOICE ON CHALLENGING BEHAVIOR DURING DELAYS-TO-REINFORCEMENT

Introduction

Functional Communication Training

Functional communication training is one of the most commonly utilized interventions for the treatment of challenging behavior among persons with IDD (Tiger et al., 2008). FCT is a variation of DRA that involves replacing challenging behavior with a socially appropriate communicative response (FCR) that produces the same reinforcer as the challenging behavior (Carr & Durand, 1985). FCT is often combined with extinction, meaning that challenging behavior no longer produces access to reinforcement. Typically, FCT begins with a pre-treatment FA to identify the variables maintaining the challenging behavior(s) (Iwata et al., 1982/1994). For example, if a FA determines that access to tangible items (e.g., tablet device) reinforces an individual's aggression, the FCT intervention would consist of teaching the individual to say, "Tablet please," instead of engaging in aggression.

Social positive reinforcement in the form of access to tangible items or attention is often responsible for the maintenance of challenging behavior among individuals with IDD. Petursdottir, Esch, Sautter, and Stewart (2010) reviewed the behavior plans of 174 individuals diagnosed with DDs. Behavior plans contained relevant demographic information as well as the functional assessment procedures used to determine the operant function of challenging behavior. Results of the review found that nearly 5% of the sample engaged in tangibly maintained challenging behavior and 30% of the sample engaged in attention maintained challenging behavior. Gerow et al. (2018) conducted a review of the available literature on FCT across disability categories and found that of the 135 participants included, 21% engaged in

challenging behavior maintained by access to tangible or edible items. Another 18% of the participants engaged in challenging behavior maintained by social positive reinforcement in the form of attention. The results of these reviews suggest that socially maintained challenging behavior is common among individuals diagnosed with IDD.

The effects of FCT on challenging behavior and communication for individuals with IDD have been well documented in the behavior analytic literature (Chezan, Wolfe, & Drasgow, 2017; Gerow et al., 2019; Tiger et al., 2008; Wong et al., 2014). FCT has been evaluated in over 300 empirical studies and is currently considered an EBP for children and adolescents with ASD and a promising practice for adults with IDD, including ASD (Heath et al., 2015; Wong et al., 2014). Byiers, Dimian, & Symons (2014), for example, evaluated the effects of FCT on challenging behavior and communication for three women diagnosed with Rett syndrome. A pre-treatment FA indicated that challenging behavior was sensitive to social-positive reinforcement in the form of access to caregiver attention and access to preferred tangible items. Prior to the implementation of FCT, levels of challenging behavior were high, and levels of appropriate communication were low for all three women. During intervention (i.e., FCT), activation of a microswitch produced praise and 10-15 s access to the relevant reinforcers. Immediately following the implementation of FCT, challenging behavior reduced to near zero levels and levels of appropriate communication increased for all three women.

In another recent investigation, Fisher et al. (2015) examined the effects of FCT plus extinction on the tangibly maintained challenging behavior of a 10-year-old male diagnosed with multiple disabilities. Moderate levels of challenging behavior and no appropriate communication were observed prior to the implementation of FCT. Immediately following the introduction of

FCT, challenging behavior reduced to near zero levels, and appropriate communication was observed at moderate levels.

During the initial implementation of FCT, reinforcement is typically provided on a fixed-ratio 1 schedule (FR-1); meaning that each emission of the FCR is reinforced. Additionally, reinforcement is delivered immediately to increase the strength of the relationship between the FCR and reinforcement. This continuous schedule of reinforcement is critical in establishing the FCR in and in most cases leads to an immediate increase in appropriate communication (Byiers et al., 2014; Fisher et al., 2015; Graff, Libby, & Green, 1998; Heath et al., 2015; Volkert et al., 2009). Although a continuous schedule of reinforcement is vital during the initial stages of FCT, it can have several adverse side effects on both communication and challenging behavior. For example, continuous schedules of reinforcement can lead to excessive emissions of the FCR, which may prevent the individual from engaging in other activities or tasks (Hagopian et al., 2011).

The most significant issue associated with the use of continuous reinforcement during FCT is that these schedules are often not feasible in applied settings and are difficult for natural change agents to maintain (Fisher, Thompson, Hagopian, Bowman, & Krug, 2000; Sidener, Shabani, Carr, & Roland, 2006). Most applied settings, such as schools, day program centers, and homes, lack the necessary resources to maintain such a dense schedule of reinforcement. In these cases, caregivers may commit critical fidelity errors such as failing to reinforce the FCR (LeBlanc et al., 2002). Failure to reinforce the FCR can lead to deterioration of the FCR and reemergence of challenging behavior (Fisher et al., 2000; Hagopian et al., 2011). To address these issues and to increase the feasibility and sustainability of FCT, schedule thinning or delay-

to-reinforcement are often added as an additional treatment component following the initial implementation of FCT.

Delay-to-Reinforcement

Delay schedules consist of inserting a delay between the emission of the FCR and delivery of reinforcement (Hagopian et al., 2011; LeBlanc et al., 2002). The initial FR-1 schedule remains intact, but the delay-to-reinforcement gradually increases. For example, if an individual emits the FCR, the caregiver would tell the individual to wait for a predetermined period of time (i.e., the delay). After the delay elapsed, the caregiver would praise appropriate waiting and give the individual access to the reinforcer. During the delay, neither challenging behavior nor emissions of the FCR result in access to reinforcement. The initial delay is brief (e.g., 1 s) and increases until a terminal schedule is reached that resembles what is acceptable in the typical environment. The purpose of implementing a delay schedule is to maintain low levels of challenging behavior and acceptable levels of the FCR to increase the generality of FCT (LeBlanc et al., 2002).

Delay schedules are often used to facilitate tolerance for delays-to-reinforcement for challenging behavior maintained by access to social positive reinforcement. In their review of schedule thinning following FCT, Hagopian et al. (2011) found that all of the studies that applied delay schedules, included participants with tangibly maintained challenging behavior. Additionally, recently published studies on the use of delay schedules following FCT have all included participants whose challenging behavior was sensitive to social positive reinforcement in the form of access to tangibles or attention (Austin & Tiger, 2015; Ghaemmaghami et al., 2016; Lehardy, Lerman, Evans, O'Connor, & LeSage, 2013; Stevenson, Ghezzi, & Valenton, 2016).

Delay schedules are ideal for decreasing reinforcement intensity and teaching tolerance for delays-to-reinforcement because they emulate typical environments most closely. Despite their appeal, challenging behavior often resurges during the delay (Briggs et al., 2018; Fisher et al., 2000; Hagopian et al., 1998; Hanley et al., 2001; Jarmolowicz & Lattal, 2014). Extinction-induced resurgence is a phenomenon in which a previously reinforced behavior remerges when an alternative behavior is placed on extinction (Epstein, 1985). Although the FCR is not directly placed on extinction, the delay resembles extinction conditions due to the temporary period of non-reinforcement. Therefore, recurrence of challenging behavior is often observed as the delay-to-reinforcement increases (Fisher et al., 2000; Hagopian et al., 1998; Hanley et al., 2001; Jarmolowicz & Lattal, 2014).

Extinction-induced resurgence has been observed in a number of studies evaluating delays-to-reinforcement. Hagopian et al. (1998) evaluated the efficacy of FCT with extinction and delays-to-reinforcement in a large sample of individuals diagnosed with ID. The initial FCT treatment evaluation resulted in a 90% reduction in challenging behavior for approximately 50% of individuals. However, when the delay-to-reinforcement was added, challenging behavior resurged in over 50% of the cases. Additional intervention components, including punishment, were needed to achieve clinically significant reductions in challenging behavior. Hanley et al. (2001) also evaluated the effects of FCT and extinction and a delay-to-reinforcement for a 29-year-old woman diagnosed with profound ID. Similar to Hagopian et al. (1998), the initial FCT treatment resulted in significant decreases in challenging behavior. However, as the delay interval was lengthened, SIB increased to unacceptable levels and additional schedule thinning procedures were required before SIB returned to initial treatment levels.

Given that the effects of FCT often diminish during delays-to-reinforcement, researchers have begun exploring variations of the traditional delay schedule. Punishment-based procedures are often implemented during delays-to-reinforcement. For example, Fisher et al. (2000) evaluated the effects of FCT plus extinction for a 7-year-old female who engaged in tangibly maintained SIB and aggression. During the initial FCT evaluation, levels of challenging behavior remained high. Therefore, the experimenters added a punishment procedure in the form of a 30 s basket hold contingent on challenging behavior. The punishment procedure remained in effect during the delay-to-reinforcement and resulted in low levels of challenging behavior. Although punishment-based procedures often produce clinically significant reductions in challenging behavior, it is generally accepted among the behavior analytic community that reinforcement-based procedures should be implemented before the use of aversive procedures (Behavior Analyst Certification Board®, 2014). Therefore, more recent studies examining delay schedules have moved toward the use of additional non-aversive procedures during delays-to-reinforcement, such as antecedent interventions.

Advancements in Delay-to-Reinforcement

Antecedent interventions are procedures used to prevent the occurrence of challenging behavior by manipulating various environmental variables (Cooper et al., 2007; Mayer et al., 2014). One antecedent strategy that has gained popularity in the schedule thinning literature is environmental enrichment. In the context of schedule thinning, this procedure involves providing an individual with access to alternative highly preferred items and/or work tasks during the delay-to-reinforcement. Fisher et al. (2000) provided access to an alternative work task to a 19-year-old male diagnosed with ASD and profound ID during delays-to-reinforcement. Levels of inappropriate sexual behavior, head rolling, and other destructive behaviors remained low during

the delay and the terminal delay of 4 min was reached. The authors hypothesized that the work task served as a distractor and may have been responsible for reductions in challenging behavior.

Hagopian et al. (2005) extended the work of Fisher et al. (2000) and examined the effects of noncontingent access to competing reinforcers during delays-to-reinforcement for three children with DDs. A pretreatment FA was conducted to identify the functional reinforcers for each child. Next, a competing stimulus assessment was conducted to identify stimuli that competed with the functional reinforcer. During the delay-to-reinforcement, the children were given free access to the competing stimuli. Access to the stimuli, for all children, resulted in low levels of challenging behavior during the delay. In 2015, Austin and Tiger replicated and extended the work of Hagopian et al. (2005) by evaluating the effects of access to competing stimuli for a child with multiply controlled challenging behavior. Two conditions, access to attention, and access to tangibles were evaluated in a multiple-baseline design. During delays in the tangible condition, the participant was given access to adult attention. Conversely, during the attention delay sessions, the participant was given noncontingent access to preferred tangibles. Access to these stimuli resulted in low levels of challenging behavior. The authors suggested that access to alternative stimuli during the delays competed with the functional reinforcer and created an abolishing operation for challenging behavior.

In a more recent evaluation of delay-to-reinforcement, Ghaemmaghami et al. (2016) conducted a comparative analysis of time-based and contingency-based delay schedules across four individuals diagnosed with DDs. During the time-based delay condition, participants had access to a variety of leisure items and academic tasks. However, the participants were not required to interact with any of the items or tasks. During the contingency-based delay condition, a member of the research team prompted the participants to engage with the leisure materials or

comply with the work task. Reinforcement was contingent on compliance with the materials throughout the delay. For all participants, the contingency-based delay was more effective than the time-based delay in promoting tolerance for delays-to-reinforcement. The time-based delay resulted in higher levels of challenging behavior and excessive levels of the FCR. Additionally, resurgence of severe topographies of challenging behavior, including SIB, was observed for some participants. Alternatively, the contingency-based delay was associated with low levels of challenging behavior and acceptable levels of the FCR. These results indicate that contingency-based delays may be more effective in establishing tolerance for delays-to- reinforcement than time-based delays with access to alternative reinforcement.

Studies examining procedural variations of delay schedules are limited and have only examined two variations to the traditional model. Additionally, findings from the existing studies are mixed and do not provide clear evidence for the support of modified delay schedules. Given the dearth of research on this topic (Hagopian et al., 2005) and the social significance of delay schedules, research exploring additional antecedent procedures within delays-to-reinforcement is warranted.

Choice

Choice making is an action in which an individual selects between two or more available stimuli and is a commonly used antecedent intervention to reduce challenging behavior (Duker, Didden, & Sigafoos, 2004). In addition to decision-making, goal setting, and self-regulation, choice is a critical component of self-determination. According to Wehmeyer's functional theory, self-determined behavior occurs when an individual engages in intentional behaviors that result in a desired outcome (Wehmeyer & Abery, 2013; Wehmeyer, Kelchner, & Richards, 2006; Wehmeyer, 2005). The critical aspect of this theory is that the desired outcome is one that is

valued by the individual as opposed to an outside agent, which is often observed among individuals with IDD (Wehmeyer & Abery, 2013). Within the context of the functional theory of self-determined learning, choice making can be conceptualized as a self-regulating behavior that empowers the individual to have control over his or her environment (Wehmeyer & Abery, 2013). When an individual independently engages in a behavior without the influence of others (e.g., making a choice), they are able to contact novel reinforcers that may reduce frustration and increase engagement in preferred activities (Bambara, Koger, Katzer, & Davenport, 1995; Vaugn & Horner, 1995). Engagement in such choice behaviors have been associated with reductions in challenging behavior, which is why choice is commonly used as a component of challenging behavior interventions.

Interventions using choice involve providing an individual with a choice of activity, work task, or allowing the individual to choose the order in which they complete a series of tasks (Zellinsky & Shadish, 2018). Choice is a powerful intervention that has been shown to reduce challenging behavior among individuals with IDD in a number of studies (Carr & Carlson, 1993; Cole & Levinson, 2002; DeLeon, Neidert, Anders, & Rodriguez-Carter, 2001; Dyer, Dunlap, & Winterling, 1990; Elliott & Dillenburger, 2017; Graff et al., 1998; Harding, Wacker, Berg, Barretto, & Rankin, 2002; Kearney & McKnight, 1997; Kern & Clemens, 2007; Rispoli et al., 2013; Romaniuk & Miltenberger, 2001; Seybert, Dunlap, & Ferro, 1996; Shogren, Faggella-Luby, Bae, & Wehmeyer, 2004; Umbreit & Blair, 1996; Vaughn & Horner, 1995).

For example, Graff et al. (1998) evaluated the effects of different reinforcement conditions on challenging behavior for two individuals diagnosed with DDs. Three reinforcement conditions (i.e., researcher choice, participant choice, and constant choice) were evaluated using a multielement design. During each session, participants were asked to complete

an arbitrary free operant task (i.e., putting dominoes into a slot container) and reinforcement was delivered on a variable ratio schedule. During the researcher choice condition, the participant was given three of the same small edibles and was allowed to select one. The participant choice condition was the same as the researcher choice condition except that the participant was given a choice of three different small edibles. Results indicated that levels of challenging behavior were overall lower in the participant versus researcher selected conditions. The authors suggested that situations, where individuals are allowed to make choices, may be preferred over nonchoice situations and may explain the differences in levels of challenging behavior between conditions.

Ip, Szymanksi, Johnston-Rodriguez, and Karls (1994) evaluated the effects of a staff implemented daily choice problem on the frequency and severity of challenging behavior among adults living in a community residential facility using a quasi-experimental pre-test, post-test design. Staff in the choice group were asked to implement a daily choice program that consisted of embedding choice making opportunities into the individual's ongoing routines (e.g., choice of activity, choice of meals). Staff in the control group were asked to implement their typical routines without the choice program. Reductions in the frequency and severity of challenging behavior were observed across individuals in the experimental group. Conversely, increases in both the frequency and severity of challenging behavior were reported for individuals in the control group.

Choice making allows individuals with IDD to express preference, which gives the individual the skill to control his or her environment (Shogren et al., 2004; Wehmeyer, Agran, & Hughes, 1998). Several researchers have suggested that the skill to make choices or express preference may be more reinforcing than the stimuli that the individual selects (Romaniuk & Miltenberger, 2001). The skill to prevent challenging behavior, and teach self-determination

skills, makes choice making an ideal intervention for the reduction of challenging behavior (Algozzine, Browder, Karvonen, Test, & Wood, 2001). Although many studies have evaluated the effects of choice on challenging behavior, few have been conducted with individuals who have socially maintained challenging behavior. Additionally, the effects of choice have not been evaluated within the context of delays-to-reinforcement following FCT.

Gaps in the Literature and Study Rationale

Reducing intensity of reinforcement has been identified as a critical component to the long-term effectiveness and sustainability of FCT. Although multiple approaches for this exist, delay-to-reinforcement procedures reflect typical environments and are feasible for most caregivers to implement. However, due to the risk of extinction-induced challenging behavior, procedural variations to delay schedules are often required to achieve optimal outcomes. Few studies have examined variations of typical delay schedules, and the findings of those studies are mixed. Additionally, only two non-aversive procedural variations (i.e., alternate activities and contingency-based delays) have been evaluated. The effects of other antecedent strategies, such as choice, have yet to be evaluated. Therefore, the purposes of the current study was to (a) evaluate if the provision of alternative activities during delays-to-reinforcement resulted in low levels of challenging behavior, and (b) to compare differences in levels of challenging behavior and communication during periods when choice of alternative activity was available versus not available. Specifically, we asked the following two research questions:

a) What is the effect of providing alternative activities on challenging behavior during delays-to-reinforcement following functional communication training?
b) How does the provision of choice of alternative activities impact challenging behavior during delays-to-reinforcement?

Method

Participants

Three participants, hereafter referred to as consumers, were recruited for this study through a nonprofit organization serving individuals with IDD in the Midwestern region of the United States. The consumers who participated in this study were included based on the following eligibility criteria: (a) consumer was diagnosed with IDD, (b) consumer was over the age of 10 years, (c) consumer received services through a local day treatment or after-school center, (d) Motivation Assessment Scale (MAS; Durand & Crimmins, 1988) scores indicated a potential tangible function, (e) consumer engaged in mild challenging behavior, and (f) consumer had the skill to make choices. For this study, mild challenging behavior was defined as behavior that was disruptive to the environment or impacted the consumer's participation in typical activities but did not produce injury to themselves or others. To maintain the safety of the consumer and the experimenter, consumers who engaged in severe aggression or SIB did not qualify for participation in this study. For this study, choice making was defined as the skill to select from two or more available options (Duker et al., 2004).

Nadine.

Nadine was a 16-year-old Hispanic female diagnosed with ASD and ID. Nadine had complex communication needs and only emitted occasional non-word utterances. Her primary form of communication was gesturing in the forms of pulling and pointing, but no formal communication system was in place. At the time of the study, she attended the after-school program four days per week for 2 hrs. each day. Staff at the after-school program reported that Nadine’s challenging behavior was chronic, and prevented her from participating in group

activities with her peers such as craft and free play. Nadine was also assigned a 1:1 staff due to the chronicity and severity of her challenging behavior.

Reba.

Reba was a 41-year-old Caucasian female diagnosed with a profound ID, intermittent explosive disorder, and nine chronic health conditions. Reba was nonverbal and used gestures as her primary form of communication. At the time of the study, Reba attended the day program center five days per week for seven hours each day. Reba had a history of engaging in chronic challenging behavior and received monthly support from a behavior specialist to address her behaviors of concern. A previous FBA that included interviews and direct observations indicated that her challenging behaviors were maintained by access to tangible items and staff attention. Interventions to reduce the target behaviors were developed but were not implemented consistently by staff. Before the onset of the study, Reba's staff and behavior specialist reported that the behavioral interventions in place were not sufficient to reduce the intensity or frequency of her behaviors.

Bryce.

Bryce was a 26-year-old Hispanic male diagnosed with a mild ID, autism spectrum disorder, attention deficit hyperactivity disorder, impulse control, and chronic rhinitis. Bryce communicated verbally and spoke using full sentences. Bryce had a history of challenging behavior including verbal and physical aggression and property destruction that resulted in removal from his place of employment and a reduction in the number of hours he spent at the day program center. Additionally, due to the chronicity of his behavior, Bryce was no longer

allowed to attend community outings with center. At the time of the study, Bryce attended the day program center four days per week for three hours each day.

Setting and Materials

All assessment and intervention sessions took place at the day program center and after-school program. For Nadine, all sessions took place in classrooms within the after-school program. Classrooms contained a variety of leisure and craft materials that were freely available to consumers. During sessions between two and eight other consumers and after-school, staff were present but did not interact with the target consumer. For Reba and Bryce, all study sessions took place in their day program classroom. The classroom housed individuals over the age of 18 who were diagnosed with IDD and engaged in challenging behavior. The classroom contained a variety of leisure and educational materials (e.g. books, activity packets, etc.) that were freely available to consumers. During sessions between two and eight other consumers and direct service providers were present, but did not interact with the target consumers.

The first author, a doctoral student in Special Education and Board Certified Behavior Analyst (BCBA®) conducted all study sessions. Study materials included, (a) relevant reinforcer(s), (b) leisure materials/activities, (c) communication materials (i.e., picture exchange cards) and (b) paper and pencils and/or iPads for data collection purposes. Specific reinforcers, leisure materials, and communication materials were selected based on consumer need and preference. Nadine's preferred item was a Dora the Explorer® storybook. Reba's preferred item was a tablet device that she used to watch television shows. Bryce's preferred activity was talking to the experimenter. All preferred items and activities were selected based on the results of a preference assessment. A variety of preferred alternative materials were used during the delay-to-reinforcement sessions. For Nadine, alternative activities included a board game and

flash cards. For Reba, alternative items and activities included a stuffed Barney™ doll and listening to music. For Bryce, alternative items included puzzles and coloring sheets.

Response Measurement and Interobserver Agreement

Data were collected on three dependent variables including challenging behavior, FCRs, and item engagement. Challenging behavior was measured as percentage of opportunities per 5 min session. An opportunity was defined as the removal of the functional reinforcer. Approximately eight opportunities were presented during each 5 min session. Percentage of opportunities was calculated by dividing the number of opportunities with challenging behavior, by the total number of opportunities and multiplying by 100. Appropriate communication was measured in the same manner as challenging behavior. Data on the frequency of communicative responses was also collected during the delay-to-reinforcement. For item engagement, each 5 min delay-to-reinforcement session was partitioned into 30 10 s intervals, to obtain an estimate of item engagement during each session. The occurrence or nonoccurrence of item engagement was then recorded during each interval. The percentage of session intervals with item engagement was calculated by dividing the total number of intervals with item engagement by the total number of intervals and multiplying by 100.

All consumers engaged in multiple topographies of challenging behavior. Nadine's challenging behaviors included screaming, grabbing, dropping, and throwing. Screaming was defined as non-word vocalizations emitted above a typical speaking volume. Grabbing was defined as extending arm towards the experimenter or preferred item. Dropping was defined as abruptly moving from a standing to a seated position, or from a seated position to lying on the floor. Throwing was defined as the forceful release of an item from the hand to the floor.

Reba's challenging behaviors included screaming, grabbing, whining, and aggression. Definitions of screaming and grabbing were the same as those described for Nadine. Whining was defined as non-word high-pitched vocalizations. Aggression was defined as hitting or scratching the experimenter. Bryce's challenging behaviors included screaming, property destruction, making verbal threats, and aggression. Screaming was defined in the same manner as for Nadine and Reba. Property destruction was defined as throwing or breaking classroom property (e.g. puzzles, art supplies). Aggression was defined as forceful contact between the hand and classroom property or another staff member in the classroom. Verbal threats were defined as verbal statements indicating harm towards another person (e.g., "I'm going to beat you up").

A FCR was selected for each consumer based on their communicative skills and input from the directors of the day program center and after-school program. Nadine was taught to mand for access to her preferred tangible items using a picture exchange system. Independent FCRs were defined as the consumer placing a 3 × 3-inch picture card in the experimenter's hand with no prompting. Reba was taught to emit the FCR by pointing to a 5 × 3-inch card that depicted a picture of a music symbol. Pointing was selected as the mode of communication for Reba because it was the existing method used at the day program center. Bryce was taught to request attention by saying "Talk." For all consumers, item engagement was defined as using the activity or tangible item for its intended purpose (e.g., coloring with crayons).

The first author collected data on all three dependent variables during 100% of sessions across all phases and for each participant. Trained graduate and undergraduate students collected reliability data on a minimum of 50% of sessions across all study phases and for each participant. Disagreements between raters were settled by discussing the disagreement until consensus was

reached. For participants who consented to video recording, sessions were randomly selected and scored for IOA. Raters were trained in data collection procedures using sample videos. Training was considered complete when agreement between the first author and the raters reached a minimum of 80% across three consecutive training sessions.

Treatment Fidelity

Data on the experimenter's adherence to study procedures was collected on a minimum of 50% of baseline, treatment, and delay-to-reinforcement sessions. Fidelity was scored using an experimenter developed task analysis that listed the intervention steps in each phase. Each step was scored as "correct," "incorrect," or "not applicable." Fidelity was calculated by dividing the number of steps scored "correct" by the total number of steps and multiplying by 100. Task analyses for baseline, FCT, and schedule thinning (no choice, choice, and no activity) sessions can be found in Appendices C9, C10, and C11.

Pre-Intervention Procedures

Preference assessment.

A stimulus preference assessment was conducted with each consumer before the FBA to identify an array of preferred items (Mayer et al., 2014) to be used during the delay-to-reinforcement. The type of preference assessment used was selected based on consumer need. A multiple stimulus without replacement (MSWO; DeLeon & Iwata, 1996) preference assessment was conducted with each participant. If challenging behavior occurred during the assessment, a free operant (FO) preference assessment was conducted. During the MSWO, between five and seven items were placed in front of the consumer. The experimenter described each item and allowed the consumer to interact with the items for 60 s. At the beginning of the assessment, the

experimenter directed the consumer to “Choose one.” Contingent on selection of one of the items, the consumer was allowed to engage with the item for 30 s. After 30 s elapsed, the experimenter told the consumer, “Ok, my turn,” removed the item, and directed the consumer to “Choose one.” If at any point the consumer attempted to select more than one item, the experimenter blocked access and said, “Remember, choose just one.” The placement of each of the items rotated between trials. This process repeated until the consumer selected each of the items. The MSWO preference assessment was conducted a minimum of two times or until a pattern of responding was observed. Data for the MSWO were collected using the *MSWO data collection sheet* (Appendix C3).

During the FO preference assessment, the experimenter placed between 5-7 items within reach of the consumer. The experimenter observed each consumer for 10 min and recorded which items the consumer selected, and how long he or she interacted with each item. During the FO preference assessment, the experimenter did not interact with the consumer. The FO preference assessment was conducted a minimum of two times or until a pattern of responding was observed. Data were collected using the *FO preference assessment sheet* (Appendix C4).

Functional behavior assessment.

The FBA consisted of an indirect assessment, direct observations, and a FA of challenging behavior. First, the experimenter administered the MAS (Durand & Crimmins, 1988) to each consumer’s direct service provider or after school-staff member. The MAS is a 16-item rating scale designed to identify the environmental variables suspected of maintaining challenging behavior (Zarcone, Rodgers, Iwata, Rourke, & Dorsey, 1991). Following the administration of the MAS, experimenters conducted a minimum of 1 hr. of direct observation using ABC narrative recording (Bijou, Peterson, & Ault, 1968). The purpose of collecting ABC

data is to observe the individual during his or her typical routines and identify relations between the target behavior and antecedents and consequences (Cooper et al., 2007). Antecedents and consequences that have a high degree of correlation with the target challenging behavior can suggest a hypothesis about the function of the target behavior (Cooper et al., 2007). Each observation session lasted no longer than 30 min and occurred in the same setting across observations. During observations, experimenters recorded each instance of the target challenging behavior(s) and the immediate events preceding and following it. Data were collected using the *ABC Data Collection Sheet* (Appendix C1). The information obtained via the ABC recording was used to develop a hypothesis of behavioral function and the FA conditions.

Functional analysis.

Each consumer participated in a TBFA of challenging behavior (Sigafoos & Saggers, 1995) based on the procedures described by Rispoli, Davis, Goodwyn, & Camargo (2013). To reduce the duration of the assessment, only variables suspected of maintaining challenging behavior were tested in the TBFA. Thus, each consumer was only exposed to between one and four conditions. Trials consisted of two components, control, and test, and each trial lasted no longer than 2 min. During the control component, the consumer had access to the putative reinforcer for 60 s. The experimenter then removed the reinforcer for 60 s or until challenging behavior occurred. Contingent on challenging behavior, the experimenter reinstated access to the reinforcer and immediately ended the trial. During the TBFA, data were collected on consumer challenging behavior using the *TBFA Data Collection Sheet* (Appendix C2).

Results of Nadine's MAS and direct observations indicated that she frequently engaged in challenging behavior when access to her storybook was denied or delayed. Thus, Nadine participated in a single TBFA condition, contingent access to tangibles. Reba's MAS and direct

observations suggested multiple functions, including access to tangibles and attention. Reba was exposed to four TBFA conditions based on these findings, including contingent access to tangibles, contingent attention, synthesized attention and tangibles, and a modified attention condition. Bryce's MAS and direct observations suggested an attention function. A TBFA was conducted first for Bryce, but the results of the assessment were unclear, and an additional pairwise FA (PFA) was conducted.

During the control component of the tangible condition, the experimenter gave the consumer access to a preferred tangible item for 60 s. No attention or task demands were presented. At the beginning of the test component, the experimenter said, "My turn," and removed the preferred tangible. Contingent on challenging behavior, the experimenter returned the item and immediately ended the trial.

During the control component of the attention condition, the experimenter provided uninterrupted access to attention for 60 s, and the consumer had access to a neutral tangible item (e.g., book, coloring pages). During the test component, the experimenter said, "Ok, I need to go help someone else," and walked away from the consumer. Contingent on challenging behavior, the experimenter approached the consumer and provided brief verbal attention (e.g. "Are you ok?") and immediately ended the trial. If the consumer did not engage in challenging behavior, the experimenter withheld attention for 60 s.

During the control component of the synthesized condition, Reba had uninterrupted access to attention from the experimenter and the tablet device. During the test condition, the experimenter said, "Ok, I need to take this to someone else," removed the tablet and walked away from the consumer. Contingent on challenging behavior, the experimenter approached the consumer, reinstated access to the tablet, and provided brief attention (e.g. "Ok, you can have it

back now"). If no challenging behavior occurred, the experimenter withheld the tablet device and attention for 60 s.

During the control component of the modified attention condition, the experimenter sat next to the consumer for 60 s and provided continuous attention. The consumer also had access to her highly preferred tangible item. During the test component, the experimenter said, "Ok, I need to go talk to someone else," and walked away from the consumer. Contingent on challenging behavior, the experimenter approached the consumer, said, "Are you ok?", and immediately ended the trial. If no challenging behavior occurred, the experimenter withheld attention for 60 s.

The PFA consisted of two conditions, attention, based on the procedures described by Iwata et al. (1994) and control. Each session was 5 min in duration, and the order of sessions was randomly selected. At the beginning of the attention condition, the experimenter said, "I'll be right back," and walked away from Bryce. Contingent on challenging behavior, the experimenter approached Bryce and provided 30 s of attention. After 30 s elapsed, the experimenter removed her attention again. This cycle repeated until the entire 5 min session elapsed. During the control condition, Bryce had access to a moderately preferred item and continuous attention from the experimenter for 5 min.

Experimental design and data analysis.

The effects of FCT and delay-to-reinforcement were evaluated using an ABAC design where A represented baseline, B represented FCT plus extinction, and C represented delay-to-reinforcement (Kennedy, 2005; Hagopian et al., 2005). A multielement design was embedded into the C phase to evaluate differences between the delay-to-reinforcement conditions (i.e., no choice, choice, and no activity). To reduce the likelihood of carryover effects, the order of

conditions was randomly selected before each data collection session (Kennedy, 2005). Results of the analyses were evaluated using visual analysis of level, trend, and variability of data within each phase (Kennedy, 2005; Maggin et al., 2013).

Baseline and Intervention Procedures

The study consisted of four phases: (a) baseline, (b) communication training, (c) FCT plus extinction, and (d) delay-to-reinforcement. All study sessions were five minutes in duration, and no more than six sessions were conducted per visit. A session termination criteria was determined for each consumer before baseline to ensure the safety of the consumer and experimenter. Termination criteria was determined with input from each consumers' direct service provider or case manager. Sessions were immediately terminated if the consumer reached the predetermined criteria.

Baseline.

Sessions began with the experimenter giving the consumer access to the putative reinforcer for 30 s. The experimenter then said, "Ok, it's time to do something else," and removed the reinforcer, or said, "I need to go help someone else," and walked away from the consumer. Contingent on challenging behavior, the experimenter reintroduced the reinforcer for 30 s. The experimenter ignored any emissions of the FCR. During the return to baseline, three probes were conducted to assess for resurgence of challenging behavior.

Communication training.

The purpose of the communication training phase was to establish the FCR in the consumer's repertoire. The entire communication phase consisted of three phases: (a) initial training, (b) training with a 5 s time delay, and (c) training with a 10 s time delay. Each

communication training session consisted of three trials, each lasting no longer than 60 s. Each phase was considered complete when the participant emitted the FCR correctly in 100% of trials across three consecutive sessions. Phase 1 consisted of a most-to-least physical or verbal prompting hierarchy. The physical prompting hierarchy consisted of full physical, partial physical, and gesture prompts. The verbal prompting hierarchy consisted of full verbal, partial verbal, and initial phoneme prompts. Trials began with the experimenter giving the participant access to the relevant reinforcer for 10 s. Then, the experimenter removed the reinforcer and immediately provided the relevant prompt (i.e., full physical, partial verbal, etc.). Contingent on appropriate communication, the experimenter gave the participant access to the reinforcer for 30 s. Procedures for Phase 2 were the same as Phase 1 except that after the experimenter removed the reinforcer, she inserted a 5 s delay. If the consumer responded incorrectly or did not respond within 5 s, the experimenter provided a full physical or verbal prompt. Procedures for Phase 3 were the same as Phase 2 except that the time delay increased to 10 s. During all phases of communication training, challenging behavior was placed on extinction.

Functional communication training plus extinction.

During treatment sessions, the experimenter sat or stood within 3 ft. of the consumer. Sessions began by giving the participant 30 s access to the putative reinforcer. The experimenter then said, “Ok, it’s time to do something different,” and removed the reinforcer or said, “I need to go help someone else,” and walked away from the consumer. Contingent on the emission of the FCR, the experimenter provided brief descriptive praise (i.e. “Thanks for asking”) and access to the reinforcer for 30 s. If the consumer did not emit the FCR within 5 s, the experimenter used a system of least-to-most (LTM) prompts to evoke appropriate responding. The experimenter ignored all instances of challenging behavior.

Delay-to-reinforcement.

During all delay-to-reinforcement sessions, the delay- to-reinforcement gradually increased until the terminal delay of 3 min was reached. The initial delay was 0 s and increased in the following manner: 5 s, 8 s, 15 s, 30 s, 60 s, 90 s, 120 s, 150 s, and 180 s (Fuhrman et al., 2018). Criteria for increasing the delay was two consecutive sessions with a 85% or greater reduction in challenging behavior relative to baseline (Hagopian et al., 1998). If challenging behavior was below the 85% reduction for three consecutive sessions (Hagopian et al., 1998), the delay returned to the previous interval. Towards the end of the study, Reba was diagnosed with Pneumonia and data collection had to be terminated. Thus, the terminal delay for Reba was modified to 120 s.

At the start of each delay-to-reinforcement session, a different colored tray was placed on the table in front of the consumer to signal the beginning of the condition. A red tray was used to signal the choice condition and green and black trays were used to signal the no choice and no activity conditions respectively. For Bryce, laminated colored sheets of paper were instead of trays to prevent Bryce from throwing the trays at staff or other consumers. Additionally, a timer was set at the beginning of the delay that sounded when the delay elapsed.

No choice. Sessions began with the experimenter giving the consumer 30 s access to the putative reinforcer. The experimenter then said, “Ok, it’s time to do something different,” and removed the reinforcer, or said, “I need to go help someone else,” and walked away from the consumer. Contingent on the emission of the FCR, the experimenter said, “Good asking, but you have to wait for ___ seconds/minutes. While you wait, you can use ___ (name of item/activity),” and offered the consumer an alternative item/activity identified via the preference assessment. During the delay, the experimenter ignored all instances of challenging behavior and any

emissions of the FCR. After the predetermined delay interval elapsed, the experimenter said, "Thanks for waiting. If you want ___ now you can ask," and reintroduced the putative reinforcer for 30 s. If challenging behavior persisted after the delay interval elapsed, the experimenter waited for a 5 s break in challenging behavior and then prompted communication using a system of LTM prompts. The alternate activity used during each no choice session was randomly selected before the start of each session.

Choice.

Choice sessions were conducted in the same manner as the No Choice sessions except that after the experimenter removed the reinforcer, she offered the consumer a choice of two activities identified via the preference assessment. If the consumer did not select between the alternative items within 10 s, the experimenter said, "Remember, you can choose one while you wait," and re-presented both items.

No activity.

During the no activity sessions, the consumer was told to wait for access to the functional reinforcer but was given no alternative activity or item. All instances of challenging behavior and appropriate communication were ignored.

Results

Ineterobserver Agreement and Treatment Fidelity

Interobserver agreement and treatment fidelity data are presented in Tables A10 and A11 respectively.

Trial-Based Functional Analysis

Results of the FAs for all participants are displayed in Figure B1.

Nadine.

High and stable levels of challenging behavior were observed across all trials of the test component of the tangible condition. During the control component, low and stable levels of challenging behavior were observed across trials, with no overlap with the test component. These findings suggest that Nadine's challenging behavior was maintained by access to preferred tangible items.

Reba.

During the test component of the tangible condition, Reba engaged in high and stable levels of challenging behavior. During the control component, no instances of challenging behavior were observed; suggesting that Reba's challenging behavior was maintained by access to the tablet. During the control component of the synthesized condition, low and stable levels of challenging behavior were observed. During the test component, high and stable levels of challenging behavior were observed, indicating that challenging behavior was maintained by both access to the tablet device and experimenter attention. When attention was evaluated in isolation, variable levels of challenging behavior were observed across both the control and test components. During the first three trials, no challenging behavior was observed during the control component. However, increases in challenging behavior were observed during the final two trials. During the test component, high levels of challenging behavior were observed in all but one trial. To clarify the results of the attention condition, a modified attention condition was conducted. Across all trials, no instances of challenging behavior were observed in either

component, suggesting that challenging behavior was maintained only by access to tangible items.

Bryce.

During the control sessions of the PFA, low levels of challenging behavior were observed, while moderate but stable levels of challenging behavior were observed during the attention sessions. A slight increase in challenging behavior during the control sessions was observed during sessions six and seven but returned to near zero levels during the remainder of the sessions. Minimal overlap was observed between conditions, with increasing differentiation observed across the last five sessions. These results suggest that Bryce's challenging behavior was maintained by social positive reinforcement, in the form of attention.

Treatment Evaluation

Nadine

Challenging behavior.

Results of the FCT and delay-to-reinforcement evaluations are displayed in Figure B2 for Nadine. Data are also presented individually for each delay-to-reinforcement condition and are displayed in Figure B3. Data on instances of resurgence across delay-to-reinforcement sessions are displayed in Table A12.

During baseline sessions, Nadine engaged in high and stable levels of challenging behavior (*M*= 93% of opportunities) across sessions. The introduction of FCT resulted in immediate decreases in challenging behavior (*M*= 8% of opportunities). The return to baseline resulted in immediate increases in challenging behavior (100% of opportunities), to similar levels observed during the initial baseline. When treatment extension began, and the delay-to-

reinforcement was between zero and 5 s, challenging behavior returned to near zero levels. Additionally, considerable overlap was observed, with 84% of data points overlapping between conditions.

Differentiation between conditions began to emerge with the implementation of the 8 s delay. Levels of challenging behavior remained stable and low in the choice condition until the terminal delay of 180 s was reached. Beginning with the 8 s delay, there was also a considerable decrease in overlap between the choice, no choice, and no activity conditions, from 84% to 65%. For the no choice and no activity conditions, increases in the duration of the delay were associated with variable increases in challenging behavior.

During the no choice condition, challenging behavior exceeded initial treatment levels during seven sessions, with challenging behavior returning to baseline levels during one session (session 73). Similar patterns of responding were observed during the no activity condition, with challenging behavior exceeding initial treatment levels during eight sessions. Moderate and variable levels of challenging behavior were observed across both conditions until the final four no choice and final three no activity sessions, at which point challenging behavior decreased to near zero levels.

A total of 26, 25, and 22 sessions were conducted across the choice, no choice, and no activity conditions respectively. Within 26 choice sessions, the terminal delay of 180 s was reached, providing evidence for the efficiency of activity choice in facilitating tolerance for delays-to-reinforcement. Conversely, by session 25 of the no choice condition, the delay-to-reinforcement had only reached 60 s. By the final no activity session (session 22), the duration of the delay had only increased to 30 s.

Communication.

During the initial baseline phase, Nadine engaged in no instances of appropriate communication. The introduction of FCT resulted in immediate and stable increases in appropriate communication (100% of opportunities). A brief return to baseline resulted in sharp decreases in the FCR (0% of opportunities). Table A13 displays the average number of FCRs emitted during delays-to-reinforcement for Nadine. Emissions of the FCR during the choice delays were observed infrequently. Only one (average .50 across sessions) emission of the FCR was observed during one 5 s delay session. Similarly, only two emissions (average 1 across sessions) of the FCR were observed during the 5 s delay. Emissions of the FCR were observed at elevated rates across all delay-to-reinforcement sessions during the no activity condition. The average number of emissions of the FCR ranged from .40 to five across all no activity sessions.

Item engagement.

Data on item engagement were collected during the choice and no choice conditions and are displayed in Table A14. Similar levels of item engagement were observed between conditions. However, slightly higher levels of item engagement were observed overall during the choice condition.

Reba

Challenging behavior.

Data for the FCT and delay-to-reinforcement evaluations are displayed in Figure B2 for Reba. Data are also presented individually for each delay-to-reinforcement condition and are displayed in Figure B4. Data on instances of resurgence across delay-to-reinforcement sessions are displayed in Table A15.

During the first baseline phase, Reba engaged in high and stable levels of challenging behavior (*M*= 87% of opportunities). The provision of FCT was associated with sharp and immediate decreases in challenging behavior (*M*= 12% of opportunities). Following a brief return to baseline, challenging behavior increased to higher and more stable levels than observed during the initial baseline phase. Moderate to high and variable levels of challenging behavior were observed during the first 30 sessions of the delay evaluation. During the initial stages of the evaluation, challenging behavior returned to baseline levels for at least one session during each condition. Additionally, the duration of the delay decreased multiple times within each condition, before stable increases in the delay interval could be made. Similar to Nadine, there was also considerable overlap (95%) between conditions during the early stages of the delay.

Differentiation between conditions began to emerge during session 47 when the duration of the delay during the choice condition was increased to 8 s. Beginning with the 8 s delay, low and stable levels of challenging behavior were observed during the choice condition until the terminal delay of 120 s was reached. A 38% decrease in overlap was also observed between the choice and other conditions, from 95% to 57%. Conversely, moderate and variable levels of challenging behavior persisted during the no choice and no activity conditions. For both the no choice and no activity conditions, challenging behavior exceeded initial treatment levels on six occasions. Stable decreases in challenging behavior were observed during the final four sessions of the no choice condition, with challenging behavior reaching near zero levels. Stable decreases in responding were observed during sessions 21-28 of the no activity condition. A slight increase in challenging behavior was observed during session 29 but decreased during the final three sessions.

A total of 37 choice 32 no choice, and 39 no activity sessions were conducted during the delay-to-reinforcement phase. Like Nadine, the terminal delay was only reached during the choice condition. By session 37 of the choice condition, the terminal delay of 120 s was reached. By session 32 of the no choice condition and session 39 of the no activity condition, a delay interval of 30 s was reached.

Communication.

During the initial baseline phase of the FCT evaluation, Reba engaged in no instances of the FCR. The application of FCT was associated with immediate and stable increases in appropriate communication (*M*= 100% of opportunities). The removal of the FCT intervention resulted in immediate decreases in appropriate communication (0% of opportunities). Data on the frequency of appropriate communication during the delay-to-reinforcement are displayed in Table A16. Patterns of responding for Reba were similar to those observed for Nadine. Overall, the number of appropriate communicative responses emitted during the delay was highest during the no activity condition. Fewer instances of the FCR were emitted during the no choice conditions compared to the no activity condition but were on average higher than the choice condition. With the exception of the 120 s delay, emissions of the FCR were lowest during the choice condition.

Item engagement.

Data on item engagement for Reba are displayed in Table A17. Considerably higher levels of item engagement were observed during the choice condition compared to the no choice condition across all delay intervals.

Bryce

Challenging behavior.

Data on Bryce's treatment evaluation are displayed in Figure B2. Data are presented for each delay-to-reinforcement condition individually in Figure B5. Data on resurgence of challenging behavior across conditions are displayed in Table A18. During baseline, Bryce engaged in high and stable levels of challenging behavior (*M*= 88% of opportunities). FCT was associated with immediate and stable decreases in challenging behavior (*M*= 0% of opportunities). During the return to baseline, challenging behavior returned at moderate but stable levels (*M*= 52% of opportunities). Stable and low levels of challenging behavior were observed with the introduction of the delay-to-reinforcement. Across all conditions, challenging behavior remained at low levels during the initial delay-to-reinforcement sessions (0 s, 5 s, 8 s, 15 s).

Communication.

During baseline, Bryce engaged in no instances of appropriate communication. FCT was associated with immediate and stable increases in appropriate communication (*M*= 100%). The removal of FCT corresponded with sharp decreases in appropriate communication (*M*= 8 percent of opportunities). Instances of appropriate communication were observed infrequently during the delay-to-reinforcement, and are displayed in Table A19. Only one instance of appropriate communication was observed during the delay during the choice and no choice conditions.

Item engagement.

Data on the percentage of 10 s intervals with item engagement across delay-to-reinforcement conditions are displayed in Table A20. Consistent with the results obtained for

Nadine and Reba, item engagement for Bryce was higher in the choice condition compared to the no choice condition.

Discussion

The purpose of the present study was to (a) examine the impact of alternate activities on challenging behavior during delays-to-reinforcement and (b) evaluate the effect of activity choice on challenging behavior during delays-to-reinforcement following FCT. The current study extends the research on delay-to-reinforcement and choice in several ways. First, relatively few studies have systematically evaluated methods for improving the efficacy of delay-to-reinforcement procedures, even though research has shown that this procedure often leads to resurgence of challenging behavior. The studies that have addressed this topic have primarily focused on the impact of alternate activities alone. The current study provides an additional demonstration of the efficacy of modified delay-to-reinforcement procedures and is the first to examine the impact of activity choice as a procedural modification.

The current study also extends the literature on choice as a treatment for challenging behavior. Most research on activity choice has been conducted with individuals who engage in challenging behavior maintained by negative reinforcement. The current study provided a novel demonstration of the use of choice to treat challenging behavior for individuals whose challenging behavior was maintained by social positive reinforcement. Results of this study showed that activity choice can be effective in maintaining low levels of challenging behavior during delays-to-reinforcement following FCT. Major findings, limitations, and directions for future research are described below.

Major Findings

Results of the current study indicate that the provision of alternate activities alone may not be sufficient to reduce challenging behavior during periods when individuals have to wait for access to a functional reinforcer. For the participant's included in this study, the application of alternative activities alone was not sufficient to reduce challenging behavior as the delay-to-reinforcement increased. Overall, higher and more variable levels of challenging behavior were observed during the no choice condition compared to the choice condition. For all participants, challenging behavior returned to baseline levels during at least one no choice session. This finding aligns with previous studies that have also demonstrated that treatment components, in addition to alternate activities, may be necessary to prevent resurgence of challenging behavior during delays-to-reinforcement (Ghaemmaghami et al., 2016).

For the participants included in this study, activity choice produced low and relatively stable levels of challenging behavior. For Nadine, no instances of resurgence were observed during the choice condition. With the exception of two sessions for Reba, challenging behavior never returned to baseline levels. This finding is significant given that delays-to-reinforcement often result in substantial increases in challenging behavior (Fisher et al., 2000; Hagopian et al., 1998; Hagopian et al., 2005; Hagopian et al., 2011; Hanley et al., 2001). There are several potential reasons for this finding.

First, choice may have functioned as an abolishing operation (AO) for challenging behavior (Michael, 1993). Although the mechanisms underlying the effectiveness of choice have yet to be conclusively determined (Rispoli et al., 2013), previous research has suggested that the act of making a choice may be more reinforcing than the stimuli the individual selects from (Graff et al., 1998). Thus in the current study, the opportunity to make a choice may have altered

the reinforcing value of the highly preferred stimuli, which resulted in a decrease in challenging behavior (Laraway, Snycerski, Michael, & Poling, 2001).

An alternative explanation may be found by examining the relationship between preference, item engagement, and challenging behavior. Choice allows individuals to express preference (Shogren et al., 2004; Wehmeyer et al., 1998). For the participants included in this study, levels of item engagement were higher during the activity choice condition compared to the no choice condition. This finding suggests that preference may influence levels of item engagement. The results of this study also indicate that there may be an association between item engagement and challenging behavior. Across participants, higher levels of item engagement were associated with lower levels of challenging behavior, which may suggest that item engagement has an abative effect on challenging behavior (Michael, 1993). However, additional research is needed to confirm this hypothesis.

The current study also demonstrated that activity choice can facilitate tolerance for delays-to-reinforcement more quickly than no choice or no activity conditions. For all participants, the terminal delay was only reached during the choice condition. Unlike the no choice and no activity conditions, the duration of the delay had to be decreased less frequently during the choice condition. This resulted in the terminal delay being reached more rapidly during this condition.

Although similar patterns of responding emerged across participants, some variation in responding was observed during the initial stages of the delay-to-reinforcement phase. For Nadine and Bryce, levels of challenging behavior remained low when the delay-to-reinforcement was initially implemented. As the duration of the delay increased, changes in patterns of responding began to occur. For Reba, the introduction of the delay-to-reinforcement procedure

resulted in immediate increases in challenging behavior. There are several potential explanations for this finding. First, Reba's staff reported that except for eating meals and occasional community outings, Reba had continuous access to her tablet device for the entire seven hours she spent at the day program center. Thus, this study marked the first time that Reba was expected to tolerate periods without access to the tablet. Given her history of continuous access to reinforcement, it is possible that the initial delay interval of 5 s was too long. It may be that more gradual increases in the delay (e.g., 0 s, 1 s, 3 s, 5 s) would have resulted in lower levels of challenging behavior. Reba was also the only participant diagnosed with a profound ID. Research has shown that individuals with more significant disabilities have greater difficulty tolerating delays-to-reinforcement (Cuskelly, Gilmore, & Jobling, 2016; Cuskelly, Zhang, & Hayes, 2003). However, there is limited empirical evidence demonstrating how individual participant characteristics, including level of ID, influence responding during delays-to-reinforcement (LeBlanc et al., 2002). Such questions were beyond the scope of this study but should be addressed in future research.

Another noteworthy finding of this study is that sharp decreases in challenging behavior began to emerge during the no choice and no activity conditions the longer the delay-to-reinforcement was in effect. For all participants, near zero levels of challenging behavior were observed during the final 10-15 no choice and no activity sessions. Similar findings have been noted in previous studies that have examined treatment persistence following DRA (Mace et al., 2010; Wacker, Berg, & Harding, 2004; Wacker et al., 2011). For example, Wacker et al. (2011) conducted a series of studies examining the long-term effects of FCT in home settings for young children diagnosed with IDD. A series of treatment challenges were embedded into FCT sessions, to evaluate the persistence of challenging behavior. When extinction challenges were

initially incorporated into treatment, challenging behavior returned to baseline levels. Over time, levels of challenging behavior decreased during both brief and extended periods of extinction.

Nevin and Wacker (2013) suggested that exposure to extinction over time weakens challenging behavior, making it less resistant to periods of extinction. The more exposure the response (i.e., challenging behavior) has to extinction, the higher the likelihood that the response will weaken (Nevin & Wacker, 2013; Wacker et al., 2004; Wacker et al., 2011; Wacker et al., 2017). This explanation may help explain why resurgence of challenging behavior is more likely to occur during initial treatment challenges, or delays-to-reinforcement (Wacker et al., 2017). Our findings support this explanation. For all participants included in this study, instances of resurgence occurred during the early delay-to-reinforcement sessions. Consistent decreases in challenging behavior began to emerge the longer the delay-to-reinforcement was in place. Thus, the findings of this study provide additional demonstrations for how repeated exposure to extinction can increase treatment persistence over time.

Decreases in challenging behavior during the no choice and no activity conditions may also be a result of exposure to variable schedules of reinforcement over time. Variable schedules of reinforcement typically result in more consistent levels of responding, and lower levels of responding over time (Cooper et al., 2007). Thus, the decreases in challenging behavior observed during the final sessions of the no choice and no activity conditions may have been a result of the schedule of reinforcement in place. However, additional replications of the procedures described in this study are warranted to further elucidate the mechanisms affecting responding.

The current study also demonstrated that, in some cases, activity choice can prevent excessive emissions of the FCR during delays-to-reinforcement. For most participants, the FCR was emitted less frequently during the activity choice condition compared to the no choice and

no activity conditions. This is an important finding given that when emitted at high rates, the FCR can become problematic in applied settings, and prevent the individual from participating in other leisure or functional activities (Hanley et al., 2001; LeBlanc et al., 2002).

Limitations and Directions for Future Research

The current study has several limitations that should be noted. First, only three individuals participated in this study which limits the generalizability of these findings to other individuals with DDs. Additionally, only one of the participants was classified as an adolescent (i.e., ages 10-17 years). Additional replications of the procedures described in this study are warranted before conclusive statements regarding the efficacy of activity choice on challenging behavior can be made for this population.

Second, similar levels of item engagement were observed across the choice and no choice conditions for some participants. Item engagement was measured using 10 s partial interval recording in order to obtain an estimate of how often engagement was occurring. However, partial interval recording may have overestimated the occurrence of item engagement during the no choice condition (Kennedy, 2005). Future research should examine the use of different measurement systems, including whole interval recording or duration recording, for a more sensitive measure of item engagement. Third, two of the three participants included in this study engaged in tangibly maintained challenging behavior. Additional examinations of these study procedures should be conducted with individuals who engage in attention-maintained challenging behavior.

Fourth, social validity data were not collected. Delay-to-reinforcement is often described as a feasible approach for decreasing reinforcement intensity because it reflects typical environments most closely (Hagopian et al., 2011; Fisher et al., 2000). However, few studies,

including the current study, have collected social validity data following implementation of this procedure. Thus, it is unclear if the procedures used in this study would be considered acceptable or feasible by the staff working at the day program and after-school centers. Although staff at both centers were given the materials to continue implementing these procedures after the conclusion of the study, data on staff adherence and long-term use of these procedures were not collected. Thus, it is unclear if these procedures made long-term improvements in the participant's behavior.

Fifth, although the order of sessions was randomly selected to reduce the likelihood of carryover effects, it possible that carryover across conditions occurred. Although carryover effects are a risk when implementing a multielement or alternating treatment design (Kennedy, 2005), future research may consider utilizing other single-case experimental designs to prevent the possibility of carryover across conditions. Sixth, although IOA scores were overall above minimum thresholds for all participants, agreement scores during select sessions were well below acceptable levels for Bryce. There are two potential explanations for these low agreement scores. First, Bryce's conversational speaking volume was higher than typical levels. Thus, during some sessions it was difficult to differentiate whether Bryce was screaming, or speaking in a typical conversational volume. Second, all agreement data for Bryce were collected via video. One limitation of collecting data via video is that the ability to collect accurate data is influenced by the quality of the video. Thus, it is possible that during some sessions, insufficient video quality impacted the raters data.

Although they were selected based on the results of a stimulus preference assessment, the alternate activities used for Reba and Bryce were not age appropriate. If used in community settings, engagement with most of the alternate items (i.e., Barney™ stuffed doll, coloring

books) could be socially stigmatizing. As part of their Medicaid-funded day program, Reba and Bryce both had treatment plans designed to teach various age-appropriate leisure and functional life skills. For example, Reba periodically participated in a dance exercise program at the center. Engagement in activities such as the one described above would be more socially acceptable and could be translated to community-based settings. Future research should consider examining the effects of such activities on challenging behavior during delays-to-reinforcement.

Finally, although FCT is a highly effective intervention for the treatment of challenging behavior, it is an intensive intervention that requires training and resources to be implemented with fidelity. Thus, future research should explore how low intensity interventions affect levels of challenging behavior in day program and after school settings. It is possible that the application of low or lower intensity interventions would eliminate the need for more intensive interventions, such as FCT. The participants included in this study engaged in few leisure activities, essentially creating an impoverished environment, which may have contributed to the elevated levels of challenging behavior observed during baseline. Thus, it is possible that teaching individual's novel leisure skills would create more enriched environments thereby reducing the need for more intensive interventions like FCT.

Implications for Practice

Despite the efficacy of many reinforcement-based interventions, reinforcement is often delivered at such high intensity, that maintaining treatment outcomes in applied settings is particularly challenging for natural change agents. When the intensity of reinforcement is decreased too quickly or in an unsystematic manner, challenging behavior is likely to resurge. Unfortunately, in many cases, increases in challenging behavior result in abandonment of empirically-supported behavioral interventions (LeBlanc et al., 2002). The results of the current

study provide initial evidence for an efficient and practical approach for reducing the intensity of reinforcement following intervention that can maintain low levels of challenging behavior. For all participant's, embedding activity choice into a delay-to-reinforcement procedure resulted in the maintenance of treatment effects. Additionally, all of the items used during the choice evaluation were found in the participant's typical placements and required no additional outside resources. Thus, these procedures could likely be adopted in applied settings (e.g., day programs, after- school centers, ABA clinics) with minimal changes to existing routines or protocols.

The current study also has important implications for facilities serving adults with IDD. Adults with developmental disabilities are given few opportunities to make choices during their daily routines (O'Donovan, Byrne, McCallian, & McCarron, 2017). The current study demonstrated a simple method for embedding choice into intervention for this population. Additionally, this study demonstrated that adding choice to treatment can result in robust decreases in challenging behavior. Staff working in the adult service field may consider including choice as an intervention component for the reduction of challenging behavior. However, given the limited external validity of this study, staff should collect ongoing data on consumer behavior to inform decisions regarding treatment.

The current study demonstrated the efficacy of reinforcement-based interventions for adults with IDD who have a history of engaging in chronic challenging behavior. As previous research has indicated, there is a lack of research on the use of evidence-based behavioral interventions for individual's 18-years and older (Sullivan, 2007). Additionally, many centers serving adults with IDD rely on ineffective treatments to manage behavioral challenges (Mills & Rose, 2010). This study provides additional evidence that high quality, empirically supported

interventions can be successfully implemented within day program centers and lead to meaningful improvements in consumer behavior.

CHAPTER IV: GENERAL DISCUSSION

The purpose of the current dissertation was to systematically evaluate procedures for decreasing reinforcement intensity following intervention using schedule thinning and delay-to-reinforcement. In study 1, a systematic review was conducted to (a) summarize the extant literature on schedule thinning following reinforcement-based intervention, and (b) determine the strength of the evidence for schedule thinning. In study 2, a single-case experimental design was used to evaluate a novel approach for promoting tolerance for delays-to-reinforcement following FCT. Three individuals with IDD were exposed to a treatment evaluation that consisted of FCT and delay-to-reinforcement. Three conditions, activity choice, no choice, and no activity were embedded into the delay-to-reinforcement phase to evaluate the impact of each condition challenging behavior. Both studies found positive effects for schedule thinning and delay-to-reinforcement and addressed several important gaps in the available literature.

To date, there has been no comprehensive systematic review of schedule thinning. Previous reviews have focused on specific reinforcement-based interventions (i.e., FCT; Hagopian et al., 2011), or focused on guidelines for reducing reinforcement intensity for practitioners (LeBlanc et al., 2002). Given that schedule thinning is a recommended treatment component, there was a need for a comprehensive summary of these procedures. Results of Study 1 indicate that schedule thinning is an effective method for maintaining treatment effects following FCT, NCR, and DRO. The effects of each experiment included in the review were classified using newly developed standards designed specifically to evaluate the effects of schedule thinning.

The standards developed for this review are unique in that they were designed to measure treatment efficacy at different points during intervention. Thus, these standards allow for a

detailed analysis of intervention effects over time and can assist researchers and practitioners in determining at what point during schedule thinning supplemental intervention is needed. Across interventions and schedule thinning approaches, schedule thinning was found to be effective in maintaining treatment effects over time. This finding is promising and suggests that the effects of evidence-based behavioral interventions can persist over time.

Results of Study 2 showed that embedding choice-making opportunities into delays-to-reinforcement can significantly reduce challenging behavior, and can facilitate tolerance for periods of non-reinforcement more rapidly than no choice or no activity conditions. Additionally, results of Study 2 showed that incorporating choice making into a delay-to-reinforcement procedure can significantly improve the efficiency of the procedure. For all three participants, the terminal delay interval was only reached during the choice making condition. For all other conditions, the final delay interval reached was at least 30 s less than the predetermined terminal delay. These findings are significant given that previous research has shown that delay-to-reinforcement often results in resurgence of challenging behavior (Fisher et al., 2000; Hanley et al., 2001; Ghaemmaghami et al., 2016). Taken together, the findings of these studies demonstrate how treatment effects can maintain as reinforcement intensity is decreased.

Although the results of the current studies provide promising evidence for the use of schedule thinning and delay-to-reinforcement, additional research on these topics is warranted. Specifically, there is a need for additional research on the use of schedule thinning and delay-to-reinforcement for individuals whose challenging behavior is maintained by automatic and negative reinforcement. Only a small percentage of studies included in Study 1, examined treatment of challenging behavior for such functions. Additionally, although the procedures described in Study 2 demonstrate a practical approach to delay fading, research is needed to

examine methods for training natural change agents in these procedures. Such research would provide important information on the sustainability and social validity of the procedures described in this study.

The findings of the current studies have important implications for practitioners and natural change agents who support individuals with DDs. Schedule thinning and delay-to-reinforcement are often recommended following reinforcement-based intervention (Hagopian et al., 2011; LeBlanc et al., 2002). However, these recommendations are often made based solely on professional conjecture (LeBlanc et al., 2002) and narrow syntheses of the schedule thinning literature. The results of Study 1 demonstrated the efficacy of schedule thinning across intervention categories, age, and disability groups, which may give practitioners greater confidence when recommending these practices to natural change agents.

Delay-to-reinforcement is often described as the most practical approach for reducing reinforcement intensity after intervention (Hagopian et al., 2011). Unfortunately, much of the literature on delay-to-reinforcement has shown that resurgence of challenging behavior often occurs as the duration of the delay increases, which often results in treatment abandonment (LeBlanc et al., 2002). Study 2 demonstrated the efficacy of a practical approach for modifying a traditional delay-to-reinforcement procedure that can be implemented in naturalistic settings. Offering participants a choice of alternative stimuli to use during the delay resulted in maintenance of treatment effects across all participants. All of the items used during the delay were selected based on consumer preference and were readily available in the participant's typical environment; which suggest that this modified delay-to-reinforcement procedure can be embedded into ongoing routines with minimal outside materials or resources.

REFERENCES

Algozzine, B., Browder, D., Karvonen, M., Test, D. W., & Wood, W. M. (2001). Effects of interventions to promote self-determination for individuals with disabilities. *Review of Educational Research, 71(2),* 219-277. doi:10.3102/00346543071002219

Austin, J. E., & Tiger, J. H. (2015). Providing alternative reinforcers to facilitate tolerance to delayed reinforcement following functional communication training. *Journal of Applied Behavior Analysis, 48(3),* 663-668. doi:10.1002/jaba.215

Bambara, L. M., Koger, F., Katzer, T., & Davenport, T. A. (1995). Embedding choice in the context of daily routines: An experimental case study. Journal of the Association for *Persons with Severe Handicaps, 20*(3), 185–195. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1996-23375-001&site=ehost-live

Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., Ogedegbe, G., Orwig, D., Ernst, D., & Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology, 23*, 443-451.

Bergstrom, R., Tarbox, J., & Gutshall, K. A. (2011). Behavioral intervention for domestic pet mistreatment in a young child with autism. *Research in Autism Spectrum Disorders, 5*(1), 218-221.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1*(2), 175-191. doi:10.1901/jaba.1968.1-175

Bowring, D. L., Totsika, V., Hastings, R. P., Toogood, S., & Griffith, G. M. (2017). Challenging behaviours in adults with an intellectual disability: A total population study and exploration of risk indices. *British Journal of Clinical Psychology, 56*(1), 16-32. doi:10.1111/bjc.12118

Briggs, A. M., Fisher, W. W., Greer, B. D., & Kimball, R. T. (2018). Prevalence of resurgence of destructive behavior when thinning reinforcement schedules during functional communication training. *Journal of Applied Behavior Analysis, 51(3),* 620-633. doi:10.1002/jaba.472

Britton, L. N., Carr, J. E., Kellum, K. K., Dozier, C. L., & Weil, T. M. (2000). A variation of noncontingent reinforcement in the treatment of aberrant behavior. *Research in Developmental Disabilities, 21*(6), 425–435. https://doi.org/10.1016/S0891-4222(00)00056-1

Byiers, B. J., Dimian, A., & Symons, F. J. (2014). Functional communication training in Rett syndrome: A preliminary study. *American Journal on Intellectual and Developmental Disabilities, 119*(4), 340-350. doi:10.1352/1944-7558-119.4.34

Carr, E. G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84*(4), 800-816. doi:10.1037/0033-2909.84.4.800

Carr, J. E., Coriaty, S., Wilder, D. A., Gaunt, B. T., Dozier, C. L., Britton, L. N.,... Reed, C. L. (2000). A Review of "Noncontingent" Reinforcement as treatment for the aberrant behavior of individuals with developmental disabilities. *Research in Developmental Disabilities, 21*(5), 377-91.

Carr, E. G., & Carlson, J. I. (1993). Reduction of severe behavior problems in the community using a multicomponent treatment approach. *Journal of Applied Behavior Analysis, 26*(2), 157-172. doi:10.1901/jaba.1993.26-157

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18(2)*, 111-26.

Carr, J. E., Severtson, J. M., & Lepper, T. L. (2009). Noncontingent reinforcement is an empirically supported treatment for problem behavior exhibited by individuals with developmental disabilities. *Research in Developmental Disabilities: A Multidisciplinary Journal, 30*(1), 44-57.

Caruthers, C. E., Lambert, J. M., Chazin, K. M., Harbin, E. R., & Houchins-Juarez, N. J. (2015). Latency-based FA as baseline for subsequent treatment evaluation. *Behavior Analysis in Practice, 8*(1), 48–51. https://doi.org/10.1007/s40617-015-0046-3

Call, N. A., Clark, S. B., Mevers, J. L., Parks, N. A., Volkert, V. M., & Scheithauer, M. C. (2018). An individualized method for establishing and thinning multiple schedules of reinforcement following functional communication training. *Learning and Motivation, 62*, 91–102. https://doi.org/10.1016/j.lmot.2017.03.006

Cole, C. L., & Levinson, T. R. (2002). Effects of within-activity choices on the challenging behavior of children with severe developmental disabilities. *Journal of Positive Behavior Interventions, 4(1)*, 29-37. doi:10.1177/109830070200400106

Cooper J. O, Heron T. E, & Heward W.L. (2007). Applied behavior analysis (2nd ed.) Upper Saddle River, NJ: Pearson; 2007.

Cuskelly, M., Gilmore, L., Glenn, S., & Jobling, A. (2016). Delay of gratification: A comparison study of children with down syndrome, moderate intellectual disability and typical development. *Journal of Intellectual Disability Research, 60*(9), 865–873. https://doi.org/10.1111/jir.12262

Cuskelly, M., Zhang, A., & Hayes, A. (2003). A mental age-matched comparison study of delay of gratification in children with Down syndrome. *International Journal of Disability, Development and Education, 50*(3), 239–251. https://doi.org/10.1080/1034912032000120435

Davies, L., & Oliver, C. (2013). The age related prevalence of aggression and self-injury in persons with an intellectual disability: A review. *Research in Developmental Disabilities: A Multidisciplinary Journal, 34*(2), 764-775.

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. doi:10.1007/s10864-017-9289-0

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. http://doi.org/10.1901/jaba.1996.29-519

DeLeon, I. G., Neidert, P. L., Anders, B. M., & Rodriguez-Catter, V. (2001). Choices between positive and negative reinforcement during treatment for escape-maintained behavior. *Journal of Applied Behavior Analysis, 34(4)*, 521-525. doi:10.1901/jaba.2001.34-521

Digenaro Reed, F. D., Hirst, J. M., & Hyman, S. R. (2012). Assessment and treatment of stereotypic behavior in children with autism and other developmental disabilities: A thirty year review. *Research in Autism Spectrum Disorders, 6*(1), 422-430.

Durand, V. M., & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18*(1), 99-117. doi:10.1007/BF02211821

Dyer, K., Dunlap, G., & Winterling, V. (1990). Effects of choice making on the serious problem behaviors of students with severe handicaps. *Journal of Applied Behavior Analysis, 23(4)*, 515-524. doi:10.1901/jaba.1990.23-515

Elliott, C., & Dillenburger, K. (2016). The effect of choice on motivation for young children on the autism spectrum during discrete trial teaching. *Journal of Research in Special Educational Needs, 16(3)*, 187-198. doi:10.1111/1471-3802.12073

Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R.,... Hatton, C. (2001). The prevalence of challenging behaviors: A total population study. *Research in Developmental Disabilities, 22*(1), 77-93. doi:10.1016/S0891-4222(00)00061-5

Epstein, R. (1985). Extinction-induced resurgence: Preliminary investigations and possible applications. *The Psychological Record, 35*(2), 143-153.

Falcomata, T. S., Roane, H. S., Feeney, B. J., & Stephenson, K. M. (2010). Assessment and treatment of elopement maintained by access to stereotypy. *Journal of Applied Behavior Analysis, 43*(3), 513–517. https://doi.org/10.1901/jaba.2010.43-513

Falcomata, T. S., Roane, H. S., Muething, C. S., Stephenson, K. M., & Ing, A. D. (2012). Functional communication training and chained schedules of reinforcement to treat challenging behavior maintained by terminations of activity interruptions. *Behavior Modification, 36*(5), 630–649. https://doi.org/10.1177/0145445511433821

Falcomata, T. S., White, P., Muething, C. S., & Fragale, C. (2012). A functional communication training and chained schedule procedure to treat challenging behavior with multiple functions. *Journal of Developmental and Physical Disabilities, 24*(6), 529–538. https://doi.org/10.1007/s10882-012-9287-z

Figueroa, R. G., Thyer, B. A., & Thyer, K. B. (1992). Extinction and DRO in the treatment of aggression in a boy with severe mental retardation. *Journal of Behavior Therapy and Experimental Psychiatry, 23*(2), 133-140. doi:10.1016/0005-7916(92)90011-7

Fisher, W. W., Greer, B. D., Fuhrman, A. M., & Querim, A. C. (2015). Using multiple schedules during functional communication training to promote rapid transfer of treatment effects. *Journal of Applied Behavior Analysis, 48(4),* 713-733. doi:10.1002/jaba.254

Fisher, W. W., Greer, B. D., Querim, A. C., & DeRosa, N. (2014). Decreasing excessive functional communication responses while treating destructive behavior using response restriction. *Research in Developmental Disabilities, 35*(11), 2614-2623. doi:10.1016/j.ridd.2014.06.024

Fisher, W. W., Thompson, R. H., Hagopian, L. P., Bowman, L. G., & Krug, A. (2000). Facilitating tolerance of delayed reinforcement during functional communication training. *Behavior Modification, 24(1),* 3-29. doi:10.1177/0145445500241001

Fritz, J. N., Jackson, L. M., Stiefler, N. A., Wimberly, B. S., & Richardson, A. R. (2017). Noncontingent reinforcement without extinction plus differential reinforcement of alternative behavior during treatment of problem behavior. *Journal of Applied Behavior Analysis, 50*(3), 590–599. https://doi.org/10.1002/jaba.395

Fuhrman, A. M., Fisher, W. W., & Greer, B. D. (2016). A Preliminary investigation on improving functional communication training by mitigating resurgence of destructive behavior. *Journal of Applied Behavior Analysis, 49*(4), 884-899.

Fyffe, C. E., Kahng, S. W., Fittro, E., & Russell, D. (2004). Functional analysis and treatment of inappropriate sexual behavior. *Journal of Applied Behavior Analysis, 37*(3), 401-404. doi:10.1901/jaba.2004.37-401

Gehrman, C., Wilder, D. A., Forton, A. P., & Albert, K. (2017). Comparing resetting to non-resetting DRO procedures to reduce stereotypy in a child with autism. *Behavioral Interventions, 32*(3), 242-247. doi:10.1002/bin.1486

Gerber, F., Bessero, S., Robbiani, B., Courvoisier, D. S., Baud, M. A., Traore, M., & Carminati, G. G. (2011). Comparing residential programmes for adults with autism spectrum disorders and intellectual disability: Outcomes of challenging behaviour and quality of life. *Journal of Intellectual Disability Research, 55*(9), 918-932. doi:10.1111/j.1365-2788.2011.01455.x

Gerow, S., Davis, T., Radhakrishnan, S., Gregori, E., & Rivera, G. (2018). Functional communication training: The strength of evidence across disabilities. *Exceptional Children.* Advance online publication. doi: 10.1177/0014402918793399

Ghaemmaghami, M., Hanley, G. P., & Jessel, J. (2016). Contingencies promote delay tolerance. *Journal of Applied Behavior Analysis, 49*(3), 548-575. doi:10.1002/jaba.333

Graff, R. B., Libby, M. E., & Green, G. (1998). The effects of reinforcer choice on rates of challenging behavior and free operant responding in individuals with severe disabilities. *Behavioral Interventions, 13(4),* 249-268. doi:10.1002/(SICI)1099-078X(199811)13:4<249::AID-BIN19>3.0.CO;2-1

Gregori, E., Rispoli, M., Gerow, S., & Lory, C. (2017). Treatment of self-injurious behavior in adults with intellectual and developmental disabilities: A systematic review. *Journal of Developmental and Physical Disabilities.* doi:10.1007/s10882-017-9568-7

Hagopian, L. P., Boelter, E. W., & Jarmolowicz, D. P. (2011). Reinforcement schedule thinning following functional communication training: Review and recommendations. *Behavior Analysis in Practice, 4(1),* 4-16.

Hagopian, L. P., Bruzek, J. L., Bowman, L. G., & Jennett, H. K. (2007). Assessment and treatment of problem behavior occasioned by interruption of free-operant behavior. *Journal of Applied Behavior Analysis, 40*(1), 89–103. https://doi.org/10.1901/jaba.2007.63-05

Hagopian, L. P., Crockett, J. L., van Stone, M., DeLeon, I. G., & Bowman, L. G. (2000). Effects of noncontingent reinforcement on problem behavior and stimulus engagement: The role of satiation, extinction, and alternative reinforcement. *Journal of Applied Behavior Analysis, 33*(4), 433–449. https://doi.org/10.1901/jaba.2000.33-433

Hagopian, L. P., Fisher, W. W., & Legacy, S. M. (1994). Schedule effects of noncontingent reinforcement on attention-maintained destructive behavior in identical quadruplets. *Journal of Applied Behavior Analysis, 27*(2), 317–325. https://doi.org/10.1901/jaba.1994.27-317

Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis, 31(2),* 211-235. doi:10.1901/jaba.1998.31-211

Hagopian, L. P., Toole, L. M., Long, E. S., Bowman, L. G., & Lieving, G. A. (2004). A comparison of dense-to-lean arid fixed lean schedules of alternative reinforcement and extinction. *Journal of Applied Behavior Analysis, 37*(3), 323–337. https://doi.org/10.1901/jaba.2004.37-323

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis, 34(1),* 17-38. doi:10.1901/jaba.2001.34-17

Hagopian, L., Kuhn, S. C., Long, E. S., & Rush, K. S. (2005). Schedule thinning following communication training: Using competing stimuli to enhance tolerance to decrements in reinforcer density. *Journal of Applied Behavior Analysis, 38(2),* 177-193. doi:10.1901/jaba.2005.43-04

Hammond, J. L., Iwata, B. A., Fritz, J. N., & Dempsey, C. M. (2011). Evaluation of fixed momentary DRO schedules under signaled and unsignaled arrangements. Journal of *Applied Behavior Analysis, 44*(1), 69–81. https://doi.org/10.1901/jaba.2011.44-69

Harding, J. W., Wacker, D. P., Berg, W. K., Barretto, A., & Rankin, B. (2002). Assessment and treatment of severe behavior problems using choice-making procedures. *Education and Treatment of Children, 25(1),* 26-46.

Heath, A. K., Ganz, J. B., Parker, R., Burke, M., & Ninci, J. (2015). A meta-analytic review of functional communication training across mode of communication, age, and disability. *Review Journal of Autism and Developmental Disorders, 2*, 155-166. doi: 10.1007/540489-014-0044-3

Holden, B., & Gitlesen, J. P. (2006). A total population study of challenging behaviour in the county of Hedmark, Norway: Prevalence, and risk markers. *Research in Developmental Disabilities: A Multidisciplinary Journal, 27*(4), 456-465.

Hong, E. R., Ganz, J. B., Mason, R., Morin, K., Davis, J. L., Ninci, J.,... Gilliland, W. D. (2016). The effects of video modeling in teaching functional living skills to persons with ASD: A meta-analysis of single-case studies. *Research in Developmental Disabilities, 57*,158-169. doi:10.1016/j.ridd.2016.07.001

Hong, E. R., Neely, L., Rispoli, M. J., Trepinski, T. M., Gregori, E., & Davis, T. (2015). A comparison of general and explicit delay cues to reinforcement for tangible-maintained challenging behaviour. *Developmental Neurorehabilitation, 18(*6), 395–401. https://doi.org/10.3109/17518423.2013.874378

Ing, A. D., Roane, H. S., & Veenstra, R. A. (2011). Functional analysis and treatment of coprophagia. *Journal of Applied Behavior Analysis, 44*(1), 151–155. https://doi.org/10.1901/jaba.2011.44-151

Ip, S. V., Szymanski, E. M., Johnston-Rodriguez, S., & Karls, S. F. (1994). Effects of staff implementation of a choice program on challenging behaviors in persons with developmental disabilities. *Rehabilitation Counseling Bulletin, 37(4),* 347-357.

Ivy, J. W., Meindl, J. N., Overley, E., & Robson, K. M. (2017). Token economy: A systematic review of procedural descriptions. *Behavior Modification, 41*(5), 708-737. doi:10.1177/0145445517699559

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27(2),* 197-209. doi:10.1901/jaba.1994.27-197

Jarmolowicz, D. P., DeLeon, I. G., & Contrucci Kuhn, S. A. (2009). Functional communication during signaled reinforcement and/or extinction. *Behavioral Interventions, 24*(4), 265–273. https://doi.org/10.1002/bin.288

Jarmolowicz, D. P., & Lattal, K. A. (2014). Resurgence under delayed reinforcement. *The Psychological Record, 64(2),* 189-193. doi:10.1007/s40732-014-0040-0

Jones, S., Cooper, S., Smiley, E., Allan, L., Williamson, A., & Morrison, J. (2008). Prevalence of, and factors associated with, problem behaviors in adults with intellectual disabilities. *Journal of Nervous and Mental Disease, 196*(9), 678-686. doi:10.9.1097/NMD.0b013e e318183f85c

Kahng, S., Iwata, B. A., & Lewin, A. B. (2002). Behavioral treatment of self-injury, 1964 to 2000. *American Journal on Mental Retardation, 107*(3), 212-21.

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

Kearney, C. A., & McKnight, T. J. (1997). Preference, choice, and persons with disabilities: A synopsis of assessments, interventions, and future directions. *Clinical Psychology Review, 17(2),* 217-238. doi:10.1016/S0272-7358(97)00004-4

Kern, L., & Clemens, N. H. (2007). Antecedent Strategies to Promote Appropriate Classroom Behavior. *Psychology in the Schools, 44(1),* 65-75. doi:10.1002/pits.20206

Kern, L., Starosta, K., & Adelman, B. E. (2006). Reducing Pica by Teaching Children to Exchange Inedible Items for Edibles. *Behavior Modification, 30*(2), 135–158. https://doi.org/10.1177/0145445505283414

Knox, M., Rue, H. C., Wildenger, L., Lamb, K., & Luiselli, J. K. (2012). Intervention for food selectivity in a specialized school setting: Teacher implemented prompting, reinforcement, and demand fading for an adolescent student with autism. *Education & Treatment of Children, 35*(3), 407-417. doi:10.1353/etc.2012.0016

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*, 26–38. doi:10.1177/0741932512452794

Lalli, J. S., Casey, S., & Kates, K. (1995). Reducing escape behavior and increasing task completion with functional communication training, extinction, and response chaining. *Journal of Applied Behavior Analysis, 28*(3), 261-268. doi:10.1901/jaba.1995.28-261

Lambert, J. M., Parikh, N., Stankiewicz, K. C., Houchins-Juarez, N. J., Morales, V. A., Sweeney, E. M., & Milam, M. E. (2019). Decreasing food stealing of child with Prader-Willi syndrome through function-based differential reinforcement. *Journal of Autism and Developmental Disorders, 49*(2), 721–728. https://doi.org/10.1007/s10803-018-3747-y

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*(3), 407–414. https://doi.org/10.1901/jaba.2003.36-407

LeBlanc, L. A., Hagopian, L. P., Maglieri, K. A., & Poling, A. (2002). Decreasing the intensity of reinforcement-based interventions for reducing behavior: Conceptual issues and a proposed model for clinical practice. *The Behavior Analyst Today, 3(3)*, 289-300. doi:10.1037/h0099991

Lehardy, R. K., Lerman, D. C., Evans, L. M., O'Connor, A., & LeSage, D. L. (2013). A simplified methodology for identifying the function of elopement. *Journal of Applied Behavior Analysis, 46*(1), 256-270. doi:10.1002/jaba.22

Lloyd, B. P., & Kennedy, C. H. (2014). Assessment and treatment of challenging behaviour for individuals with intellectual disability: A research review. *Journal of Applied Research in Intellectual Disabilities, 27*(3), 187-199. doi:10.1111/jar.12089

Logue, A. W. (1988). Research on self-control: An integrating framework. *Behavioral and Brain Sciences, 11*, 665-709.

Lowe, K., Allen, D., Jones, E., Brophy, S., Moore, K., & James, W. (2007). Challenging behaviours: Prevalence and topographies. *Journal of Intellectual Disability Research, 51*(8), 625-636. doi:10.1111/j.1365-2788.2006.00948.x

Lundqvist, L. (2013). Prevalence and risk markers of behavior problems among adults with intellectual disabilities: A total population study in Orebro County, Sweden. *Research in Developmental Disabilities: A Multidisciplinary Journal, 34*(4), 1346-1356.

Lustig, N. H., Ringdahl, J. E., Breznican, G., Romani, P., Scheib, M., & Vinquist, K. (2014). Evaluation and treatment of socially inappropriate stereotypy. *Journal of Developmental and Physical Disabilities, 26*(2), 225-235. doi:10.1007/s10882-013-9357-x

Maggin, D. M., Briesch, A. M., & Chafouleas, S. M. (2013). An application of the What Works Clearinghouse standards for evaluating single-subject research: Synthesis of the self-management literature base. *Remedial and Special Education, 34*(1), 44-58. doi:10.1177/0741932511435176

Mace, F. C., McComas, J. J., Mauro, B. C., Progar, P. R., Taylor, B., Ervin, R., & Zangrillo, A. N. (2010). Differential reinforcement of alternative behavior increases resistance to extinction: Clinical demonstration, animal modeling, and clinical test of one solution. *Journal of the Experimental Analysis of Behavior, 93*(3), 349–367. https://doi.org/10.1901/jeab.2010.93-349

MacNaul, H. L., & Neely, L. C. (2018). Systematic review of differential reinforcement of alternative behavior without extinction for individuals with autism. *Behavior Modification, 42*(3), 398-421. doi:10.1177/0145445517740321

Martinez, J. R., Werch, B. L., & Conroy, M. A. (2016). School-based interventions targeting challenging behaviors exhibited by young children with autism spectrum disorder: A systematic literature review. *Education and Training in Autism and Developmental Disabilities, 51*(3), 265-280.

Matson, J. L., & Boisjoli, J. A. (2009). The token economy for children with intellectual disability and/or autism: A review. *Research in Developmental Disabilities: A Multidisciplinary Journal, 30*(2), 240-248.

Matson, J. L., Dixon, D. R., & Matson, M. L. (2005). Assessing and treating aggression in children and adolescents with developmental disabilities: A 20-year overview. *Educational Psychology*, *25*(2), 151-181.

Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. *Behavior analysis for lasting change* (3rd ed.). Cornwall-on-Hudson, NY: Sloan Publishing; 2014.

Mazaleski, J. L., Iwata, B. A., Vollmer, T. R., Zarcone, J. R., & Smith, R. G. (1993). Analysis of the reinforcement and extinction components in DRO contingencies with self-injury. *Journal of Applied Behavior Analysis, 26*(2), 143-156. doi:10.1901/jaba.1993.26-143

Michael, J. (2000). Implications and refinements of the establishing operation concept. *Journal of Applied Behavior Analysis, 33*, 401-410

Michael, J. (1993). Establishing operations. *The Behavior Analyst, 16*,191–206

Murphy, G. H., Beadle-Brown, J., Wing, L., Gould, J., Shah, A., & Holmes, N. (2005). Chronicity of Challenging Behaviours in People with Severe Intellectual Disabilities and/or Autism: A Total Population Sample. *Journal of Autism and Developmental Disorders,* 35(4), 405-418. doi:10.1007/s10803-005-5030-2

Nevin, J. A., & Wacker, D. P. (2013). Response strength and persistence. In G. J. Madden (Ed.), *APA Handbook of Behavior Analysis* (Vol. 2, pp. 109-128).

O'Donovan, M. A., Byrne, E., McCallion, P., & McCarron, M. (2017). Measuring choice for adults with an intellectual disability—A factor analysis of the adapted Daily Choice Inventory Scale. *Journal of Intellectual Disability Research, 61*(5), 471–487. https://doi.org/10.1111/jir.12364

Parker, R. I., Vannest, K. J. & Brown, L. (2009) 'The improvement rate difference for single-case research', *Exceptional Children, 75*, 135–150.

Parker, R. I., Vannest, K. J., Davis, J. L. & Sauber, S. B. (2011) 'Combining non-overlap and trend for single case research: Tau-U', *Behavior Therapy*, *42*, 284–299.

Peery, A., & Fisher, W. W. (2001). Behavioral economic influences on treatments designed to decrease destructive behavior. Journal of Applied Behavior Analysis, 34(2), 211–215. https://doi.org/10.1901/jaba.2001.34-211

Petscher, E. S., Rey, C., & Bailey, J. S. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities: A Multidisciplinary Journal, 30*(3), 409-425.

Petursdottir, A. I., Esch, J. W., Sautter, R. A., & Stewart, K. K. (2010). Characteristics and hypothesized functions of challenging behavior in a community-based sample. *Education and Training in Autism and Developmental Disabilities, 45*(1), 81-93.

Pipkin, C. S. P., 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

Piazza, C. C., Hanley, G. P., Bowman, L. G., Ruyter, J. M., Lindauer, S. E., & Saiontz, D. M. (1997). Functional analysis and treatment of elopement. *Journal of Applied Behavior Analysis, 30*(4), 653–672. https://doi.org/10.1901/jaba.1997.30-653

Poppes, P., Van der Putten, A. J., & Vlaskamp, C. (2014). Addressing challenging behavior in people with profound intellectual and multiple disabilities: Analyzing the effects of daily practice. *Journal of Policy and Practice in Intellectual Disabilities, 11*(2), 128-136. doi:10.1111/jppi.12078

Rakap, S. (2015). Effect sizes as result interpretation aids in single-subject experimental research: Description and application of four nonoverlap methods. *British Journal of Special Education, 42*(1), 11-33.

Rapp, J. T., Vollmer, & Timothy, R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities: A Multidisciplinary Journal, 26*(6), 527-547.

Rasmussen, K., & O'Neill, R. E. (2006). The effects of fixed-time reinforcement schedules on problem behavior of children with emotional and behavioral disorders in a day-treatment classroom setting. *Journal of Applied Behavior Analysis, 39*(4), 453–457. https://doi.org/10.1901/jaba.2006.172-05

Reichle, J., Johnson, L., Monn, E., & Harris, M. (2010). Task Engagement and Escape Maintained Challenging Behavior: Differential Effects of General and Explicit Cues when Implementing a Signaled Delay in the Delivery of Reinforcement. *Journal of Autism and Developmental Disorders, 40*(6), 709–720. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ882644&site=ehost-live

Reynolds, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior, 4*, 57-71. doi:10.1901/jeab.1961.4-57

Richman, D. M., Barnard-Brak, L., Grubb, L., Bosch, A., & Abby, L. (2015). Meta analysis of noncontingent reinforcement effects on problem behavior. *Journal of Applied Behavior Analysis, 48*(1), 131-152. doi:10.1002/jaba.189

Richman, D. M., & Hagopian, L. P. (1999). On the effects of "quality" of attention in the functional analysis of destructive behavior. *Research in Developmental Disabilities, 20(*1), 51–62. https://doi.org/10.1016/S0891-4222(98)00031-6

Rispoli, M., Brodhead, M., Wolfe, K., & Gregori, E. (2018). Trial-based functional analysis informs treatment for vocal scripting. *Behavior Modification, 42*(3), 441-465. doi:10.1177/0145445517742882

Rispoli, M. J., Davis, H. S., Goodwyn, F. D., & Camargo, S. (2013). The use of trial-based functional analysis in public school classrooms for two students with developmental disabilities. *Journal of Positive Behavior Interventions, 15*(3), 180–189. https://doi.org/10.1177/1098300712457420

Rispoli, M., Lang, R., Neely, L., Camargo, S., Hutchins, N., Davenport, K., & Goodwyn, F. (2013). A comparison of within- and across-activity choices for reducing challenging behavior in children with autism spectrum disorders. *Journal of Behavioral Education, 22(1),* 66-83. doi:10.1007/s10864-012-9164-y

Roane, H. S., Falcomata, T. S., & Fisher, W. W. (2007). Applying the behavioral economics principle of unit price to DRO schedule thinning. *Journal of Applied Behavior Analysis, 40*(3), 529–534. https://doi.org/10.1901/jaba.2007.40-529

Roane, H. S., Fisher, W. W., Sgro, G. M., Falcomata, T. S., & Pabico, R. R. (2004). An alternative method of thinning reinforcer delivery during differential reinforcement. *Journal of Applied Behavior Analysis, 37*(2),

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

Rooker, G. W., Jessel, J., Kurtz, P. F., & Hagopian, L. P. (2013). Functional communication training with and without alternative reinforcement and punishment: An analysis of 58 applications. *Journal of Applied Behavior Analysis, 46*(4), 708–722. https://doi.org/10.1002/jaba.76

Saini, V., Fisher, W. W., & Pisman, M. D. (2017). Persistence during and resurgence following noncontingent reinforcement implemented with and without extinction. *Journal of Applied Behavior Analysis, 50*(2), 377-392. doi:10.1002/jaba.380

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). 'The quantitative synthesis of single-subject research: Methodology and validation': Reply to Owen White. RASE: *Remedial & Special Education, 8*(2), 40-42. doi:10.1177/074193258700800208

Seybert, S., Dunlap, G., & Ferro, J. (1996). The effects of choice-making on the problem behaviors of high school students with intellectual disabilities. *Journal of Behavioral Education, 6(1),* 49-65. doi:10.1007/BF02110477

Shogren, K. A., Faggella-Luby, M. N., Bae, S. J., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior: A meta-analysis. *Journal of Positive Behavior Interventions, 6(4),* 228-237. doi:10.1177/10983007040060040401

Shukla, S., & Albin, R. W. (1996). Effects of extinction alone and extinction plus functional communication training on covariation of problem behaviors. *Journal of Applied Behavior Analysis, 29*(4), 565-68.

Sidener, T. M., Shabani, D. B., Carr, J. E., & Roland, J. P. (2006). An evaluation of strategies to maintain mands at practical levels. *Research in Developmental Disabilities, 27(6),* 632-644. doi:10.1016/j.ridd.2005.08.002

Sigafoos, J., & Saggers, E. (1995). A discrete-trial approach to the functional analysis of aggressive behaviour in two boys with autism. *Australia & New Zealand Journal of Developmental Disabilities, 20*(4), 287-297

Simmons, J. N., Smith, R. G., & Kliethermes, L. (2003). A multiple-schedule evaluation of immediate and subsequent effects of fixed-time food presentation on automatically maintained mouthing. *Journal of Applied Behavior Analysis, 36*(4), 541–544. https://doi.org/10.1901/jaba.2003.36-541

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. Oxford: Appleton-Century.

Slocum, S. K., Grauerholz-Fisher, E., Peters, K. P., & Vollmer, T. R. (2018). A multicomponent approach to thinning reinforcer delivery during noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis, 51*(1), 61-69. doi:10.1002/jaba.427

Smith, S. W., Daunic, A. P., & Taylor, G. G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education & Treatment of Children, 30*(4), 121–134. https://doi.org/10.1353/etc.2007.0033

Soares, D. A., Harrison, J. R., Vannest, K. J., & McClelland, S. S. (2016). Effect size for token economy use in contemporary classroom settings: A meta-analysis of single-case research. *School Psychology Review, 45*(4), 379-399.

Stevenson, M. T., Ghezzi, P. M., & Valenton, K. G. (2016). FCT and delay fading for elopement with a child with autism. *Behavior Analysis in Practice, 9*(2), 169-173. doi:10.1007/s40617-015-0089-5

Sullivan, R. C. (2007). Position paper on the national crisis in adult services for individuals with autism: A call to action. Bethesda: *Autism Society of America*, pp. 1–16

Tucker, M., Sigafoos, J., & Bushell, H. (1998). Use of Noncontingent reinforcement in the treatment of challenging behavior: A review and clinical guide. *Behavior Modification, 22*(4), 529-47.

Tiger, J. H., Fisher, W. W., & Bouxsein, K. J. (2009). Therapist- and self-monitored DRO contingencies as a treatment for the self-injurious skin picking of a young man with Asperger syndrome. *Journal of Applied Behavior Analysis, 42*(2), 315-319. doi:10.1901/jaba.2009.42-315

Tiger, J. H., Hanley, G. P., & Bruzek, J. (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice, 1(1)*, 16-23.

Torres-Viso, M., Strohmeier, C. W., & Zarcone, J. R. (2018). Functional analysis and treatment of problem behavior related to mands for rearrangement. *Journal of Applied Behavior Analysis, 51*(1), 158-165. doi:10.1002/jaba.437

Totsika, V., Toogood, S., Hastings, R. P., & Lewis, S. (2008). Persistence of challenging

behaviours in adults with intellectual disability over a Period of 11 years. *Journal of Intellectual Disability Research, 52*(5), 446-457.

Umbreit, J., & Blair, K. (1996). The effects of preference, choice, and attention on problem behavior at school. *Education & Training in Mental Retardation & Developmental Disabilities, 31(2),* 151-161.

Vaughn, B. J., & Horner, R. H. (1997). Identifying instructional tasks that occasion problem behaviors and assessing the effects of student versus teacher choice among these tasks. *Journal of Applied Behavior Analysis, 30(2),* 299-312. doi:10.1901/jaba.1997.30-299

Volkert, V. M., Lerman, D. C., Call, N. A., & Trosclair-Lasserre, N. (2009). An evaluation of resurgence during treatment with functional communication training. *Journal of Applied Behavior Analysis, 42(1),* 145-160. doi:10.1901/jaba.2009.42-145

Vollmer, T. R., Borrero, J. C., Lalli, J. S., & Daniel, D. (1999). Evaluating self-control and impulsivity in children with severe behavior disorders. *Journal of Applied Behavior Analysis, 32*(4), 451–466. https://doi.org/10.1901/jaba.1999.32-451

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis, 26*(1), 9-21. doi:10.1901/jaba.1993.26-9

Wacker, D. P., Berg, W. K., & Harding, J. W. (2004). *Maintenance effects of functional communication training*. Washington, DC: Department of Health and Human Services, National Institute on Child Health and Human Development.

Wacker, D. P., Harding, J. W., Berg, W. K., Lee, J. F., Schieltz, K. M., Padilla, Y. C., & ... Shahan, T. A. (2011). An evaluation of persistence of treatment effects during long-term treatment of destructive behavior. *Journal of the Experimental Analysis of Behavior, 96*(2), 261-282.

Wacker, D. P., Harding, J. W., Morgan, T. A., Berg, W. K., Schieltz, K. M., Lee, J. F., & Padilla, Y. C. (2013). An evaluation of resurgence during functional communication training. *Psychological Record, 63*(1), 3-20.

Wacker, D. P., Schieltz, K. M., Berg, W. K., Harding, J. W., Dalmau, Y. C. P., & Lee, J. F. (2017). The long-term effects of functional communication training conducted in young children's home settings. *Education & Treatment of Children, 40*(1), 43–56. https://doi.org/10.1353/etc.2017.0003

Wehmeyer, M. L. (2005). Self-determination and individuals with severe disabilities: Re-examining meanings and misinterpretations. *Research and Practice for Persons with Severe Disabilities, 30*(3), 113–120. https://doi.org/10.2511/rpsd.30.3.113

Wehmeyer, M. L., & Abery, B. H. (2013). Self-determination and choice. *Intellectual and Developmental Disabilities, 51*(5), 399–411. https://doi.org/10.1352/1934-9556-51.5.399

Wehmeyer, M. L., Agran, M., & Hughes, C. (1998). *Teaching self-determination to students with disabilities: Basic skills for successful transition*. Baltimore, MD, US: Paul H Brookes.

Wehmeyer, M. L., Kelchner, K., & Richards, S. (1996). Essential characteristics of self-determined behavior of individuals with mental retardation. *American Journal on Mental Retardation, 100*(6), 632–642. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1996-01616-006&site=ehost-live

Weston, R., Hodges, A., & Davis, T. N. (2018). Differential reinforcement of other behaviors to treat challenging behaviors among children with autism: A systematic and quality review. *Behavior Modification, 42*(4), 584-609. doi:10.1177/0145445517743487

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders, 45(7),* 1951-1966.

Zarcone, J. R., Rodgers, T. A., Iwata, B. A., Rourke, D. A., & Dorsey, M. F. (1991). Reliability analysis of the Motivation Assessment Scale: A failure to replicate. *Research in Developmental Disabilities, 12*(4), 349-360.http://dx.doi.org/10.1016/0891-4222(91)90031-M

Zelinsky, N. M., & Shadish, W. (2018). A demonstration of how to do a meta-analysis that combines single-case designs with between-groups experiments: The effects of choice making on challenging behaviors performed by people with disabilities. *Developmental Neurorehabilitation, 21(4),* 266-278. doi:10.3109/17518423.2015.1100690

APPENDIX A. TABLES

Table A1

Search terms

Category One	Category Two
Differential reinforcement	Schedul*
Differential reinforcement of alternative behavior	Schedule thinning
Differential reinforcement of other behavior	Delay to reinforcement
Differential reinforcement of zero responding	Multiple schedule*
Omission training	Delay schedule*
Differential reinforcement of low rates	Chained schedule*
Differential reinforcement of diminishing rates	Response restriction
Differential reinforcement of high rates	Signal* delay
Noncontingent reinforcement	Unsignal* delay
Functional communication training	Demand fading
Functional equivalence training	Delay fading
Token economy	Delay to reinforcement
Differential negative reinforcement of alternative behavior	Response chaining
	Delayed reinforcement
	Prompt fading

Note. All terms were searched with quotation marks to capture the exact terms.

Table A2

WWC Basic Design Standards

	Meets Standards	Meets Standards with Reservations	Does not Meet Standards
Design Standard 1	Researcher systematically manipulated the IV	N/A	Researcher failed to systematically manipulate the IV
Design Standards 2A	IAA was collected		IAA was not collected
Design Standard 2B	IAA was collected on a minimum of 20% of data points across all study phases (baseline and intervention)	IAA was collected on a minimum of 20% of data points across all study phases (baseline and intervention combined)	IAA was collected on less than 20% of data points across study phases
Design Standard 2C	IAA was above 80% or 0.6 Kappa	N/A	IAA was below minimal thresholds
Design Standard 3	Researcher made at least three attempts to demonstrate an experimental effect at three different points in time	N/A	Researcher made less than three attempts to demonstrate an experimental effect at three different points in time
Design Standard 4	Experiment included at least 5 data points per phase (baseline and intervention)	Experiment included between 3-4 data points per phase	Experiment included less than 3 data points per phase

Note. Standards were adapted based on the procedures described by Maggin et al. (2013) and Hong et al. (2016); DV=dependent variable; IAA=interassessor agreement; IV=independent variable.

Table A3

WWC Evidence Standards

22	Baseline Analysis	Within Phase Analysis	Between Phase Basic Effects	Between Phase Experimental Effect	Overall Effectiveness
Do the data document a pattern of behavior in need of change?	Yes	N/A	N/A	N/A	N/A
Do the data demonstrate a predictable pattern?	Yes	N/A	N/A	N/A	N/A
Is the variability sufficiently consistent?	Yes	N/A	N/A	N/A	N/A
Is the trend either stable or moving away from the therapeutic direction?	Yes	N/A	N/A	N/A	N/A
Does each phase, including baseline, have at least three data points?	N/A	Yes	N/A	N/A	N/A
Do the data within each phase non-baseline document a predictable data pattern?	N/A	Yes	N/A	N/A	N/A
Is the variability sufficiently consistent?	N/A	Yes	N/A	N/A	N/A
Is the trend either sufficiently low or moving in the hypothesized direction?	N/A	Yes	N/A	N/A	N/A
Does between phase data document the presence of basic effects?	N/A	N/A	Yes	N/A	N/A
Is the level discriminably different between the first and last three data points in adjacent phases?	N/A	N/A	Yes	N/A	N/A

Table A3

WWC Evidence Standards

	Baseline Analysis	Within Phase Analysis	Between Phase Basic Effects	Between Phase Experimental Effect	Overall Effectiveness
Is the trend discriminably different between the first and last three data points in adjacent phases?	N/A	N/A	Yes	N/A	N/A
Is there an overall change in level between baseline and treatment phases?	N/A	N/A	Yes	N/A	N/A
Is there an overall change in trend between baseline and treatment phases?	N/A	N/A	Yes	N/A	N/A
Report the design	N/A	N/A	N/A	N/A	Yes
How many opportunities were there to demonstrate a treatment effect?	N/A	N/A	N/A	N/A	Yes
How many treatment effects were demonstrated?	N/A	N/A	N/A	N/A	Yes
Data points per phase: 5 data points indicates strong evidence; 3-4 data points indicates moderate evidence; less than 3 data points per phase indicates no evidence	N/A	N/A	N/A	N/A	Yes
Total demonstrations of treatment effect. There must be at least three demonstrations of an effect.	N/A	N/A	N/A	N/A	Yes
Ratio of effects to non-effects. No instances of a non-effect indicates strong evidence; ratio of 3 effects to 1 non-effect indicates moderate evidence; if the ratio of effects to non-effects is greater than 3:1 the experiment demonstrates no evidence	N/A	N/A	N/A	N/A	Yes

Note. Definitions for each Evidence Standard were obtained from Maggin et al. (2013); N/A= not applicable.

Table A4

Evidence standards for schedule thinning

	2 Strong Evidence	1 Moderate Evidence	0 No Evidence
Overall Effect			
ES 1: Between Phase Effect – Treatment	80-100% of data points in the schedule thinning phase overlap with or are below the data last treatment phase	79-51% of data points in the schedule thinning phase overlap with or are below the data in the last treatment phase	50% or less of the data points in the schedule thinning phase overlap with or are below the data in the last treatment phase
ES 2: Between Phase Effect – Baseline	0% of data points in the schedule thinning phase overlap with or exceed data in the last baseline phase	1-30% of data points in the schedule thinning phase overlap with or exceed the data in the last baseline phase	More than 30% of data points in the schedule thinning phase overlap with or exceed the data in the last baseline phase
ES 3: Data Points	The schedule thinning phase contains five or more data points	The schedule thinning phase contains 3-4 data points	The schedule thinning phase contains less than 3 data points
ES 4: Overall Effect	All previous ES received scores of 2	One or more of the previous ES received a score of 1, and no ES received a score of 0	One or more of the previous ES received a score of 0
Initial Effect The following standards refer to the 30% of data points in the schedule thinning phase			
ES 4: Between Phase Effect – Treatment	80-100% of data points in the schedule thinning phase overlap with or are below the data last treatment phase	79-51% of data points in the schedule thinning phase overlap with or are below the data in the last treatment phase	50% or less of the data points in the schedule thinning phase overlap with or are below the data in the last treatment phase
ES 5: Between Phase Effect – Baseline	0% of data points in the schedule thinning phase overlap with or exceed data in the last baseline phase	1-30% of data points in the schedule thinning phase overlap with or exceed the data in the last baseline phase	More than 30% of data points in the schedule thinning phase overlap with or exceed the data in the last baseline phase
ES 6: Overall Effect	All previous ES received scores of 2	One or more of the previous ES received a score of 1, and no ES received a score of 0	One or more of the previous ES received a score of 0

Table A4

Evidence Standards for Schedule Thinning

Terminal Effect			
The following standards apply to the last 30% of data points of the schedule thinning phase			
ES 7: Between Phase Effect – Treatment	80-100% of data points in the schedule thinning phase overlap with or are below the data last treatment phase	79-51% of data points in the schedule thinning phase overlap with or are below the data in the last treatment phase	50% or less of the data points in the schedule thinning phase overlap with or are below the data in the last treatment phase
ES 8: Between Phase Effect – Baseline	0% of data points in the schedule thinning phase overlap with or exceed data in the last baseline phase	1-30% of data points in the schedule thinning phase overlap with or exceed the data in the last baseline phase	More than 30% of data points in the schedule thinning phase overlap with or exceed the data in the last baseline phase
ES 9: Overall Effect	All previous ES received scores of 2	One or more of the previous ES received a score of 1, and no ES received a score of 0	One or more of the previous ES received a score of 0

Note. Schedule thinning evidence standards are based on existing standards developed by Kratochwill et al., 2010; Maggin et al., 2013; and Neely et al., 2018.

Table A5

Interrater agreement

Study phase	Percent agreement
Inclusion evaluation	90
Basic Design Standard Evaluation	88
Single case evidence evaluation	93
Schedule thinning evidence evaluation	85
Descriptive evaluation	86

Table A6

Results of the design evaluation

	Meets without reservations	Meets with reservations	Does not meet
Design Standard 1	173	N/A	0
Design Standard 2A	166	N/A	7
Design Standard 2B	24	120	31
Design Standard 2C	135	N/A	12
Design Standard 3	72	N/A	101
Design Standard 4	63	85	25
Overall	3	43	127

Note. N/A= not applicable.

Table A7

Results of the evidence evaluation

Evidence Standard	Number of experiments
Criteria 1: Data Points per Phase	
• Five or more	16
• 3-4	60
• Less than three data points	0
Criteria 2: Total Demonstrations of Effect	
• At least three demonstrations	65
• Less than three demonstrations	11
Criteria 3: Ratio of Effects to Non-Effects	
• No instances of non-effect	64
• The ratio of effects to non-effects is less than or equal to 3:1	1
• The ratio of effects to non-effects is greater than 3:1	11
Overall Effect	
• Strong Effect	12
• Moderate Effect	53
• No Effect	11

Table A8

Results of the schedule thinning evidence evaluation

		Overall Effect				Initial Effect			Terminal Effect		
	Participant	Between phase effect 1	Between phase effect 2	Data points per phase	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect
Britton et al., 2000	Rob	2	2	2	2	2	2	2	2	2	2
Britton et al., 2000	Todd	1	2	2	1	0	2	0	2	2	2
Britton et al., 2000	Victor (DRO)	2	0	2	0	2	0	0	2	0	0
	Victor (NCR)	2	0	2	0	2	0	0	2	0	0
Call et al., 2018	Jose	0	1	2	0	0	1	0	1	1	1
Caruthers et al., 2015	Alison	2	1	2	1	1	0	0	2	2	2
Davis et al., 2012	Eli	2	2	2	2	2	2	2	2	2	2
Falcomata, Roane, et al., 2012	Steph	2	2	2	2	2	2	2	2	2	2
Falcomata, Roane, et al., 2012	Mike	2	2	1	1	2	2	2	2	2	2
Falcomata, White, et al., 2012	Danny	2	2	2	2	2	2	2	2	2	2
Falcomata, Muething, et al., 2013	Alonzo	2	2	2	2	2	2	2	2	2	2
Falcomata, Muething, et al., 2013	Joe	2	2	2	1	2	2	2	2	2	2

Table A8

Results of the schedule thinning evidence evaluation

		Overall Effect				Initial Effect			Terminal Effect		
	Participant	Between phase effect 1	Between phase effect 2	Data points per phase	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect
Falcomata, Muething, et al., 2013	Alonzo experiment 2	2	2	2	2	2	2	2	2	2	2
Falcomata, Roane, Feeney, et al., 2010	John	2	2	2	2	2	2	2	2	2	2
Fisher et al., 2015	Jacob (room B)	2	2	2	2	2	2	2	2	2	2
Fisher et al., 2015	Jacob (room C)	2	2	2	2	2	2	2	2	2	2
Fisher et al., 2015	Jacob (room A)	2	2	2	2	2	2	2	2	2	2
Fisher et al., 1996	Mat	2	2	2	2	2	2	2	2	2	2
Fritz et al., 2017	Charley	2	2	2	2	2	2	2	2	2	2
Fritz et al., 2017	Gilbert	2	2	2	2	2	2	2	2	2	2
Fritz et al., 2017	Dyson	2	2	2	2	2	2	2	2	2	2
Fyffe et al., 2004	Matt	1	2	2	1	1	2	1	2	2	2
Hagopian et al., 1994	Wanda	2	1	2	1	1	2	1	2	2	2
Hagopian et al., 1994	Laurie	0	0	2	0	0	0	0	2	0	0
Hagopian et al., 1994	Lynn	2	2	2	2	2	2	2	2	2	2

Table A8

Results of the schedule thinning evidence evaluation

		Overall Effect				Initial Effect			Terminal Effect		
	Participant	Between phase effect 1	Between phase effect 2	Data points per phase	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect
Hagopian et al., 1994	Glenda	2	0	2	0	2	0	0	2	0	0
Hagopian et al., 1998	Case 17	2	0	2	0	1	0	0	2	0	0
Hagopian et al., 1998	Case 19	2	2	2	2	2	2	2	2	2	2
Hagopian et al., 2000	Jack	0	1	2	0	0	0	0	0	2	0
Hagopian et al., 2000	Rex	0	1	2	0	0	1	0	1	2	1
Hagopian et al., 2000	Alex	1	1	2	1	0	0	0	1	1	1
Hagopian et al., 2000	Emily	1	1	2	1	0	0	0	2	2	2
Hagopian et al., 2004	Brent	2	1	2	1	2	2	2	2	2	2
Hagopian et al., 2004	Jason	2	1	2	1	2	0	0	2	2	2
Hagopian et al., 2004	Sally	0	1	2	0	0	0	0	0	0	0
Hagopian et al., 2005	Stephen (FCT w/ EXT)	1	2	2	1	2	2	2	0	2	0
Hagopian et al., 2005	Stephen (FCT w/ EXT plus CS)	2	2	2	2	2	2	2	2	2	2
Hagopian et al., 2005	James (FCT w/ EXT)	0	2	2	0	2	2	2	0	2	0

Table A8

Results of the schedule thinning evidence evaluation

		Overall Effect				Initial Effect			Terminal Effect		
	Participant	Between phase effect 1	Between phase effect 2	Data points per phase	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect
Hagopian et al., 2005	James (FCT w/EXT plus CS)	2	2	2	2	2	2	2	1	2	1
Hagopian et al., 2005	Matt (FCT w/ EXT)	2	1	2	1	2	2	2	2	1	2
Hagopian et al., 2005	Matt (FCT w/ EXT plus CS)	2	2	2	2	2	2	2	2	2	2
Hagopian et al., 2007	Maxwell	0	1	2	0	0	1	0	1	1	1
Hagopian et al., 2007	Kelly	1	2	2	1	1	2	1	2	2	2
Hagopian et al., 2007	Perry	0	1	2	0	0	1	0	1	1	1
Hammond et al., 2011	Alex	2	2	2	2	2	2	2	2	2	2
Hanley et al., 2001	Karen (delay-to-reinforcement)	1	2	2	1	2	2	2	0	2	0
	Karen (multiple schedule)	2	2	2	2	1	2	1	2	2	2
Hanley et al., 2001	Jake (multiple schedule)	0	1	2	0	1	1	1	1	1	1
	Jake (mixed schedule)	0	0	2	0	0	0	0	1	2	1
Hanley et al., 2001	Julie (multiple schedule)	2	2	2	2	2	2	2	2	2	2
	Julie (mixed schedule)	2	1	2	1	2	2	2	1	2	1
Ing et al., 2011	Sara	1	2	2	1	2	2	2	2	2	2

Table A8

Results of the schedule thinning evidence evaluation

		Overall Effect				Initial Effect			Terminal Effect		
	Participant	Between phase effect 1	Between phase effect 2	Data points per phase	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect
Jarmolowicz et al., 2009	Multiple schedule signaled FR1	0	2	2	0	0	2	0	0	2	0
	Multiple schedule signaled EXT	0	2	2	0	0	2	0	2	2	2
	Multiple schedule signaled FR1/EXT	0	2	2	0	0	2	0	0	2	0
Kern et al., 2006	Orlando	2	1	2	1	2	2	2	2	1	1
	Matthew	2	2	2	2	2	2	2	2	2	2
Lalli et al., 1997	Donny	2	1	2	1	2	0	0	2	2	2
	Harry	2	0	2	0	2	0	0	2	0	0
Lalli et al., 1998	Val	2	2	2	2	2	2	2	2	2	2
Lambert et al., 2019	Leah	2	1	2	1	2	2	2	2	2	2
Peery & Fisher, 2001	Ann	2	2	2	2	2	1	1	2	2	2
Piazza et al., 1997	Ray	2	2	2	2	1	2	1	1	2	1
Piazza et al., 1997	Ty	2	2	1	1	2	2	2	2	2	2
Rasmussen et al., 2006	Josh	1	1	2	1	1	1	1	2	2	2
Rasmussen et al., 2006	Mike	0	2	0	0	N/A	N/A	N/A	N/A	N/A	N/A

Table A8

Results of the schedule thinning evidence evaluation

		Overall Effect				Initial Effect			Terminal Effect		
	Participant	Between phase effect 1	Between phase effect 2	Data points per phase	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect	Between phase effect 1	Between phase effect 2	Overall effect
Rasmussen et al., 2006	Chad	0	2	1	0	2	2	2	0	2	0
Richman et al., 1999	Don	1	2	2	1	1	2	1	1	2	1
Roane et al., 2007	Fred (experiment 1)	0	0	2	0	0	2	0	0	0	0
Roane et al., 2007	Fred (experiment 2)	2	2	2	2	2	2	2	2	2	2
Rooker et al., 2013	Application 36	1	1	2	1	1	1	1	0	1	0
Rooker et al., 2013	Application 6	1	2	1	1	1	2	1	0	2	0
Simmons et al., 2003	Alicia	1	2	2	1	2	2	2	1	2	1
Slocum et al., 2016	Abby	2	0	2	0	1	0	0	2	0	0
Slocum et al., 2018	Clancy	2	2	2	2	2	2	2	2	2	2
Slocum et al., 2018	Korey	2	2	2	2	2	2	2	2	2	2
Slocum et al., 2018	Reginald	2	2	2	2	2	2	2	2	2	2
Tiger et al., 2009	Jack	2	2	2	2	2	2	2	2	2	2
Torres-Viso et al., 2008	Amy	1	2	2	1	2	2	2	2	2	2
Vollmer et al., 1999	Dale	2	2	2	2	2	2	2	2	2	2

Note. N/A= not applicable.

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Britton et al., 2000	Rob	Male Adult ID only	Tangible	NCR plus EXT	Time based	Increasing intervals	None	FT-45 s	FT-5 min	Yes
Britton et al., 2000	Todd	Male Child ASD only	Attention	NCR plus EXT	Time based	Increasing intervals	None	FT-1 min 40 s	FT-5 min	Yes
Crauthers et al., 2015	Alison	Female Young child ASD only	Escape	FCT plus EXT	Response based	Multiple schedule	DRO embedded during the schedule thinning phase	1 min	5 min	Yes
Davis et al., 2012	Eli	Male Adult Multiple	Escape	FCT	Response based	Delay-to-reinforcement	None	30 s	1 min	Yes
Falcomata, Roane, et al., 2012	Steph	Female Child ASD only	Other	FCT plus EXT	Response based	Delay-to-reinforcement with demand fading embedded into the delay	Modification 1: access to leisure materials Modification 2: work tasks presented	5 min	10 min	No
	Mike	Male Child ASD only	Other	FCT plus EXT	Response based	Delay-to-reinforcement with demand fading embedded into the delay	Modification 1: access to leisure materials Modification 2: work tasks presented	5 min	10 min	No
Falcomata, White, et al., 2012	Danny	Male Child ASD only	Multiple	FCT plus EXT	Response based	Demand fading	None	0 math problems	20 math problems	No
Falcomata, Muething, et al., 2013	Alonzo (experiment 2)	Male Child ASD only	Multiple	FCT plus EXT	Response based	Delay-to-reinforcement	None	5 s	5 min	No

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Falcomata, Muething, et al., 2013	Joe	Male Adolescent ASD only	Multiple	FCT plus EXT	Response based	Delay-to-reinforcement	None	5 s	5 min	No
	Alonzo (experiment 3)	Male Child ASD only	Multiple	FCT plus EXT	Response based	Demand fading	None	FI- 5 min	VI-5 min	No
Falcomata, Roane, Feeney, et al., 2010	John	Male Young child ASD only	Other	FCT plus EXT	Response based	Delay-to-reinforcement and response restriction	None	Setting 1: 2 s Setting 2: 2 s	Setting 1: 10 s Setting 2: 10 min	No
Fisher et al., 2015	Jacob	Male Child Other DD	Tangible	FCT plus EXT	Response based	Multiple schedule	Attention provided during the EXT component of the multiple schedule	S^{Δ} 1 min	S^{Δ} 5 min	No
Fisher et al., 1996	Mat	Male Young child Multiple	Attention	NCR plus EXT	Time based	Increasing intervals	None	FT-40 s	FT-5 min	No
Fritz et al., 2017	Charley	Male Child ASD only	Tangible	NCR plus EXT	Time based	Other	None	3 reinforcers per min	0.2 reinforcers per min	No
	Gilbert	Male Child ASD only	Tangible	NCR plus EXT	Time based	Other	None	3 reinforcers per min	0.2 reinforcers per min	No
	Dyson	Male Child ASD only	Tangible	NCR plus EXT	Time based	Other	None	3 reinforcers per min	0.2 reinforcers per min	No
Fyffe et al., 2004	Matt	Male Child Multiple	Attention	FCT plus EXT	Response based	Response restriction	None	5 s	5 min	No

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Hagopian et al., 1994	Wanda	Female Young child Multiple	Attention	NCR	Time based	Increasing intervals	None	FT-10 s	FT-5 min	No
	Lynn	Female Young child Multiple	Attention	NCR	Time based	Increasing intervals	None	FT-10 s	FT-5 min	No
Hagopian et al., 1998	Case 19	Gender not reported Child Multiple	Multiple	FCT plus EXT	Response based	Demand fading	Punishment procedure added	FR-1	VR-7	No
Hagopian et al., 2000	Alex	Male Child Multiple	Tangible	NCR	Time based	Increasing intervals	DRO and EXT added	FT-30 s	NCR only: FT-35 s NCR, DRO, and EXT: FT-7 min	No
	Emily	Female Young child Multiple	Tangible	Treatment package	Time based	Increasing intervals	None	FT-25 s	FT- 1 min, 55 s	No
Hagopian et al., 2004	Brent	Male Child Multiple	Multiple	FCT plus EXT	Response based	Multiple schedule	None	S^{Δ} 0 s	S^{Δ} 9 min	No
	Jason	Male Adolescent Multiple	Tangible	NCR plus EXT	Time based	Increasing intervals	None	FT-15 s	FT-4 min	No
Hagopian et al., 2005	Stephen	Male Adolescent Multiple	Attention	FCT, FCT plus EXT	Response based	Delay-to-reinforcement	Competing stimuli	5 s	FCT plus EXT only: 15 s Competing stimuli: 4 min	No
	James	Male Adolescent Multiple	Tangible	FCT, FCT plus EXT	Response based	Delay-to-reinforcement	Competing stimuli	15 s	5 min	No

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Hagopian et al., 2005	Matt	Male Child Multiple	Multiple	FCT, FCT plus EXT	Response-based	Delay-to-reinforcement	Competing stimuli	1 min	9 min	No
Hagopian et al., 2007	Kelly	Female Adolescent Multiple	Other	Treatment package	Response based Time based	Multiple schedule	None	S^{Δ} 30 s	S^{Δ} 9 min	No
Hammond et al., 2011	Alex	Male Adolescent Multiple	Tangible	DRO	Time based	Increasing intervals	None	30 s	5 min	No
Hanley et al., 2001	Karen	Female Adult ID only	Tangible	FCT plus EXT	Response based	Delay-to-reinforcement Fixed interval Multiple schedule	None	Delay-to-reinforcement: 1 s Fixed interval: 1 s Multiple schedule: S^{Δ} 15 s	Delay-to-reinforcement: 25 s Fixed interval: 58 s Multiple schedule: S^{Δ} 4 min	No
Hanley et al., 2001	Julie	Female Adult Multiple	Attention	FCT plus EXT	Response based	Mixed schedule Multiple schedule	None	Mixed schedule: S^{Δ} 15 s Multiple schedule: S^{Δ} 15 s	Mixed schedule: S^{Δ} 2 min Multiple schedule: S^{Δ} 4 min	No
Ing et al., 2011	Sara	Female Child ASD only	Automatic	NCR plus EXT	Time based	Increasing intervals	None	Continuous	FT-30 s	No
Kern et al., 2006	Orlando	Male Child ID only	Tangible	DRA	Response based	Demand fading and delay-to-reinforcement	None	Demand fading: FR-1 Delay-to-reinforcement: 10 s	Demand fading: FR-3 Delay-to-reinforcement: 20 s	No

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Kern et al., 2006	Matthew	Male Adult Multiple	Tangible	Treatment package	Response based	Delay-to-reinforcement	None	5 s	40 s	No
Lambert et al., 2019	Leah	Female Child Other DD	Tangible	Treatment package	Time based Response based	Delay-to-reinforcement	None	3 s	36 s	Yes
Lalli et al., 1997	Donny	Male Young child Other DD	Tangible	NCR plus EXT	Time based	Increasing intervals	None	1 min 30 s	10 min	Yes
Lalli et al., 1998	Val	Female Child Multiple	Attention	NCR plus EXT	Time based	Increasing intervals	None	FT-30 s	FT-5 min	No
Peery & Fisher, 2001	Ann	Female Adolescent Multiple	Escape	Treatment package	Response based	Demand fading	None	FR-1	FR-20	No
Piazza et al., 1997	Ray	Male Adolescent Multiple	Multiple	DRO	Time based	Increasing intervals	None	50 s	5 min	No
	Ty	Male Young child Multiple	Attention	Treatment package	Response based	Demand fading	None	5 s of appropriate walking	30 s of appropriate walking	No
Rasmussen & O'Neill, 2006	Josh	Male Adolescent Other DD	Attention	NCR plus EXT	Time based	Increasing intervals	None	FT-10 s	FT- 1 min 15 s	No
Richman et al., 1999	Don	Male Child Multiple	Attention	NCR plus EXT	Time based	Increasing intervals	DRA	Continuous	FT-30 s	No
Roane et al., 2007	Fred (exp 2)			DRO plus EXT	Time based	Increasing intervals	Magnitude of reinforcer adjusted	10 s	3 min	No

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Rooker et al., 2013	Application 36	Gender not reported Child ID only	Escape	FCT plus EXT	Response based	Multiple schedule	NCR	S^{Δ} 1 min	S^{Δ} 10 min	No
	Application 6	Gender not reported Child ID only	Attention	FCT plus EXT	Response based	Multiple schedule	NCR and punishment	S^{Δ} 30 s	S^{Δ} 15 min	No
Simmons et al., 2003	Alicia	Female Adult Diagnosis not specified	Automatic	NCR plus EXT	Time based	Increasing intervals	None	10 s	10 min	No
Slocum et al., 2018	Clancy	Male Child Diagnosis not specified	Tangible	NCR plus extinction	Time based	Increasing intervals in which access to the functional reinforcer was unavailable	Verbal warning	10 s	2 min	No
	Korey	Male Young child Diagnosis not specified	Attention	NCR plus extinction	Time based	Increasing intervals in which access to the functional reinforcer was unavailable	None	10 s	2 min	No
	Reginald	Male Adolescent Diagnosis not specified	Tangible	NCR plus extinction	Time based	Increasing intervals in which access to the functional reinforcer was unavailable	None	10 s	2 min	No

Table A9

Results of the descriptive evaluation

	Participant	Demographic Information	Function of Challenging Behavior	Intervention Type	Reinforcement Schedule	Schedule Thinning Procedures	Procedural Modifications	Initial Schedule	Terminal Schedule	Treatment Fidelity
Tiger et al., 2009	Jack	Male Adult ASD only	Automatic	DRO	Time based	Increasing intervals	Self-monitoring	5 min	15 min	Yes
Torres-Viso et al., 2008	Amy	Female Adolescent Multiple	Other	FCT plus EXT	Response based	Multiple schedule	DRO	S^{Δ} 50 s	S^{Δ} 15 min	No
Vollmer et al., 1999	Dale	Male Child Multiple	Tangible	FCT plus EXT	Time based	Delay-to-reinforcement	None	0 s	15 s	No

Note. ASD= autism spectrum disorder; DRA= differential reinforcement of alternative behavior; DRO= differential reinforcement of other behavior; EXT= extinction; FCT= functional communication training; NCR= noncontingent reinforcement.

Table A10

Interobserver agreement data

	Baseline	FCT	Delay-to-reinforcement
Challenging behavior			
Nadine	100	94 (range= 83-100)	95 (range= 66-100)
Reba	89 (range= 67-100)	100	86 (range= 50-100)
Bryce	83 (range= 66-100)	86 (range= 66-100)	100
Communication			
Nadine	100	100	91 (range= 66-100)
Reba	100	100	96 (range= 75-100)
Bryce	100	100	100
Item engagement			
Nadine	N/A	N/A	95 (range= 44-100)
Reba	N/A	N/A	92 (range= 63-100)
Bryce	N/A	N/A	100

Note. Numbers represent the percentage agreement between primary and secondary raters. FCT= functional communication training; N/A= not applicable.

Table A11

Treatment fidelity

	Baseline	FCT	Delay-to-reinforcement
Nadine	100%	100%	100%
Reba	100%	100%	100%
Bryce	100%	100%	100%

Note. FCT= functional communication training.

Table A12

Instances of resurgence during the delay-to-reinforcement for Nadine

Delay interval	Choice	No choice	No activity
0 s	N/A	N/A	N/A
5 s	N/A	N/A	N/A
8 s	N/A	N/A	N/A
15 s	N/A	N/A	N/A
30 s	N/A	1 (range= 0-1)	N/A
60 s	N/A	N/A	N/A
90 s	N/A	N/A	N/A
120 s	N/A	N/A	N/A
150 s	N/A	N/A	N/A
180 s	N/A	N/A	N/A

Note. N/A= not applicable.

Table A13

Average number of communicative responses during the delay-to-reinforcement for Nadine

Delay interval	Choice	No choice	No activity
0 s	0	0	0
5 s	0.5 (range= 0-1)	1 (range= 0-1)	0.4 (range= 0-1)
8 s	0	0	2 (range= 0-6)
15 s	0	0	3.5 (range= 2-4)
30 s	0	0	5 (range= 2-9)
60 s	0	N/A	N/A
90 s	0	N/A	N/A
120 s	0	N/A	N/A
150 s	0	N/A	N/A
180 s	0	N/A	N/A

Note. N/A= not applicable.

Table A14

Average percentage of 10 s intervals with item engagement per session for Nadine

Delay interval	Choice	No choice
0 s	N/A	N/A
5 s	25 (range= 17-27)	22 (range= 20-23)
8 s	27	27 (range= 23-27)
15 s	31 (range= 20-37)	32 (range= 20-40)
30 s	47	43.4 (range= 40-47)
60 s	53 (range= 47-60)	47
90 s	63 (range= 57-67)	N/A
120 s	67 (range= 57-70)	N/A
150 s	70	N/A
180 s	63	N/A

Note. N/A= not applicable.

Table A15

Instances of resurgence during the delay-to-reinforcement for Reba

Delay interval	Choice	No choice	No activity
0 s	N/A	N/A	N/A
5 s	1 (range= 0-1)	2 (range= 0-1)	2 (range= 0-1)
8 s	1 (range= 0-1)	N/A	N/A
15 s	N/A	N/A	N/A
30 s	N/A	N/A	1 (range= 0-1)
60 s	N/A	N/A	N/A
90 s	N/A	N/A	N/A
120 s	N/A	N/A	N/A

Note. N/A= not applicable.

Table A16

Average number of communicative responses during the delay-to-reinforcement for Reba

Delay interval	Choice	No choice	No activity
0 s	0	0	0
5 s	1.4 (range= 0-3)	2 (range= 0-8)	4 (range= 0-13)
8 s	0.42 (range= 0-1)	4 (range= 0-8)	13 (range= 6-18)
15 s	1.8 (range= 0-7)	8.3 (range= 10-15)	5.6 (range= 7-10)
30 s	0	5.6 (range= 1-18)	11.25 (range= 5-19)
60 s	4.6 (range= 0-10)	N/A	N/A
90 s	4.6 (range= 0-11)	N/A	N/A
120 s	14.5 (range= 6-23)	N/A	N/A

Note. N/A= not applicable.

Table A17

Average percentage of 10 s intervals with item engagement per session for Reba

Delay interval	Choice	No choice
0 s	N/A	N/A
5 s	18 (range=17-27)	10 (range= 0-27)
8 s	24 (range=13-33)	10 (range= 0-20)
15 s	28 (range=30-33)	23 (range=10-37)
30 s	45 (range= 40-50)	9 (range= 0-33)
60 s	35 (range= 23-60)	N/A
90 s	57 (range= 47-63)	N/A
120 s	75 (range= 70-90)	N/A

Note. N/A= not applicable.

Table A18

Instances of resurgence during the delay-to-reinforcement for Bryce

Delay interval	Choice	No choice	No activity
0 s	N/A	N/A	N/A
5 s	N/A	N/A	N/A
8 s	N/A	N/A	N/A
15 s	N/A	N/A	N/A
30 s	N/A	N/A	N/A
60 s	N/A	N/A	N/A
90 s	N/A	N/A	N/A
120 s	N/A	N/A	N/A
180 s	N/A	N/A	N/A

Note. N/A= not applicable.

Table A19

Average number of communicative responses per session during the delay for Bryce

Delay interval	Choice	No choice	No activity
0 s	N/A	N/A	N/A
5 s	N/A	N/A	1 (rang= 0-1)
8 s	N/A	N/A	N/A
15 s	1 (range= 0-1)	N/A	N/A
30 s	N/A	N/A	N/A
60 s	N/A	N/A	N/A
90 s	N/A	N/A	N/A
120 s	N/A	N/A	N/A
150 s	N/A	N/A	N/A
180 s	N/A	N/A	N/A

Note. N/A= not applicable.

Table A20

Average percentage of 10 s intervals with item engagement per session for Bryce

Delay interval	Choice	No choice
0 s	N/A	N/A
5 s	20	3 (rang= 0-3)
8 s	19	15
15 s	20	N/A
30 s	N/A	N/A
60 s	N/A	N/A
90 s	N/A	N/A
120 s	N/A	N/A
180 s	N/A	N/A

Note. N/A= not applicable.

APPENDIX B. FIGURES

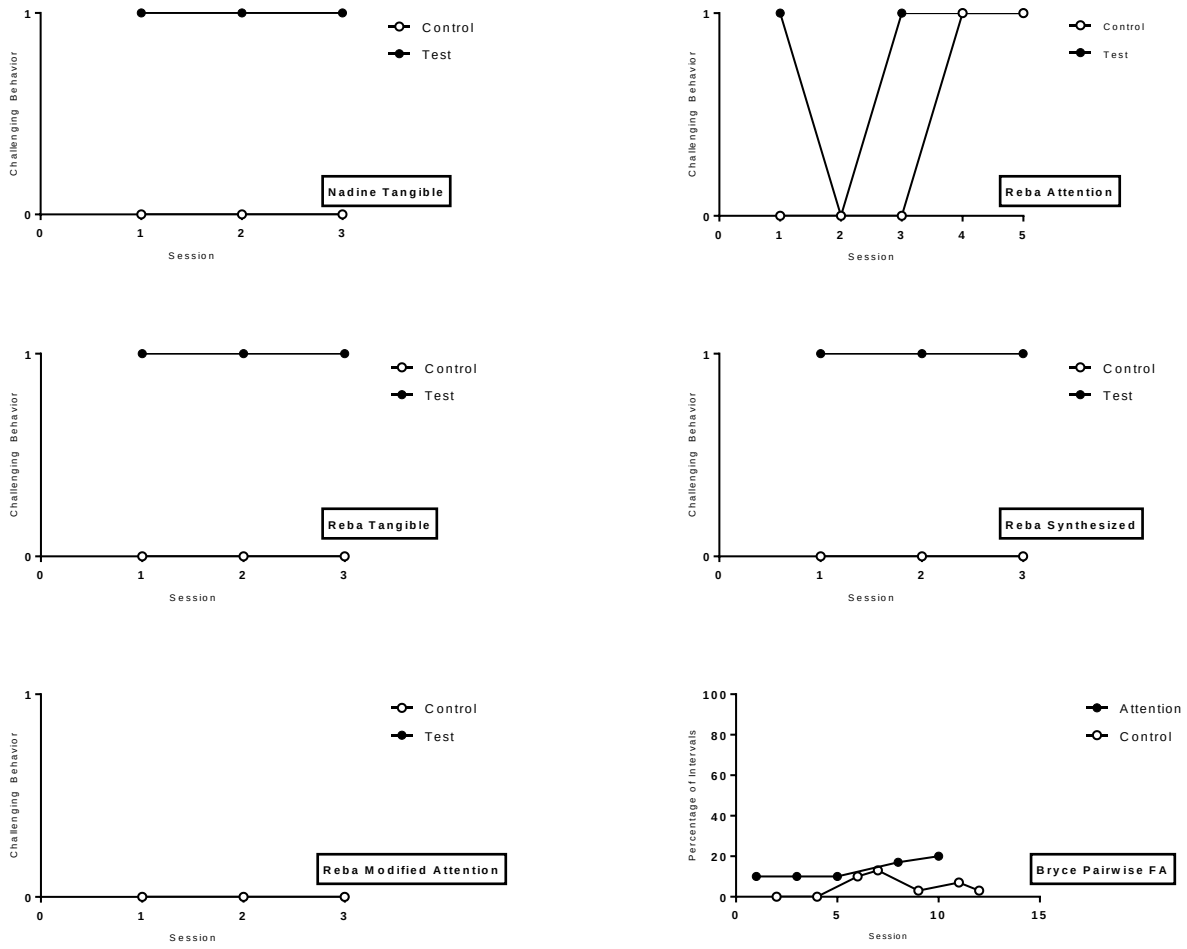

Figure B1. Results of the functional analyses for each participant. Results of the tangible condition of the TBFA for Nadine are displayed on the top left panel. Results of the attention condition of the TBFA for Reba are displayed on the top right panel. The left middle panel displays the results of the tangible condition of the TBFA for Reba. The middle right panel depicts the results of the synthesized condition of the TBFA for Reba. Results of the modified attention condition of the TBFA for Reba are displayed in the bottom left panel. The results of the PFA for Bryce are displayed in the bottom right panel.

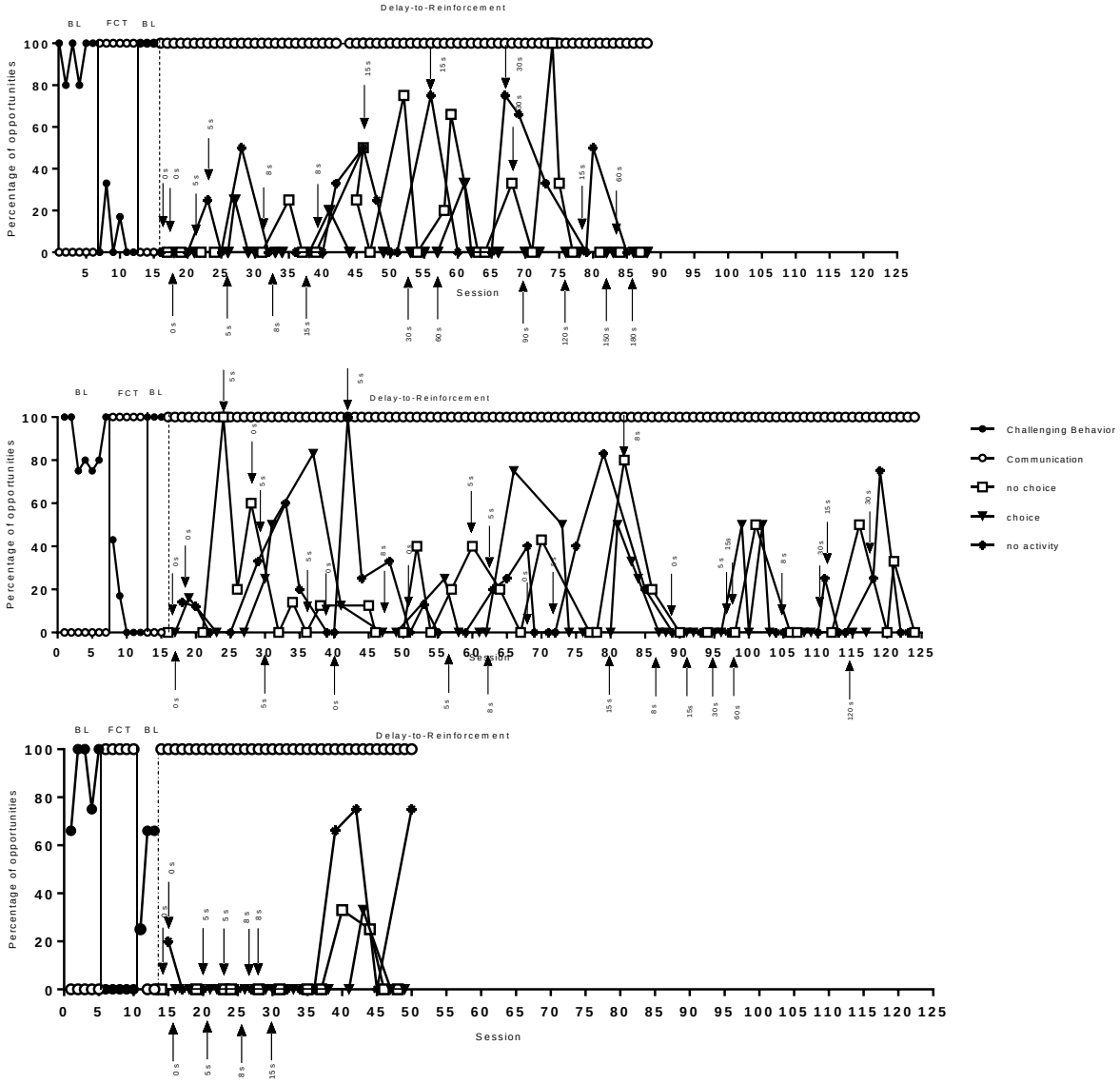

Figure B2. Results of the FCT and delay-to-reinforcement evaluation for all participants.

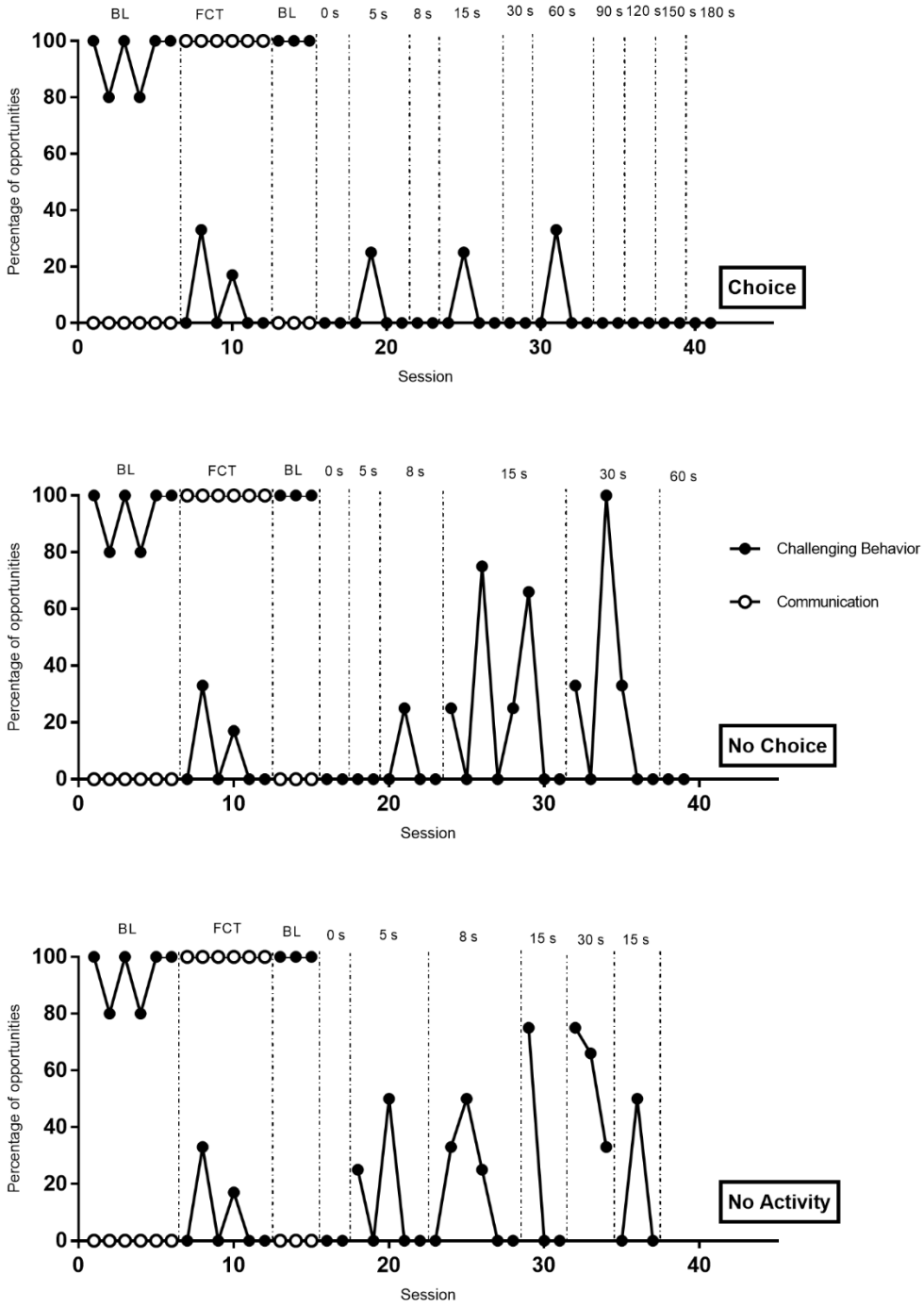

Figure B3. Results of the delay-to-reinforcement across individual conditions for Nadine.

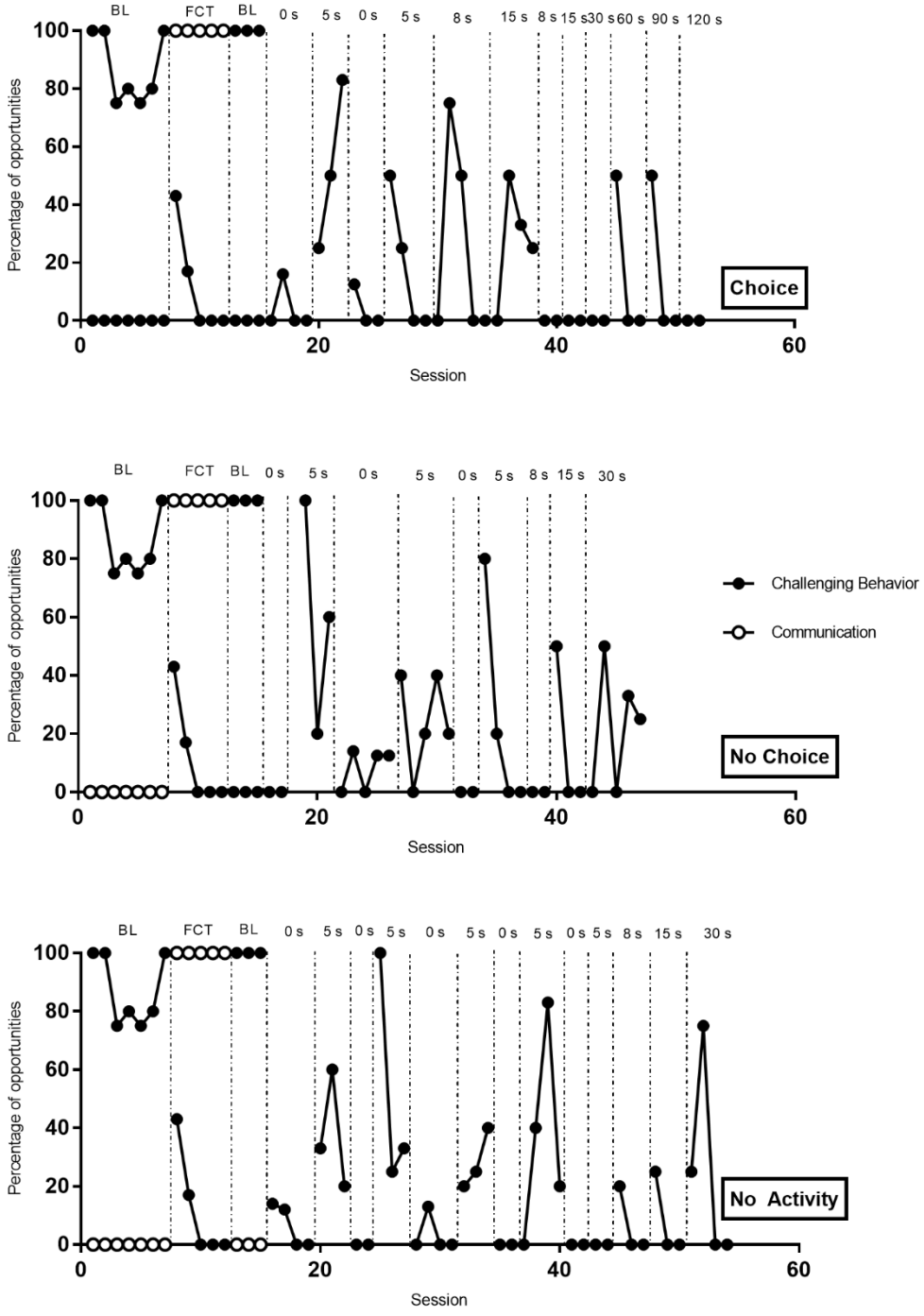

Figure B4. Results of the delay-to-reinforcement across individual conditions for Reba.

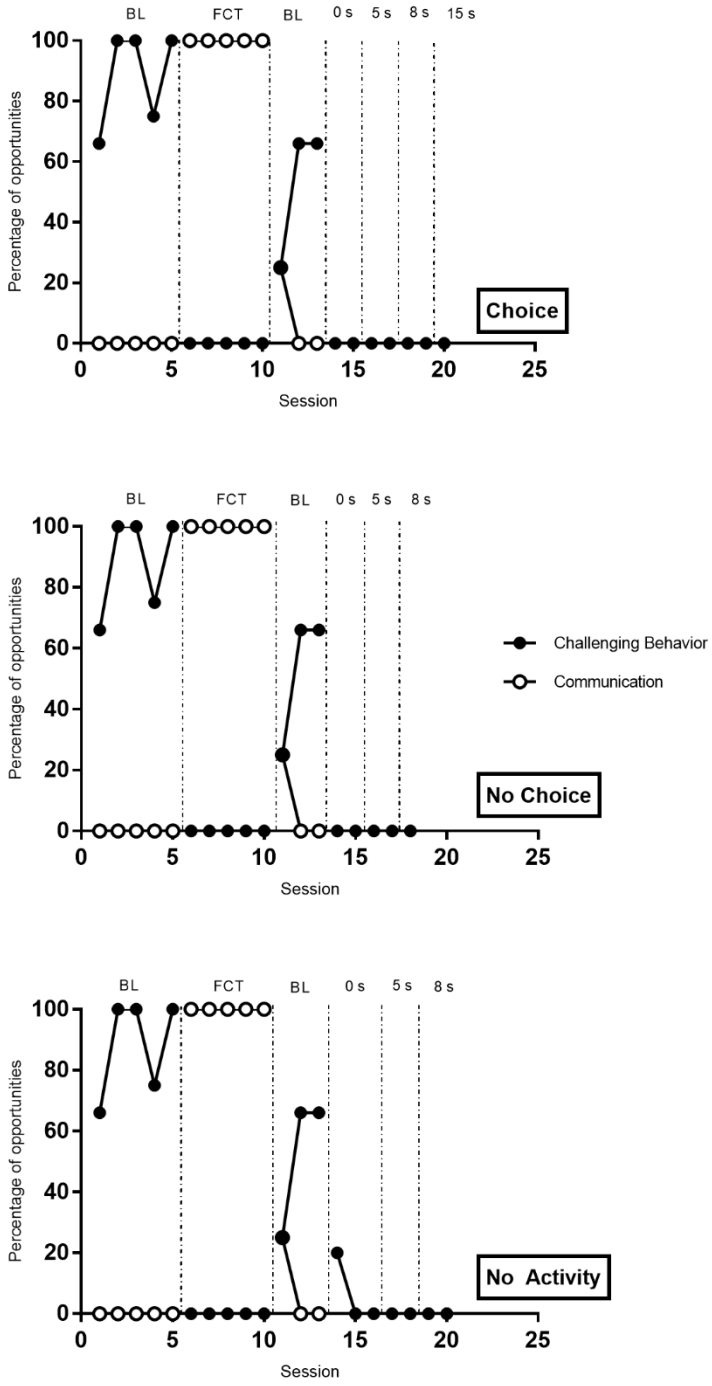

Figure B5. Results of the delay-to-reinforcement across individual conditions for Bryce.

APPENDIX C. DATA COLLECTION SHEETS

ABC Data Collection

Reviewer: Consumer: Date: Primary/Rely Observation Time:

Target Behaviors:

Setting:

Time	Setting	Antecedent	Behavior	Consequence	Notes

TBFA Data Collection Sheet
Consumer Behavior

Reviewer: Consumer: Date: Primary/Rely
Condition:

Target Behavior(s):
Directions

- Mark + if challenging behavior occurs
- Mark – if no challenging behavior occurs

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10
Control										
Test										

MSWO Preference Assessment Sheet

Reviewer: Consumer: Date: Primary/Rely

Session Number:

Directions

- Record the item the consumer selects under “item”
- Record the rank of each item

Item	Rank	Notes

Free Operant Preference Assessment

Reviewer: Consumer: Date: Primary/Rely

Session Number:

Directions

- Record the item the consumer selects under “item”
- Record the number of minutes and seconds the consumer interacts with the item under “duration”

Item	Duration	Notes
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	
	_____ min _____ sec	

FCT Data Collection Sheet

Reviewer: Consumer: Date: Primary/Rely
Baseline/Intervention/Schedule Thinning Session Number: Delay Interval:

Directions:
- Circle CB or IE if it occurred at any point during the interval
- Tally the number of appropriate communicative responses per session

	Minute 1	**Minute 2**	**Minute 3**	**Minute 4**	**Minute 5**
0-10 s	CB IE	CB IE	CB IE	CB IE	CB IE
11-20 s	CB IE	CB IE	CB IE	CB IE	CB IE
21-30 s	CB IE	CB IE	CB IE	CB IE	CB IE
31-40 s	CB IE	CB IE	CB IE	CB IE	CB IE
41-50 s	CB IE	CB IE	CB IE	CB IE	CB IE
51-60 s	CB IE	CB IE	CB IE	CB IE	CB IE
Percentage of Intervals with CB					
Percentage of intervals with IE					
Frequency of appropriate communication					

% Opp										
CB										
Com										

Communication Training

Reviewer: Consumer: Date: Primary/Rely

Phase:

Directions

- + correct
- - incorrect

Session Number	Trial 1	Trial 2	Trial 3	% Trials Correct

Baseline Fidelity

Reviewer: Interventionist: Consumer: Date:

Primary/Rely Session Number:

Directions

- Mark + for correct
- Mark – for incorrect
- Mark N/A if not applicable

Interventionist Behavior	Score
Interventionist provides access to the putative reinforcer for 30 s	
Interventionist says, “Ok, time to do something different,” and removes the reinforcer	
Contingent on challenging behavior, interventionist reinstates reinforcer for 30 s	
If no challenging behavior occurs, interventionist continues to withhold reinforcer	
Interventionist ignores all instances of appropriate communication	
Total Correct	
Percentage Correct	

Functional Communication Training Fidelity

Reviewer: Interventionist: Consumer: Date:

Primary/Rely Session Number:

Directions

- Mark + for correct
- Mark – for incorrect
- Mark N/A if not applicable

Interventionist Behavior	Score
Interventionist provides access to the putative reinforcer for 30 s	
Interventionist says, “Ok, time to do something different,” and removes the reinforcer	
Contingent on appropriate responding, interventionist reinstates reinforcer for 30 s	
Contingent on appropriate responding, interventionist provides specific praise	
If consumer does not respond within 5 s, interventionist prompts appropriate communication using MTL prompting	
If consumer engages in challenging behavior upon removal of the reinforcer, interventionist waits for a 5 s break in challenging behavior before prompting appropriate communication	
Interventionist ignores all instances of challenging behavior	
Total Correct	
Percentage Correct	

Schedule Thinning: No Choice Fidelity

Reviewer: Interventionist: Consumer: Date:

Primary/Rely Session Number:

Directions

- Mark + for correct
- Mark – for incorrect
- Mark N/A if not applicable

Interventionist Behavior	Score
Interventionist provides access to the putative reinforcer for 30 s	
Interventionist says, “Ok, time to do something different,” and removes the reinforcer	
Interventionist says, “Good asking, but you can’t use this right now. You have to wait ___ minutes/seconds. While you wait you can use _____” and provides consumer with alternate activity	
After predetermined delay elapses, interventionist reinstates putative reinforcer for 30 s	
Interventionist ignores all instances of appropriate communication	
Total Correct	
Percentage Correct	

Schedule Thinning: Choice Fidelity

Reviewer: Interventionist: Consumer: Date:

Primary/Rely Session Number:

Directions

- Mark + for correct
- Mark – for incorrect
- Mark N/A if not applicable

Interventionist Behavior	Score
Interventionist provides access to the putative reinforcer for 30 s	
Interventionist says, “Ok, time to do something different,” and removes the reinforcer	
Interventionist says, “Good asking, but you can’t use this right now. You have to wait __ minutes/seconds. While you wait you can use one of these,” offers two alternate activities, and provides consumer with alternate activity	
After predetermined delay elapses, interventionist reinstates putative reinforcer for 30 s	
If consumer does not make a choice within 10 s, interventionist re-presents items and says, “Remember, you can choose one.”	
Interventionist ignores all instances of appropriate communication	
Total Correct	
Percentage Correct	