Evaluating the Effects of Omission Errors during Token Economies

by

Odessa Luna

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 5, 2017

Keywords: token economy, tokens, treatment integrity, omission errors, differential reinforcement of alternative behavior

Approved by

Sacha T. Pence, Chair, Assistant Professor of Psychology
John Rapp, Professor of Psychology
Chris Newland, Professor of Psychology

Abstract

A token economy is a differential-reinforcement of alternative-behavior (DRA) procedure that can be used to decrease problem behavior and increase appropriate behavior and involves the delivery of tokens on a specified schedule following appropriate behavior. When implementing token economies, the behavior-change agent can engage in treatment integrity errors, including omission errors (i.e., failing to perform a component of the protocol). During a token economy, omission errors occur when a practitioner fails to deliver a token following appropriate behavior. The purpose of this study was to evaluate the extent to which differing levels of omission errors (100%, 80%, 60%, 40%, 20%, and 0%) influenced the efficacy of token economies across fixed-ratio and variable-ratio schedules of reinforcement on appropriate and problem behavior. As levels of integrity decreased, participants engaged in increases in problem behavior and decreases in appropriate behavior. However, this pattern was not replicated within subject for one participant. The outcomes suggest that omission errors within a token economy can be detrimental to treatment outcomes for individuals who engage in escape-maintained aggression.

Table of Contents

Abstract ii

List of Tables iv

List of Figures v

Chapter 1: Literature Review 1

Functional Behavior Assessment 1

Differential Reinforcement 2

Token Economies 4

Treatment Integrity 13

Chapter 2: Experimental Procedures 20

Results 31

Discussion 38

References 48

List of Tables

Table 1 63

Table 2 64

List of Figures

Figure 1 65

Figure 2 66

Figure 3 67

Figure 4 68

Figure 5 69

Chapter 1: Literature Review

Autism spectrum disorder (ASD) is a developmental disability that is associated with social, communication, and behavioral challenges (American Psychiatric Association, 2013). Childhood diagnosis of ASD is on the rise in the United States, with close to 1 in 50 children receiving a diagnosis every year (Blumberg et al., 2013). Moreover, 1 in 100 children are diagnosed with an intellectual disability and 1 in 6 children are diagnosed with a developmental disability. Children with disabilities are likely to engage in moderate to severe problem behavior, including self–injurious behavior, aggression, and property destruction (Boyle et al., 2011; Lecavalier, 2006; Maulik, Mascarenhas, Mathers, Dua, & Saxena, 2010). Because of the detrimental effects problem behavior has on the individual's quality of life, reduction of problem behavior and the strengthening of adaptive behavior with individuals with intellectual and developmental disabilities are of top priority. In fact, individuals who engage in problem behavior are at increased risk for developing deficits in academic, communication, and social skills (Olson & Hoza, 1993).

Functional Behavior Assessment

Interventions based on a functional behavior assessment (FBA) are more effective than non–function–based interventions to increase appropriate behavior and decrease problem behavior (Beavers, Iwata, & Lerman, 2013; Carr & Durand, 1985; Hanley, Iwata, & McCord, 2003; Iwata, Pace, Cowdery, & Miltenberger, 1994). The goal of a FBA is to identify the environmental variables, including antecedent events that occasion problem behavior and consequence events that maintain problem behavior. Under the Individuals with Disabilities Education Improvement Act (IDEA), FBAs are mandated for children with disabilities who engage in problem behavior at school that impedes the child's learning or the learning of others

(Individuals with Disabilities Education Act, 2004). The functional assessment process can consist of indirect assessments, interviews, descriptive assessments, and functional analyses. Although indirect assessments, interviews, and descriptive assessments provide information about problem behavior (e.g., temporal patterns of problem behavior) and the environmental events surrounding problem behavior, a functional analysis is necessary to determine the function of problem behavior (Lerman & Iwata, 1993; Pence, Roscoe, Bourret, & Ahearn, 2009).

A functional analysis involves the systematic manipulation of antecedent and consequence events to determine the effect these variables have on problem behavior (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). Individuals are exposed to environmentally analogous conditions to determine behavioral sensitivity to antecedent arrangements and putative reinforcers. Elevated responding in one or more test conditions compared to a control condition (e.g., demands are not presented and preferred items and adult attention are freely available) allows for conclusions to be made about the function of behavior. Common test conditions for social functions include attention, escape from demands, and access to tangible items. The test conditions are compared to a control condition. Practitioners and researchers use the results of the functional analysis (i.e., identified function of problem behavior) to inform behavior intervention plans. For example, if a child engages in aggression to escape work, a practitioner may design an intervention where the staff provides breaks contingent on appropriate behavior and reinforces appropriate break requests

Differential Reinforcement

Differential reinforcement involves delivering the reinforcer contingent upon either the absence of problem behavior or the occurrence of an alternative behavior as well as the discontinuation of the functional reinforcer maintaining the problem behavior (Beavers et al., 2013; Carr & Durand, 1985; Hanley et al., 2003; Lerman & Iwata, 1996). One procedural variation is differential reinforcement of alternative behavior (DRA). Typically, a DRA intervention entails withholding reinforcement for problem behavior (i.e., placing the behavior on extinction) and delivering the functional reinforcer contingent on an alternative, replacement behavior (Hanley, Iwata, & Thompson, 2001; Mace et al., 2010; Petchser, Ray, & Bailey, 2009; Vollmer, Roane, Ringdahl, & Marcus, 1999). For example, if a child engages in self–injurious behavior to gain access to a toy, a DRA intervention would involve the practitioner providing the toy following appropriate behavior (e.g., asking for the toy or complying with demands). If the child engages in self–injurious behavior, the practitioner would not provide the toy. In general, DRA procedures are one of the most frequently used interventions to decrease problem behavior with individuals with ASD and developmental disabilities (Petscher et al., 2009). In addition, DRA interventions meet the Division 12 Task Force on the Promotion and Dissemination of Psychological Procedures criteria for empirically supported interventions (Lennox, Miltenberger, Spengler, & Erfanian, 1988; Petscher et al., 2009).

The alternative behavior targeted as part of the DRA intervention can be a communicative response (e.g., requests for attention, breaks, or access to tangible items) or replacement behavior (e.g., hand raising, compliance with adult directives, sharing toys; Petscher et al., 2009; Piazza, Moes, & Fisher, 1996). When possible, the targeted appropriate response should be in the learner's current repertoire, require less effort than the problem behavior, and have a high probability of reinforcement in the natural environment (Flynn & Lo, 2016; Friman, 1990; Hanley et al., 2001; Miltenberger, 2012). Appropriate responses are typically reinforced on a dense schedule of reinforcement during initial implementation of the intervention prior to thinning to leaner schedules of reinforcement (Dixon, Benedict & Larson, 2001; Mitteer, Romani, Greer, & Fisher, 2015; Piazza et al., 1996; Roane, Lerman & Vorndran, 2001; Vollmer et al., 1999).

One procedural variation of DRA is functional communication training (FCT), which involves teaching the individual to request for the functional reinforcer that maintains problem behavior appropriately. For example, if an individual engages in aggression maintained by escape from demands, the practitioner would teach the individual to request, “I need a break” to replace the aggression. Aggression would no longer produce escape from demands; rather the alternative, communicative behavior would be reinforced with escape (Carr & Durand, 1985). Typically, the practitioner needs to reinforce the communicative response on a dense schedule of reinforcement during acquisition and early phases of the FCT intervention (Fisher, Kuhn, & Thompson, 1998; Hagopian, Kuhn, Long, & Rush, 2005; Jarmolowicz, DeLeon, & Kuhn, 2013; Tiger, Hanley, & Bruzek, 2008). However, it is unlikely that a dense schedule will be maintained in the natural environment if the communicative response occurs at high rates or in situations when it cannot be reinforced (Geiger, Carr, & LeBlanc, 2010; Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997). As a result, levels of problem behavior may increase as the association between the communicative response and reinforcer is extinguished over time. Similarly, it is impractical for individuals with escape–maintained behavior to avoid activities such as schoolwork or independent–living tasks (e.g., showering, brushing teeth) continuously by requesting a break. Therefore, practitioners may use other differential reinforcement procedures to increase compliance with demands.

Token Economies

Within the home, school, and community, behavior–change agents often want the individual to engage in a specific competing response (e.g., compliance) to earn the functional reinforcer. For example, a teacher could request that the child completes a worksheet with 10 problems prior to receiving a break. A token economy can be used to help increase a specific response and help to bridge the delay between appropriate behavior and the functional reinforcer. Within a token economy, tokens are delivered contingent on the occurrence of appropriate behavior. Following a designated number of tokens, the individual can exchange the tokens for a backup reinforcer, such as access to preferred stimuli (e.g., edibles, toys), an activity (e.g., computer, video games), or a privilege (e.g., homework pass, walk with a teacher). The components of a token economy intervention package include token identification, target behavior selection, schedule of reinforcement, backup reinforcer identification, exchange rate, and staff training (Gillis & Pence, 2015).

Token identification. One of the first steps when implementing a token economy is to identify the stimuli to use for tokens. It is helpful if the tokens are durable, inexpensive, easy to make, and unique. Tokens can include poker chips (Sran & Borrero, 2010), tallies (Breyer & Allen, 1975; Phillips, 1968), stars (Conyers et al., 2004; Tarbox, Ghezzi, & Wilson, 2002), laminated icons (Kahng, Boscoe, & Bryne, 2003; Fiske et al., 2015), pennies (Cowdery, Iwata, & Pace, 1990), and imitation dollar bills (LeBlanc, Hagopian, & Maglieri, 2000).

Client preferences can be considered and the practitioner may decide to individualize tokens for the client (Gillis & Pence, 2015). For example, if a client prefers characters from a movie, laminated pictures of the characters can be used as tokens. For some clients, using their interests when designing the tokens can influence levels of appropriate behavior compared to other stimuli that are arbitrarily selected by the researcher (Carnett et al., 2014; Charlop–Christy & Haymes, 1998).

Generalized conditioned reinforcement and token pairing. The token is a previously neutral stimulus that becomes a generalized conditioned reinforcer due to being associated with other reinforcing stimuli (e.g., unconditioned and conditioned reinforcers). Unconditioned reinforcers are stimuli that do not need to be learned throughout the lifespan and can include edibles and liquids. Conditioned reinforcers require a learning history and the stimulus needs to be paired with established reinforcers to function as a reinforcer. Conditioned reinforcers can include tangible items (e.g., iPad, toys), praise, and breaks from demands. Within a token economy, it is hypothesized that tokens become associated with many other stimuli and therefore, their effectiveness does not depend upon a specific backup reinforcer or establishing operation (Gillis & Pence, 2015; Massetti & Fabiano, 2005).

Ivy, Neef, Meindl, and Miller (2016) conducted a preliminary evaluation on how motivating operations (pre-session access and restriction) influence the effectiveness of a token, unconditioned item (edibles), and conditioned stimuli (tangibles) as a reinforcer during a progressive-ratio reinforcer assessment. During pre-session access, participants had unlimited free access to preferred items (edibles and preferred items) that ended when the participants said "done" or did not consume or interact with the items for 30 s. During pre-session restriction, the conditioned and unconditioned preferred items were restricted for 24 hours prior to the evaluation. Pre-session access resulted in lower responding for both participants across all conditions (token, conditioned, and unconditioned reinforcement). Pre-session restriction produced increased responding, regardless if conditioned or unconditioned stimuli were restricted. These results suggest that tokens may be sensitive to motivating operations (e.g., pre-session restriction), and future studies should continue to evaluate the role of motivating operations and their influence on token economy effectiveness.

Despite general recommendations to establish tokens as reinforcers, there are research gaps in evaluating the necessary token training procedures prior to token economy implementation. LeBlanc et al. (2000) conducted token training with an individual with an intellectual disability who engaged in inappropriate social behavior. Initially, the researchers immediately delivered praise and a token following compliance. The researchers provided rules outlining that the token could be exchanged for preferred items and conducted three trials where the participant earned a token and exchanged it for a backup reinforcer. Finally, LeBlanc et al. asked the participant comprehension questions regarding his token economy (e.g., "What do you do to earn tokens?" and "What can you get for earning tokens?"). Once the participant correctly answered these questions, token training was considered finished.

Becraft and Rolider (2015) used two phases to train the token economy to an individual with Seckel Syndrome (a genetic disorder associated with microcephaly and dwarfism) and mild intellectual disability. Initially, the researchers delivered a token on a time-based schedule, and prompted the participant to exchange the token for a high–preferred edible using least-to-most prompting. In the next phase, the researchers delivered a token contingent on folding notecards. Token were exchangeable for a small edible. Despite outlining the general phases, the researchers did not provide information regarding their criterion to move phases or when token training was complete.

Frequently, investigations of token economies either do not report token pairing procedures or report recruiting participants who have a history of earning tokens (Carnett et al., 2014; Fiske et al., 2015; Gilley & Ringdahl, 2014; Karracker, 1977; Plavnick et al., 2010; Tarbox et al., 2002; Truchlicka, MacLaughlin, & Swain, 1998). Studies with participants who have a history of tokens often do not report information regarding the duration of exposure to tokens or other variables (e.g., prior production and exchange ratios, assessments to determine if the token is a reinforcing stimulus, etc.). In addition, few investigations have looked at or reported data verifying the tokens functioned as reinforcers (i.e., used reinforcer assessments). In one notable exception, Fiske et al. (2015) used a progressive-ratio (PR) reinforcer assessment to determine the reinforcing value of tokens from previous token economies (laminated pictures of pennies that were currently used in the participant's classroom) compared to unconditioned reinforcers (edibles) and unpaired tokens (laminated smiley faces) with two children with autism. For one participant, responding in the paired token and edible condition were similar. Both conditions produced differentially higher levels of responding when compared to the unpaired token condition, suggesting that tokens with a history may function as reinforcers when compared to tokens without a history. The other participant showed variable responding in both the paired token and unpaired token conditions, while the edible condition had stable, consistent responding. For this participant, modifications to the current token economy intervention were warranted given that the paired token did not support consistent higher levels of responding. In addition, such findings suggest that a history with a token (i.e., use of token in a classroom) does not necessarily mean that the token is established as a reinforcer.

Currently, investigations using token economies provide little information regarding the types of pairing procedures used. There are two types of pairing procedures within the literature to condition a token as a generalized conditioned reinforcer: stimulus–stimulus pairing and response–stimulus pairing. These pairing procedures have been evaluated in the basic literature (e.g., DiCiano & Everitt, 2004) and with clinical populations (Dozier, Iwata, Thomason–Sassi, Wordsell, & Wilson, 2012; Greer, Pistoljevic, Cahill, & Du, 2011). During stimulus–stimulus pairing, a reinforcer or an established secondary (conditioned) reinforcer is paired with a neutral stimulus and delivered to the individual contingent on a time–based schedule (e.g., every 30 s) regardless of the individual's responding. For example, Greer et al. (2011) used stimulus–stimulus pairing to condition adult voices as reinforcers to increase the amount of story–listening behavior exhibited by three children with ASD. During response–stimulus pairing, an unconditioned reinforcer or an established conditioned reinforcer is paired with a neutral stimulus and delivered contingent on a target response. For example, Hanney (2015) used response-stimulus pairing procedures to condition preschool-aged peers as reinforcers to children with developmental disabilities who had deficits in social skills. Three participants had successfully conditioned peers in the response-stimulus condition, one participant indicated a change in overall preference for peers as compared to baseline, and two participants had no successfully conditioned peers.

Target behavior selection and identification. When beginning a token economy, the practitioner must identify the target responses and the operational definitions must be clear, concise, and objective (Kazdin & Bootzin, 1972; Phillips, 1968). Interventions using token economies have decreased problem behavior, including tics (Capriotti, Brandt, Ricketts, Espil, & Woods, 2012), inappropriate social interactions (LeBlanc et al. 2000), and aggression (Himle, Woods, & Bunaciu, 2008; LePage et al., 2003). Token economies can also be used to increase appropriate behavior, including food acceptance (Kahng et al., 2003), refraining from skin picking (Touissant & Tiger, 2012), on-task classroom behavior (Donaldson, DeLeon, Fisher, & Kahng, 2014; Jowett Hirst, Dozier, & Payne, 2016), compliance to adult directives (Eluri, Andrade, Trevio, & Mahmoud, 2016; Schmidt et al., 2016, Wadsworth, Hansen, & Willis, 2015), engagement (Bellone, Dufrene, Tingstrom, Olmi & Barry, 2014), reading skills (Lahey & Drabman, 1974; Plavnick, Thompson, Englert, Mariage, & Johnson, 2016), appropriate manipulation of competing sensory items (Cavalari, DuBard, & Luiselli, 2014), healthy life skills like jump roping and walking (Alstot, 2012; Krentz, Miltenberger, & Valbuena, 2016), social skills (Gilley & Ringdahl, 2014; LeBlanc et al., 2000), appropriate driving speed (Mullen, Maxwell, & Bedard, 2015), listener skills (Jenkins, Hirst, & DiGennaro Reed, 2015), room cleanliness (Taylor & Mudford, 2012), and verbal behavior (Dickman, Bright, Montgomery, & Miguel, 2012; Taylor & Mudford, 2012).

Backup reinforcers. A crucial component of token economies is the identification of backup reinforcers. Backup reinforcers are preferred stimuli that that are "purchased" using the tokens earned. Backup reinforcers can be a variety of stimuli, including tangible items (e.g., toys, trinkets, electronic devices), special privileges (e.g., homework passes, a walk in the hall, visiting a favorite teacher or friend), and edible items. Backup reinforcers should be identified prior to token economy implementation and individualized to the participant (Gillis & Pence, 2015).

Practitioners can use the information obtained from functional analyses and preference assessments to identify stimuli that could function as reinforcers in the token economy. If a functional analysis indicates that the individual's problem behavior is maintained by escape from work, backup reinforcers can include activities that involve the removal of aversive work situations, such as breaks, leaving class early, or homework passes. Recently, researchers have used token economies with children who engaged in behavior maintained by "mand compliance" (Eluri et al., 2016; Schmidt et al., 2016). Mand compliance, a blended behavior function that includes attention, escape, and tangible consequences, refers to a functional reinforcer that consists of a therapist ceasing the current activity and following the direction of the child within a given situation. For example, an adult may be playing with a toy in a non-preferred manner (e.g., pretending a superhero is having a tea party). When the child screams "I don't like that," the adult would cease the play activity, provide the toy to the child, and say, "You can play your way." To address problem behavior maintained by mand compliance, researchers have developed interventions during which the child must tolerate and comply with adult-directed play and demands in the absence of problem behavior. For example, the participant earns tokens following compliance that are exchangeable for 1 min of access to "her way" during which the therapist will comply with appropriate mands (Eluri et al., 2016; Schmidt et al., 2016).

Token schedule of reinforcement. Determining the criterion for token delivery (i.e., production ratio) must be outlined prior to implementing the token economy. The production ratio designates the number of responses that must occur or the duration that must elapse prior to delivery of a token (Hackenberg, 2009). Once the criterion for token delivery is met, the practitioner should immediately deliver the token. General recommendations are to use dense schedules of reinforcement when beginning a token economy and then gradually thin to intermittent schedules of reinforcement (Cooper, Heron, & Heward, 2007; Gillis & Pence, 2015; Miltenberger, 2012).

During a DRA intervention, tokens are earned contingent on appropriate, replacement behavior, and tokens are not provided when problem behavior occurs (Brigham, Finfrock, Breunig, & Bushell, 1972; Hobbs & Holt, 1976; Johnston & Johnston, 1972; Phillips, 1968). Frequently, the target appropriate behavior is reinforced on a fixed-ratio (FR) 1 schedule (e.g., Becraft & Rolider, 2015; Carnett et al., 2014; Dickman et al., 2012; Everett, Hayward, & Meyers, 1974; Gilley & Ringdahl, 2014; Jenkins, Hirst, & DiGennaro Reed 2015; Krentz et al., 2016; Tarbox et al., 2002; Taylor & Mudford, 2012; Schmidt et al., 2016). Other studies have used intermittent schedules, including variable-interval (VI) schedules (e.g., Breyer & Allen, 1975; Vargo & Ringdahl, 2015) and fixed-interval (FI) schedules (e.g., DiGennaro, Martens, &

Kleinmann, 2007), higher fixed-ratio and variable-ratio schedule values (Field, Nash, Handwerk, & Friman, 2004), and have increased the duration of differential-reinforcement-of-other-behavior schedules (e.g., Fox, Hopkins, & Anger, 1987, McLaughlin & Malby, 1972; Phillips, 1968; Schmidt et al., 2016; Shillingsburg, Lomas, & Bradley, 2012; Touissant & Tiger, 2012). For example, Breyer and Allen (1975) used a VI 8-min schedule where the first instance of listening, seatwork, or voluntary reading following the interval was reinforced with a token. Kelleher (1958) trained chimpanzees to engage in lever pressing that produced access to tokens on a FR 30 schedule with exchange requirement of 50 tokens.

Exchange ratio. Another consideration when outlining components of a token economy is the exchange ratio, which is the number of tokens needed to exchange for the backup reinforcer. Prior to implementing a token economy, Gillis and Pence (2015) suggest that behavioral observations could inform exchange prices. During these observations, the researcher can determine initial levels of appropriate behavior and problem behavior and make the exchange rate an obtainable amount that would allow the child to contact the backup reinforcer. Within the current literature, some studies do not indicate exchange ratios (e.g., Becraft & Rolider, 2015; Karraker, 1977), while other researchers use arbitrary, preselected exchange ratios of 5 tokens (e.g., Gilley & Ringdahl, 2014; LeBlanc et al., 2000), 10 tokens (e.g., Carnett et al., 2004; Donaldson et al., 2014; Jowett Hirst et al., 2016), or 15 tokens (Eluri et al., 2016). Other exchange procedures can be arranged based on the passage of time. For instance, Zlomke and Zlomke (2003) implemented a token economy within a classroom where participants could exchange every 2 hours. Exchanges can also be designed to occur following the completion of designated activity. For example, Taylor and Mudford (2012) set up exchange periods for participants in a substance-abuse rehabilitation setting, in which participants were allowed to exchange tokens either after the completion of a therapy session, meal, or chore.

Researchers have arranged exchange ratios to occur in a distributed format (e.g., one instance of completion results in 30-s access to an item) or an in an accumulated manner (e.g., individual can not exchange until 10 tokens are earned for 5-min access backup reinforcer; DeLeon et al., 2014; Dickman et al., 2012). DeLeon et al. (2014) compared distributed and accumulated token exchange ratios with four individuals with intellectual disabilities who engaged in problem behavior. Accumulated schedules supported higher rates of task completion with three participants. When participants were given a choice of accumulated or distributed exchange procedures in a concurrent-chain arrangement, participants consistently selected accumulated schedules when stimuli were activity-based items (e.g., electronics, music), suggesting a preference for the accumulated arrangement with activity-based backup reinforcers.

Treatment Integrity

Though token economies can effectively increase appropriate behavior and decrease problem behavior, little is known about the extent to which behavior–change agents are implementing token economies with high procedural fidelity and how failure to implement the procedures with high fidelity affects the token economy. One notable concern when implementing DRA procedures is the possibility of treatment relapse or resurgence. Resurgence is the phenomenon in which a previously extinguished behavior, like problem behavior, temporarily increases when alternative reinforcement is discontinued. In fact, several investigations have demonstrated resurgence following the discontinuation of DRA (or FCT) in experimental (Mace et al., 2010), translational (Smith, Smith, Shahan, Madden, & Twohig, 2017), and clinical preparations (Pritchard, Hoerger, Mace, Penney, & Harris, 2014). It is a paramount importance to understand how clinical populations may respond when treatment is discontinued.

One way of conceptualizing resurgence is in the framework of procedural fidelity. Procedural fidelity, or treatment integrity, is the extent to which an intervention is implemented in the manner in which it is designed (DiGennaro Reed, Reed, Baez, & Maguire, 2011; Gresham, 1989). Within the natural environment (e.g., classroom, home), behavior–change agents may make treatment integrity errors that can affect treatment outcomes (e.g., Carroll, Kodak, & Fisher, 2013). Understanding the influence of errors on treatment outcomes may guide recommendations to practitioners to make interventions that are durable to treatment failures. When compared to the abundance of research on the effectiveness of behavioral interventions, there are few well–designed and controlled investigations of treatment integrity errors (e.g., Carroll et al., 2013; Pence & St. Peter, 2015; St. Peter, Byrd, Pence, & Foreman, 2016; St. Peter Pipkin, Vollmer, & Sloman, 2010; Vollmer et al., 1999).

There are three types of treatment integrity errors: omission errors, commission errors, and blended errors (St. Peter Pipkin et al., 2010). Omission errors consist of the behavior–change agent failing to perform a component of the protocol. For example, when using a token system, a teacher may fail to deliver a token when the child engages in appropriate behavior. A commission error is the addition of a component that has not been programmed. An example of a commission error is said to occur when an individual provides a token for a response not designated in the procedure. Finally, blended errors occur when the behavior–change agent adds a programmed component and simultaneously removes a programmed component. For example, a teacher may provide the backup reinforcer (e.g., a piece of candy) rather than providing the token for appropriate behavior. In this example, the teacher omitted the token delivery while simultaneously delivering an unprogrammed backup reinforcer.

Skill acquisition. Treatment integrity errors can influence how individuals with disabilities acquire skills. For example, DiGennaro Reed et al. (2011) examined the effects of commission errors during discrete-trial instruction on the acquisition of receptive identification of nonsense shapes with three individuals with ASD. The researchers evaluated the effects of delivering a token following an error at differing levels of integrity (100%, 50%, and 0% commission errors). For all participants, the high-integrity condition had the highest level of independent responses for all participants. Two of three participants had no difference in independent responding between 50% (commission errors on 50% of opportunities) and low-integrity (commission errors on all opportunities) conditions. Carroll et al. (2013) evaluated three conditions (control, high integrity, and low integrity) with six children with ASD on an expressive task (e.g., saying a letter when shown on a card) or a play task. All participants mastered skills in the high-integrity condition and one participant mastered skills in the low-integrity (errors occurred on 67% of trials) condition. In addition, reinforcer-delivery errors (delayed reinforcement) during skill-acquisition programming with children with ASD can affect correct responding. Carroll, Kodak, and Adolf (2016) observed that both participants mastered the skill (labeling a two-dimensional stimulus) with immediate praise and tangible reinforcement, while only one participant acquired skills when reinforcement was delayed. Jenkins et al. (2015) also demonstrated that when commission errors in the form of reinforcing an incorrect response with a token occurred during discrete-trial teaching, three of four participants had impaired acquisition of listener-responding skills. Following the commission error evaluation, researchers returned to 100% correct fidelity. Authors found that high-fidelity instruction after a period of commission errors reversed the adverse effects of poor instruction.

High levels of integrity are necessary when teaching verbal behavior, specifically requesting items (i.e., mands). Pence and St. Peter (2015) evaluated the effects of providing items noncontingently (i.e., regardless if an independent request occurs). The target request was acquired in the fewest sessions in the 100% integrity condition. Participants mastered the skill with 70% integrity (the item was delivered independent of responding on 30% of trials). However, requests were not acquired with the participant did not have to mand for the item (0% integrity) or was only required to ask for the item on 40% of trials.

Problem behavior. One method to study levels of treatment integrity and its influence on appropriate and problem behavior is parametric analyses (e.g., Leon, Borerro, & DeLeon, 2016; St Peter Pipkin et al., 2010; St Peter et al., 2016; Vollmer et al., 1999). For example, Vollmer et al. (1999) evaluated how treatment integrity errors (0%, 25%, 50%, 75%, or 100% of opportunities) during DRA influenced problem behavior and appropriate behavior with four participants with problem behavior maintained by access to items or escape from demands. During the DRA intervention at full implementation (i.e., 100% integrity), the researcher implemented extinction following every instance of problem behavior (i.e., prompted through the demand or item continued to be restricted) and reinforced every instance of appropriate behavior (requests for breaks or toys). Integrity failures consisted of providing the functional reinforcer for problem behavior (commission error) or failing to reinforce appropriate behavior (omission error). Participants engaged in appropriate behavior when integrity was high (e.g., 75% and 100% integrity). However, problem behavior increased with low to moderate (0%, 25%, or 50%) integrity.

St. Peter Pipkin et al. (2010) used a translational approach to examine the effects of different types of errors (omission, commission, and blended) during a DRA intervention. In the first experiment, participants had to click on red (defined as appropriate behavior) or black (defined as problem behavior) circles. Participants that were exposed to 80% and 60% integrity with omission, commission, or blended errors engaged in higher rates of red-circle clicking (appropriate behavior) than black-circle clicking (problem behavior) compared to baseline. Increases in black-circle clicking occurred when integrity decreased to 40% and 20%. In addition, the efficacy of DRA treatments decreased based on different types of treatment errors. Participants who experienced commission errors (reinforcement of problem behavior) and blended errors (reinforcement of problem behavior and failing to reinforce appropriate behavior) engaged in decreased levels of appropriate behavior compared to participants who experienced omission errors (no reinforcement following appropriate behavior). However, this effect was only seen when integrity was relatively low (20 and 40% levels).

In the second phase of their study, St. Peter Pipkin et al. (2010) replicated the procedures used in the human-operant preparation by evaluating the effects of blended errors (a combination of omitting reinforcers for appropriate behavior and providing reinforcers for problem behavior) on levels of on-task behavior. For example, in the 80% integrity phase, 80% of on-task behavior and 20% of off-task behavior resulted in adult attention. The participant engaged in more appropriate behavior during the higher levels of integrity (i.e., 80% and 60% conditions). Subsequently, the participant engaged in more problem behavior during the 40% and 20% conditions.

Delays to reinforcers can also be conceptualized as a treatment integrity error that may be detrimental in maintaining appropriate behavior, and it is important to understand how immediacy of a conditioned reinforcer influences its effectiveness. Leon et al. (2016) examined how delayed primary and conditioned reinforcers influenced responding on mastered academic tasks for three children with intellectual disabilities. The researchers evaluated three conditions: edibles, tokens with postsession exchange (tokens exchanged at the end of the session), and tokens with posttrial exchange (token immediately exchanged after response requirement met). Within each condition, participants were exposed to a no-reinforcement baseline, immediate reinforcement (either edible or token), and delayed reinforcement at six different delays. Greater response persistence occurred during delayed edible reinforcement. However, when token delivery was delayed, rapid decreases in responding occurred. For two participants, decreases in responding were observed with a 3-s delay. These results suggest that when tokens are delayed (even as little as 3 s) there may be a weakening of the conditioned reinforcing effects.

In sum, previous investigations (Leon et al., 2016; St. Peter Pipkin et al., 2010; Vollmer et al., 1999) suggest that DRA is relatively durable behavioral intervention as indicated by low levels of problem behavior (similar to 100% integrity conditions) with moderate levels of integrity. However, levels of integrity at or below 40% are detrimental to treatment outcomes. In addition, different kinds of treatment integrity errors have differential effects on treatment outcomes. St. Peter Pipkin et al. (2010) concluded that commission (delivering a reinforcer following problem behavior) and blended errors (combination of delivering reinforcers following problem behavior and omitting reinforcers for appropriate behavior) are more detrimental than omission errors (failure to reinforce appropriate behavior). Leon et al., 2016 found that if delays occurred between the response and the delivery of tokens, it was detrimental to participant responding on mastered tasks. Finally, the resurgence literature demonstrates that previously eliminated behavior can recur when reinforcement is discontinued for appropriate behavior. It is crucial to understand how individuals who engage in problem behavior respond in situations in which a treatment is implemented with low levels of fidelity.

Although levels of treatment integrity are important predictors of treatment outcomes, to date, relatively few studies have systematically evaluated the effects of treatment integrity on token-reinforcement interventions. Token economies can be used in a variety of settings to increase appropriate behavior and decrease problem behavior. One omission error that occurs when implementing a token economy involves failing to provide tokens at the specified schedule following appropriate behavior. However, the effects of omission errors on appropriate behavior and problem behavior during a token economy are unknown in a clinical population. The first purpose of the current investigation is to evaluate the extent to which treatment integrity errors in the form of omission of token delivery influence the efficacy of token economies. One variable that could influence the durability of a token economy is the programmed schedule of reinforcement. For example, tokens can be delivered on a continuous schedule of reinforcement using a FR 1 schedule of token delivery for appropriate behavior or a variable schedule of reinforcement using a VR 3 schedule of token delivery for appropriate behavior. The second purpose of the current investigation is to evaluate the effect of omission errors when appropriate behavior is reinforced on a continuous (FR 1) or a variable (VR 3) schedule of reinforcement.

Chapter 2: Experimental Procedures

Participants and Setting

Six children with developmental disabilities participated. All participants were referred for the assessment and treatment of aggression. Teacher and parent report suggested that aggression could be maintained by escape from demands. Nate was a 4-year-old Caucasian male diagnosed with ASD. He communicated vocally using three to four-word phrases. Greg was a 4-year-old Caucasian male diagnosed with ASD. He communicated vocally using two to three-word phrases. James was an 8-year-old African American male diagnosed with developmental delay. He communicated vocally using full sentences (5+ words). Ray was an 8-year-old Caucasian male with a traumatic brain injury. He communicated using one-word approximations (e.g., saying "mo" for more) and using gestures (i.e., pointing) to an item he wanted. James was an 8-year-old African American male diagnosed with developmental delay. He communicated vocally using full sentences (5+ words). Kyle was a 9-year-old African American male diagnosed with ASD. He also communicated using ProloQuo2Go application on a tablet with two to three-word phrases. Kari was an 8-year female diagnosed with ASD and Noonan Syndrome who communicated vocally using full sentences (5+ words). All participants were ambulatory and could follow at least 5 one-step directions.

Three to six appointments were conducted per week with each appointment consisting of three to nine sessions. Sessions were conducted in a therapy room with a table and at least two chairs on an university campus or in a quiet area of the participant's school. The room was void of any potentially distracting, extraneous visual (e.g., bright posters, tangible items within sight) and auditory (e.g., loud noises, music) stimuli.

Data Collection, Interobserver Agreement, and Procedural Fidelity `

Paper-and-pencil data collection were used to record occurrences of the dependent variables. The target problem behavior for all participants was aggression. Aggression was defined as any actual or attempted instance of punching, hitting, slapping, or kicking another individual from at least a 16 cm distance, any instance of biting, scratching, pinching, pushing, shoving, grabbing, or throwing an object that made contact with another person. Compliance was defined as the participant completing the specified response (independently or following a gestural or model prompt) with the adult directive within 5 s. Data were collected on aggression and compliance across all conditions. Session duration was 10 min for all conditions.

Undergraduate and graduate research assistants were trained to collect data by the researcher. The researcher provided direct instruction by reviewing the operational definitions for target behavior, data–collection procedures, and procedures for calculating interobserver agreement (IOA). The researcher conducted role–plays with the trainees and had them practice data–collection procedures. The researcher provided corrective feedback following each role–play. Trainees practiced collecting data by observing participants displaying the target behavior in situ or from video recordings. Trainees were considered reliable data collectors after they recorded behavior with at least 90% accuracy across three consecutive sessions.

Two trained data collectors independently recorded data during in-vivo sessions or from video recordings. IOA was calculated using the mean count-per-interval method (Cooper et al., 2007). Each session was divided into 10–s bins and the number of responses scored within each bin was compared across observers. The smaller number of responses in the bin was divided by the larger number of responses to obtain a proportion. The proportions were summed and divided by the number of intervals and then multiplied by 100 to obtain a percentage. IOA was collected on 47.5% of sessions for Nate. Mean IOA was 97.4% (range, 81.5% to 100%) for aggression and 91% (range, 76.6% to 100%) for compliance. IOA was calculated for 43.3% of sessions for Greg. Mean was 98.5% (range, 80.8% to 100%) for aggression and 92.8% (range, 77.5% to 100%) for compliance. IOA was calculated for 31.8% of sessions for Ray. Mean IOA was 96.7% (range, 79.4%-100%) for aggression and 93.1% (range, 75 to 100%) for compliance. IOA was calculated for 33.7% of sessions for James. Mean IOA was 96.5 (range, 76.8% to 100%) for aggression and 88.5% (range, 76.6% to 100%) for compliance. IOA was collected for 45.1% of sessions for Kyle. Mean IOA was 96.5% (range, 77.9% to 100%) for aggression and 92% (range, 78.3% to 100%) for compliance. IOA was collected for 10% of sessions for Kari. Mean IOA was 95.2% (range, 89.2% to 100%) for aggression and 94.6% (range, 90.1% to 100%) for compliance.

Graduate and undergraduate research assistants were trained to collect procedural fidelity using similar procedures as described for training data collection. Observers were required to collect procedural fidelity data for a minimum of three sessions with at least 90% accuracy to be considered reliable data collectors. Observers collected data on accuracy which with the researcher (1) delivered demands; (2) used three-step prompting for compliance; (3) implemented the correct consequence following aggression; and (4) delivered the correct consequence following appropriate behavior (e.g., provided a token). During the phases in which treatment integrity errors were programmed, integrity scores were also based on the researcher's adherence to the programmed consequences (e.g., refraining from providing a programmed token). Treatment fidelity scores were calculated by dividing the correct researcher responses by the total number of correct and incorrect responses in a session. This fraction was converted into a percentage. Procedural fidelity was collected on 28.4% of sessions and averaged 95.5% (range, 75% to 100%) across all participants.

Independent Variables and Experimental Design

A reversal design was used to demonstrate the effects of treatment integrity errors on levels of compliance and aggression. The programmed omission error consisted of the researcher failing to deliver a token following compliance. Five levels of omission errors (failing to deliver a token) were evaluated: 80%, 60%, 40%, 20%, and 0% integrity. Nate, Greg, Kyle, and Kari were exposed to descending levels of treatment integrity initially and then ascending levels of treatment integrity. James and Ray were exposed to ascending levels of treatment integrity initially followed by descending levels.

Procedure

Preference assessment. A multiple–stimulus without–replacement (MSWO) preference assessment (DeLeon & Iwata, 1996) was used to identify highly and moderately preferred stimuli. Seven items were identified based on parental and teacher report of the participant's preferred edibles or tangibles. During the MSWO preference assessment, the researcher presented an array of seven items. After selecting an item, the participant was allowed to consume the edible or manipulate the tangible item for 30 s (after which the selected item was removed and placed out of sight). The remaining items in the array were rotated, and the researcher provided the individual with an opportunity to select another item. The researcher blocked any attempts to select more than one item. If no item was selected, the researcher prompted the participant to select an item and waited at least 5 s before removing the remaining items. The researcher repeated these steps until all items were selected. These procedures were repeated for seven blocks of trials. The percentage of selection for each item was calculated by dividing the percentage of opportunities an item was selected by the number of trials that the item was presented. Edibles and tangible items identified as highly preferred and moderately preferred were used in the token-economy procedures. In addition, if the participant requested an item that was part of the original preference assessment, the researcher allowed the participant to earn this as a backup reinforce during the parametric evaluation.

Functional analysis. Functional analyses were conducted to determine the maintaining variable of aggression. Based on parental interviews and observations, a pairwise functional analysis (Iwata & Dozier, 2008) comparing escape and control (play) conditions was conducted for Nate, Kari, and Ray. Following an initial standard multielement functional analysis (Iwata et al., 1982/1994; Rooker, Iwata, Harper, Fahmie, & Camp, 2011) with Greg and James during which aggression was not evoked, we conducted a pairwise functional analysis with modifications to the escape condition. Finally, we completed a standard multielement functional analysis with Kyle for his aggression that included escape, attention, and tangible test conditions, and an additional pairwise comparison of tangible and control conditions.

During the escape condition, a minimum of three different classes of tasks (e.g., motor directives, math, and listener skills) was used. Tasks were individually determined for each participant and were based on parental or teacher report of difficulty and association with aggression. Labeling, listener skills, and motor imitation tasks were used with Nate. Fine motor imitation, listener skills, and handwriting tasks were used with Greg. For James and Kari, the escape condition included academic tasks involving multiplication, multi-step addition and subtraction, and writing tasks. Simple motor directives, imitation, and listener tasks were given with Ray. Finally, tasks for Kyle were one-digit subtraction, addition, and fine-motor imitation demands. During the escape condition, a researcher presented instructions to perform a task using a least-to-most prompting hierarchy (verbal, gestural, and physical guidance, with 5 s between prompts. Contingent on aggression, the researcher stated, “You can have a break” or “Okay, you don’t have to,” and provided a 30–s break. In the attention condition, the researcher diverted attention and pretended to be busy with another task. The participant had continuous access to low–preferred materials. The researcher provided attention, such as a brief reprimand, consolation statement (e.g., “Are you ok?”), or a redirection statement (e.g., “Keep safe hands”) when the participant engaged in aggression. In the access to tangibles condition, the participant had 2-min access to tangibles prior to beginning the session. The researcher removed preferred items and stated, “My turn.” Contingent on aggression, the researcher allowed access to the tangibles for 30 s. During the control (play) condition, the participant had continuous access to highly preferred materials, adult attention was delivered every 30 s, and no demands were delivered.

Baseline. Baseline sessions were identical to the escape condition of the functional analysis. The researcher presented instructions to perform a task using a least-to-most prompting hierarchy (verbal, gestural, and physical guidance, with 5 s between prompts). If the participant engaged in compliance (following the initial directive or the model/gesture prompt), the researcher provided nondescriptive praise (e.g., “Good.”). Contingent on aggression, the researcher stated, “You can have a break” and provided a 30–s break. During the 30–s break, the researcher removed work materials and oriented away from the participant.

Token training. Following baseline, the participant underwent token training, consisting of token pairing, a reinforcer assessment, and schedule thinning.

Token pairing. The researcher sat next to the participant and asked the participant what she or he would like to earn while holding up a minimum of two items identified as preferred from the preference assessment. Following a selection, the researcher stated, “When you follow directions, you can earn tokens for [preferred item],” and began the session. The researcher presented a directive and used a least-to-most prompting hierarchy for compliance (outlined in baseline). Contingent upon compliance (independently or following the model), the researcher delivered praise on a FR 1 schedule and provided a token to the participant. Following the token delivery, the researcher vocally prompted the participant to say, “Let’s trade” and provided manual guidance for the participant to trade the token for a backup reinforcer.

During the pairing sessions, vocal and manual guidance prompts to exchange the token were faded using a most–to–least, time–delay prompting procedure. Prompt steps included the following: (1) immediate physical guidance to trade-in the tokens and full-vocal prompt of “Let’s trade;” (2) 5-s delay to physical guidance and full-vocal prompt; (3) 10-s delay to physical guidance and full-vocal prompt. The criteria to increase prompt steps was 10 consecutive trials with correct (either independently or at the prescribed prompt step) token trade-in. A pairing trial began once the researcher delivered the instruction for the target response and ended once the participant consumed the edible item or if 30 s elapsed. At the end of each pairing trial, a 5–s inter–trial interval occurred prior to beginning the next trial. Each pairing session contained 50 pairing trials. Following each pairing session, the researcher conducted the reinforcer assessment.

Reinforcer assessment. The reinforcer assessment was conducted using a progressive–ratio (PR) schedule as described by Roane et al. (2001). Nate, Greg, James, Kyle, and Kari were required to complete a predetermined number of arbitrary responses (e.g., inserting clothespins into a bin) that increased on an additive of one PR schedule (i.e., PR 1, 2, 3, 4, 5, etc.). Sessions ended when responding ceased for 3 min or after 45 min of total session duration. During the token condition, the researcher delivered one token after each schedule requirement completion. At the end of the session, the researcher delivered two high or moderate-preferred items (e.g., 2 min with tangibles or two pieces of an edible), regardless of the number of tokens earned during the assessment. The token condition was compared to a control condition in which a mundane object with no pairing history (e.g., piece of paper) was provided after each schedule requirement completion. Following the end of the session and independent of responding, the researcher delivered two high and moderate-preferred items regardless of the number of mundane items earned during the assessment. A minimum of three sessions of each condition were conducted. Differentiation between the data paths (i.e., higher responding in the token condition compared to the control condition) indicated that the tokens functioned as reinforcers.

If the token did not function as a reinforcer, another token pairing session was conducted. Nate and Greg needed additional pairing sessions to establish the token as a reinforcer during the reinforcer assessment. Nate completed an additional two pairing sessions (a total of 150 trials), and Greg completed an additional pairing session (a total of 100 trials). Kyle, James, and Kari participated in one pairing session (a total of 50 trials). For all participants (except Ray) differentially higher levels of task completion occurred in the token condition when compared to the control following the pairing session(s), suggesting that tokens functioned as reinforcers for these participants.

For Ray, the reinforcer assessment was modified to a concurrent arrangement in which two pieces of paper were presented on a table. Contingent on tapping the piece of paper, the associated consequence (token or piece of paper) was provided following each response. Ray engaged in differentially higher response rates for tapping the paper associated with token delivery, indicating that the token functioned as a reinforcer.

Schedule thinning. Schedule thinning occurred to increase the number of tokens the participant was expected to earn before exchanging for the backup reinforcer. Initially, the participant had a token board with four tokens on the board and a space for one token. The researcher presented task demands as described in baseline. Following compliance, the researcher delivered a token and allowed the participant to place the token on the board. If the participant did not place the token on the board within 5 s, the researcher gestured or modeled token placement on the board. If the participant did not place the token on the board after the gesture or model prompt, the researcher provided physical guidance for the participant to place the token on the board. Once the participant placed the last token on the token board, the researcher waited 5 s before prompting, “Let’s trade.” If the participant did not hand the tokens to the researcher, she gestured to the board and repeated the directive of “Let’s trade.” If no response or an incorrect response occurred, the researcher manually guided the participant to trade–in the tokens by removing all five tokens and placing them in the researcher’s hand. Following two successful trials with token exchanges in the absence of aggression on the FR 1 (token) FR 1 (exchange) schedule, the token exchange ratio was thinned.

Greg, Ray, and Kyle were assigned to a token economy with a continuous schedule of token delivery for appropriate behavior (FR 1 token/FR 5 exchange). In the continuous-delivery token economy, the researcher reinforced every instance of compliance with a token and five tokens were required to exchange for a backup reinforcer. The researcher increased thinning steps after two consecutive exchanges in the absence of aggression.

Nate, James, and Kari were assigned to a token economy with a variable schedule of token delivery for appropriate behavior (VR 3 token/FR 5 exchange). In the variable-delivery token economy, the researcher reinforced compliance on a VR 3 scheduled where, on average, every third instance of compliance was reinforced (range, 2–5 responses) with a token and five tokens was required to exchange for a backup reinforcer. Nate, James, and Kari proceeded through the progression of thinning in the following manner: FR 1 token FR 1 exchange, FR 1 token FR 2 exchange, FR 1 token FR 3 exchange, FR 1 token FR 4 exchange, FR 1 token FR 5 exchange, VR 2 (range, 1-3) token FR 5 exchange, and VR 3 (range, 2-5) token FR 5 exchange. The thinning step was increased after two consecutive trials in the absence of aggression. The schedule of reinforcement that outlined which instances of compliance were followed by a token was generated by the procedures outlined in Bancroft and Bourret (2008).

To help control for additional exposure to tokens in the VR 3 FR 5 group, participants in the FR 1 FR 5 group (Greg, Ray, and Kyle) continued to earn tokens at the FR 1 FR 5 ratio that were yoked to a the number of tokens earned by a participant in the VR 3 FR 5 group (Nate, James, and Kari).

Token Evaluation (100% integrity)

During the token evaluation, the researcher used the same directives and prompts as outlined in baseline. If the participant engaged in compliance, the researcher provided a token on either a FR 1 or VR 3 schedule (depending on the group the participant was assigned to). Once the participant earned five tokens and handed the tokens to the researcher, the researcher provided access to preferred edibles or tangible items for 30 s. If the participant did not hand the tokens to the researcher, she gestured to the board and repeated the directive of "Let's trade." If no response or an incorrect response occurred, the researcher manually guided the participant to trade–in the tokens by removing all five tokens and placing them in the researcher's hand. If the participant engaged in aggression, the researcher immediately manually guided the correct response to the current directive and did not deliver a token following that instance (FR 1), or count that instance towards the response requirement (VR 3).

Parametric Evaluation

During the parametric evaluation, the researcher followed the procedures outlined in the token evaluation (100% integrity) with the exception that omission errors were programmed. If the participant engaged in compliance that met the schedule requirement (FR 1 or VR 3), the researcher provided a token depending on the designated level of integrity (i.e., 80%, 60%, 40%, 20%, or 0%). During all integrity failure phases, the researcher held a clipboard with a computer–generated sequence of which instances of compliance should result in a token delivery. If an omission error was programmed, the researcher did not provide a token once the production ratio was met and presented the next directive. If no error was programmed for the opportunity to earn a token, the researcher delivered a token.

Five levels of reduced treatment integrity were evaluated: 80%, 60%, 40%, 20%, and 0%. During the 80% integrity condition, 8 of 10 opportunities in which the participant met the production ratio were reinforced with a token. In the 60% integrity condition, 6 of 10 opportunities in which the participant met the production ratio were reinforced with a token.

During the 40% integrity condition, 4 of 10 opportunities in which the participant met the production ratio were reinforced by the delivery of a token. In the 20% integrity condition, 2 of 10 opportunities in which the participant met the production ratio were reinforced with a token. Finally, in the 0% integrity condition, 0 of 10 opportunities in which the participant met the production ratio were reinforced with a token.

Results

Two highly preferred items were identified for each participant from the preference assessment. Legos™ and Airheads™ were highly preferred for Nate. Greg's preferred items were bubbles and Honeybuns™. James preferred squishy ball and gummy worms. Ray preferred the music camera and M&Ms™. Kyle's highly preferred items were Thomas the Train™ and candy corn. Kari preferred the iPad™ and M&Ms™. High-preferred tangible items were included in the play condition of the functional analysis for each participant. For Kyle, high-preferred items were also used in the restricted access condition of the functional analysis. For all participants highly and moderately preferred items were included as backup reinforcers during the token training and the token economy parametric evaluation.

Figure 1 shows the results of the functional analysis for Nate, Greg, James, Ray, Kyle, and Kari. Nate, Greg, Jake, Ray, and Kari's aggression occurred at differentially higher levels during the escape condition compared to the control condition, indicating that the participants' aggression was maintained by escape from demands. Kyle had elevated levels of aggression during the escape and tangible conditions compared to the control condition, suggesting his aggression was maintained by escape from demands and tangibles.

Table 1 shows the mean percentage of compliance aggregated for each procedural integrity phase for Greg, Kyle, and Ray. Figure 2 displays the percentage of compliance with directives during Greg (top panel), Kyle (middle panel), and Ray's (bottom panel) parametric analysis of omission errors (i.e., failure to deliver a token) for a FR 1 schedule of token delivery. During baseline, Greg's engaged in moderate to low levels of compliance (*M*s = 16.4% and 34.3%, in the first and second baseline phase, respectively). Levels of compliance increased during the 100% token economy phases (*M*s = 94.1% and 90.7%, respectively). When the level of procedural integrity decreased to 80%, compliance remained at high levels (*M* = 87.5%). During the 60%, 40%, and 20% phases, levels compliance were variable and gradually decreased as integrity degraded (*M*s = 54.5%, 32.9%, and 9.5%, respectively). During the 0% integrity phase, compliance remained at low levels (*M* = 10.3%). Overall, levels of compliance gradually increased as integrity levels increased during the 20%, 40%, and 60% phases, with decreasing trends in compliance across sessions within each phase (*M*s = 18.1%, 21.8%, and 34.3%, respectively). With the return to the 80% integrity phase, Greg's compliance increased to higher levels (*M* = 73.7%). During the final 0% integrity phase, Greg's compliance decreased across the phase to low levels (*M* = 23.3%). During the return to 100% integrity, Greg engaged in moderate to high levels of compliance (*M* = 71%).

During baseline, Kyle (Figure 2, middle panel) engaged in low levels of compliance (*M* = 17.1% and 10.6%, respectively). During the 100% and 80% token economy phases, compliance increased to moderate levels (*M*s = 75.6%, 74.7%, and 60%, respectively). As integrity decreased, mean levels of compliance decreased in the 60% phase (*M* = 34.7%) and remained at similar levels in the 40% (*M* = 45.1%) and 20% (*M* = 41.8%) phases. During the 0% phase, compliance decreased to low levels (*M* = 22.4%). With the return to 100% integrity, levels of compliance increased to moderate levels (*M* = 62.9%). Kyle engaged in moderate to low levels of compliance during the reversal to the 0% phase (*M* = 19.9%; range, 5.9% to 47.1%). Low levels of compliance continued during the 20% (*M* = 19.7%) and 40% (*M* = 8%) integrity phases. During the 60% integrity, compliance increased initially; however, compliance decreased across sessions in this phase to low levels (*M* = 22.9%). Kyle engaged in moderate to high levels of compliance during the 80% (*M* = 83.8%) and 100% (*M* = 95.3%) phases.

During baseline, Ray (Figure 2, bottom panel) engaged in low levels of compliance (*M*s = 11.7% and 8.7%, respectively). At 100% integrity, levels of compliance increased (*M* = 57.4%). During the 0% phase, compliance initially increased from baseline, but decreased across the phase (*M* = 13%), remained low in the 20% phase (*M* = 14%), and slightly increased in the 40% phase (*M* = 19.8%). As levels of integrity increased, Ray gradually engaged in higher levels of compliance in the 60% (*M* = 29.1%), 80% (*M* = 40%), and 100% phases (*M* = 54%). When integrity returned to 0%, levels of compliance decreased across the phase (*M* = 21.7%). With 100% integrity was reimplemented, compliance increased (*M* = 42.5%) to moderate, but variable, levels (range, 9.1% to 67.7%). Ray's mean level of compliance decreased slightly when integrity was reduced and remained at similar levels during the 80% (*M* = 37.5%), 60% (*M* = 32.9%), 40% (*M* = 33.4%), and 20% (*M* = 33.2%) phases. Lower levels of compliance occurred in the 0% phase (*M* = 21.7%). Ray engaged in higher and more stable levels of compliance in the final 100% integrity phase (*M* = 56.2%).

Table 1 shows the mean rates of aggression aggregated for each procedural integrity phase for Greg, Kyle, and Ray. Figure 3 displays levels of aggression during the parametric analysis for Greg (top panel), Kyle (middle panel), and Ray (bottom panel). During baseline, Greg engaged in moderate levels of aggression (*M*s = 2.4 resp/min and 1.4 resp/min, respectively) Aggression decreased to low levels during the 100% integrity phases (*M*s = 0 resp/min and 0.1 resp/min, respectively). Aggression remained at low levels during the 80% (*M* = 0.02 resp/min) and 60% (*M* = 0.1 resp/min) integrity phases. During the 40%, 20%, and 0% integrity phases, levels of aggression increased gradually from the preceding phase (*M*s = 1.0, 1.2, and 2.3 resp/min, respectively). Levels of aggression remained at moderate levels during the return to 20% (*M* = 1.5 resp/min) and 40% (*M* = 1.0 resp/min) integrity phases. Aggression decreased at 60% integrity (*M* = 0.6 resp/min). Aggression decreased to low levels during 80% integrity (*M* = 0.1 resp/min). When 0% integrity was reintroduced, aggression increased to higher levels (*M* = 1.0 resp/min). Greg engaged in zero levels of aggression in the final 100% integrity phase.

In baseline, Kyle (Figure 3, middle panel) engaged in aggression at increasing levels (*M*s = 1.8 resp/min and 2.2 resp/min, respectively). Aggression decreased to low levels during the 100% integrity phases (*M*s = 0.01 resp/min and 0.03 resp/min, respectively). During the 80% integrity phase, aggression initially increased, but decreased across the phase (*M* = 0.2). During the 60%, and 40% integrity phases, levels of aggression increased from the preceding phase with increasing levels of aggression across the phase (*Ms* = 0.3 and 2.3 resp/min, respectively). Aggression continued to occur at moderate levels with 20% (*M* = 0.9 resp/min) and 0% (*M* = 1.6 resp/min) integrity. During the 100% integrity phase, aggression decreased to low levels (*M*= 0.2 resp/min), and increased when 0% integrity was reimplemented (*M* = 4.0 resp/min). Aggression occurred at moderate levels during the 20%, 40%, and 60% integrity phases (*Ms* = 0.2, 0.5, 0.6 resp/min, respectively). Aggression occurred at lower levels with 80% integrity (*M* = 0.1 resp/min). Finally, during the 100% integrity phase, aggression occurred at zero levels.

During baseline, Ray (Figure 3, bottom panel) engaged in increasing levels of aggression (*M*s = 1.7 resp/min and 1.2 resp/min, respectively). Aggression decreased to low levels during the 100% integrity phase (*M* = 0.01 resp/min). Aggression increased during the 0% and 20% integrity phases (*M*s = 0.5 resp/min and 0.6 resp/min, respectively). During the 40% integrity phase, aggression decreased across the phase to low levels (*M* = 0.3 resp/min). Aggression remained low during the 60% (*M* = 0.02 resp/min) and 80% (*M* = 0.01 resp/min) integrity phases. When integrity increased to 100%, aggression remained at lower levels (*M* = 0.1 resp/min). During the 0% integrity phase, aggression increased to high levels (*M*= 1.4 resp/min) and then decreased to low levels (*M*= 0.2 resp/min) with the return to 100% integrity phase. During the 80% phase, aggression increased (*M*= 1.0 resp/min). Similar levels of aggression were observed in the 60% and 40% phases (*M*s = 0.7 and 0.7 resp/min, respectively), with increased variability at 40% integrity (range, 0 to 4.2 resp/min). Aggression decreased in the 20% integrity phase (*M* = 0.2 resp/min), and then increased in the 0% integrity phase (*M*= 0.5 resp/min). In the final 100% integrity phase, aggression decreased to low levels (*M*= 0.1 resp/min).

Table 2 shows the mean percentage of compliance aggregated for each procedural integrity phase for Kari, Nate, and James. Figure 4 shows the percentage of compliance during Kari (top panel), Nate (middle panel), and James' (bottom panel) parametric analysis of omission errors when tokens were delivered under a VR 3 schedule of reinforcement for compliance. During baseline, Kari engaged in low levels of compliance (*M*s = 0% and 12.2%, in the first and second baseline phase, respectively). During the 100% integrity phase, levels of compliance increased (*M*s = 88.9% and 75.8%, respectively). When the level of procedural integrity decreased to 80% and 60%, compliance maintained at high levels (*M*s = 87.8% and 86.8%, respectively). Overall, as levels of integrity decreased during the 40%, 20%, and 0% phases, decreasing trends of compliance occurred (*M*s = 39.8%, 30.9%, and 3%, respectively). With the reimplementation of 100% integrity, levels of compliance increase (*M* = 97.6%). During the return to the 0% and 20% integrity phases, compliance decreased and occurred at low levels (*M*s = 15.3% and 3.2%, respectively). Compliance increased in the 40% integrity phase (*M* = 77.8%). During the 60%, 80%, and 100% integrity phases, compliance remained high (*M*s = 93.8%, 78.8%, 82.6%, respectively).

During baseline, Nate (Figure 4, middle panel) engaged in low levels of compliance (*M*s = 8% and 30.8%, respectively). During the 100% token economy phase, levels of compliance increased (*M*s = 89.9% and 59.5%, respectively). When levels of procedural integrity decreased, compliance maintained at high levels during the 80%, 60%, 40%, 20% phases (*M*s = 88%, 96.8%, 94.5%, and 84.4%, respectively). During 0% integrity, compliance was variable and decreased to low percentages (*M* = 46.1%). With the return to 100% integrity, compliance increased to high levels (*M* = 92.2%). When 0% integrity was reintroduced, Nate engaged in variable levels of compliance (*M* = 27.9%; range, 0.4% to 81.7%). Nate engaged in moderate to high levels of compliance continued during the 20% (*M* = 87.7%) and 40% (*M* = 62.7%) integrity phases. At 60% integrity, compliance increased to high levels (*M* = 91.1%). Nate engaged in high levels of compliance during the 80% and 100% phases (*M*s = 97.4% and 95.2%, respectively).

During baseline, James (Figure 4, bottom panel) engaged in low levels of compliance (*M*s = 4% and 9.8%, respectively). With 100% integrity phase, levels of compliance increased (*M* = 98.1%). Compliance decreased at 0% integrity (*M* = 14.5%). Compliance initially occurred at high levels, but decreased across the phases during 20% (*M* = 73.1%) and 40% (*M* =

70.6%) integrity. Compliance increased across the 60% phase to high levels (*M* = 69.3%). James engaged in high levels of compliance during the 80% and 100% phases (*M*s = 95.7% and 95.3%, respectively). Levels of compliance decreased at 0% integrity, (*M* = 29%). During 100% integrity, compliance increased (*M* = 97.8%). James's mean level of compliance was similar during the return to 80% (*M* = 97.8%), 60% (*M* = 98.0%), 40% (*M* = 97.7%), 20% (*M* = 98.0%), and 0% (*M* = 95.4%) phases. James continued to engage in high levels of compliance in the 100% integrity phase (*M* = 98.4%). During the final 0% integrity phase, high levels of compliance maintained (*M* = 97.4%).

Table 2 shows the mean rates of aggression aggregated for each procedural integrity phase for Kari, Nate, and James. Figure 5 displays levels of aggression during the parametric analysis for Kari (top panel), Nate (middle panel), and James (bottom panel). During baseline, Kari engaged in moderate levels of aggression (*M*s = 2.3 resp/min and 1.5 resp/min, respectively). During the 100% integrity phases, aggression decreased (*M*s = 0 resp/min and 0.1 resp/min, respectively). Aggression continued to occur at low levels at 80% (*M* = 0.1 resp/min) and 60% (*M* = 0.0 resp/min) integrity. During the 40%, 20%, and 0% integrity phases, levels of aggression increased from moderate to high levels across phase (*M*s = 1.8, 2.4, and 10.5 resp/min, respectively). During the 100% phase, aggression decreased to zero levels. Levels of aggression increased to high levels in the 0% (*M* = 10.5 resp/min) and 20% (*M* = 15.4 resp/min) integrity phases. Aggression decreased when treatment integrity increased to 40% (*M* = 0.01 resp/min), and aggression remained at zero levels across the 60%, 80%, and 100% phases.

During baseline, Nate (Figure 5, middle panel) engaged in moderate levels of aggression (*M*s = 2.1 resp/min and 2.7 resp/min, respectively). Aggression decreased to low levels during the 100% integrity phases (*M*s = 0.2 resp/min and 0.9 resp/min, respectively). When levels of procedural integrity decreased, aggression occurred at low levels across the 80%, 60%, and 40% phases (*M*s = 0.1, 0, and 0.04 resp/min, respectively). During the 20% integrity phase, aggression increased slightly at the end of the phase (*M* = 0.8 resp/min). In the 0% phase, aggression occurred at high levels (*M* = 4.2 resp/min). Aggression decreased to low levels (*M* = 0.2 resp/min) with reintroduction of 100% integrity. When returning to 0% integrity, aggression increased to high levels (*M* = 3.7 resp/min). Aggression decreased in the return to 20% integrity (*M* = 0.5 resp/min) and remained at similar levels at 40% integrity (*M* = 0.6 resp/min). Nate engaged in low levels of aggression during the 60% (*M* = 0.1 resp/min), 80% (*M* = 0.01 resp/min), and 100% (*M* = 0.04 resp/min) phases.

James (Figure 5, bottom panel) engaged in moderate levels of aggression during baseline (*M*s = 3.8 resp/min and 2.7 resp/min, respectively). Aggression decreased to zero levels with 100% integrity. During the 0% integrity phases, levels of aggression increased (*M* = 2.9 resp/min). In the 20% phase, aggression occurred at low levels (*M* = 0.1 resp/min). With the implementation of 40% integrity, aggression was initially low and increased across the phase (*M* = 1.8 resp/min). During the 60% integrity phase, aggression decreased to low levels (*M* = 1.1 resp/min). Aggression occurred at zero levels with 80% and 100% integrity. When integrity returned to 0%, aggression occurred at higher levels (*M*= 2.2 resp/min). Aggression decreased to zero levels with the return to 100% integrity. Aggression remained at zero levels across the remaining phases, including the return to 80%, 60%, 40%, 20%, and 0% integrity.

Discussion

The present study examined the parametric effects of one type of omission error (i.e., failure to deliver a token following compliance) on percentages of compliance and rates of aggression during a token economy for children with disabilities who engaged in aggression maintained by escape from demands. Participants completed token training, which consisted of response-stimulus pairing, a reinforcer assessment, and schedule thinning. With the implementation of the token economy (FR or VR token production) at 100% integrity, participants engaged in decreased levels of problem behavior and increased levels of compliance.

Greg, Kyle, and Ray were assigned to a token economy that used a FR1 schedule of token production. High levels of compliance and low levels of aggression occurred in the 100% integrity conditions. During the 80% integrity condition, Greg, Kyle, and Ray had moderate to high levels of compliance; however, levels of aggression differed across participants. For example, Greg's aggression remained low during this phase, and Kyle and Ray had slight increases during one of the exposures to 80% integrity. Lower levels of compliance and increases in aggression occurred in the 60% and 40% integrity conditions for Greg, Kyle, and Ray. During 20% integrity, Greg's compliance decreased rapidly, and Kyle and Ray's level of compliance corresponded to levels seen in the previous phase (either 40% or 0% integrity). In the 0% integrity phase, Greg, Kyle, and Ray engaged in the lowest levels of compliance and the highest levels of aggression.

Kari, Nate, and James were assigned to token economies with a VR 3 schedule of token production. High levels of compliance and low levels of aggression occurred in the 100%, 80%, and 60% integrity conditions. The first exposure to 40% integrity resulted in decreased levels of compliance and increased levels of aggression for Kari and James, while Nate continued to have increased levels of compliance and lower levels of aggression. The second reiteration of 40% resulted in higher levels of compliance and no occurrences of aggression for Kari and James, while Nate had variable levels of compliance and increased aggression. During 20% integrity, all three participants engaged in decreased levels of compliance and increased aggression; however, this pattern did not replicate during James' second exposure to 20% integrity. Finally, the first and second exposure to 0% integrity led to the lowest levels of compliance and elevated aggression for Kari, Nate, and James. However, James continued to engage high levels of compliance during additional exposures to 0% integrity.

The present study adds to the literature on the use of DRA to decrease problem behavior and increase appropriate behavior. Past experimental and translation studies have evaluated challenges to a procedural variation of DRA, FCT, to target problem behavior (Mace et al., 2010; Nevin et al., 2016; Pritchard et al., 2014; Vollmer et al., 1999). Frequently, researchers train a functional communication response to teach the participant to request the functional reinforcer in an appropriate manner. For example, Nevin et al. (2016) taught a participant to engage in a communicative response to ask for a preferred tangible item. The authors evaluated the influence of differing schedules of reinforcement following the participant's mand (i.e., request) for the restricted item. Additionally, Vollmer et al. (1999) conducted a parametric evaluation of treatment failures (i.e., blended errors of reinforcing problem behavior and failing to reinforce the FCR) during FCT with individuals with problem behavior. The current investigation targeted compliance with adult directives as the alternative behavior instead of teaching a functional communicative response. Compliance is a socially significant behavior for stakeholders (e.g., parents, teachers, caregivers) with school-aged children with escape-maintained behavior. Relatively little research has looked at the effects of treatment integrity errors during DRA interventions that target compliance. This study evaluated one treatment integrity error (failing to deliver a token following compliance). However, given that DRA interventions targeting compliance may be important for individuals entering or attending school, future research should evaluate different treatment integrity errors that could occur. For example, future research could evaluate the failure to allow the participant to exchange the tokens for backup reinforcers or the delivery of tokens independent of compliance.

In the current study, we used a DRA intervention that incorporated a token economy to target compliance and used the functional reinforcer of escape in combination with an edible or tangible as the backup reinforcer. This token economy intervention was able to maintain low rates of aggression, while increasing and stabilizing levels of compliance with six participants when levels of integrity were at 100%. When levels of integrity decreased, beginning at 80% (for participants in the FR token economy) and 60% integrity (for participants in the VR token economy), levels of compliance also decreased. All participants engaged in increased levels of aggression as level of integrity decreased, suggesting that omission errors in the form of failures to deliver a token for compliance influenced levels of problem behavior. These increases in problem behavior were observed even though extinction was in place during the parametric evaluation. These findings diverge from St. Peter Pipkin et al. (2010), who concluded that omission errors (failing to reinforce appropriate behavior) were not associated with high levels of inappropriate behavior in their translation study.

A goal of the present study was to provide descriptive data on the effects of the token-production schedules on problem behavior and compliance. One benefit of a token economy is that it allows participants to access breaks during consumption of the backup reinforcer, while still including a contingency for compliance. In addition, a token economy is associated with delays to reinforcement in that participants must meet an exchange requirement prior to trading in for the backup reinforcer. The present study demonstrated that participants in a FR 1 token production token economy were sensitive to omission errors beginning at 80%; two participants engaged in lowered levels of compliance and increased rates of aggression. However, we did not see consistent treatment degradation (increases in aggression or decreases in compliance) with the participants in the VR 3 token production until 40% integrity, suggesting that a variable schedule of reinforcement may buffer the negative impacts of an omission error (failure to deliver a token) at moderate levels of integrity. In fact, two participants in the VR 3 token economy, Kari and James, engaged low levels of aggression and high levels of compliance in phases of degraded integrity during the second or third exposures (40% for Kari and 60%, 40%, 20%, and 0% for James) even though prior reiterations of these integrity levels were associated with increased aggression and low compliance. However, any conclusions about the interactive effects of different schedules of reinforcement and treatment integrity errors are limited in the current study given that these observations are across different groups of participants. Future studies should replicate these procedures using a within-subject approach. For example, replications of FR and VR schedules of token reinforcement within participants may provide information about the effectiveness variable schedules of reinforcement to reduce the negative impact of treatment integrity errors.

This study replicated the results of prior research showing that the discontinuation of reinforcement for the alternative response (e.g., compliance) in a DRA intervention and continuing extinction for problem behavior results in resurgence (e.g., Craig & Shahan, 2016; Mace et al., 2010; Nevin et al., 2016; Pritchard et al., 2014; Smith et al., 2017). Typically, resurgence studies consist of three phases: (1) target response (e.g., problem behavior, clicking mouse, pressing a lever) is reinforced; (2) an alternative response (e.g., appropriate behavior) is reinforced while the target response is placed on extinction; (3) reinforcers for both responses are discontinued. Previous investigations found that problem behavior occurs at levels higher than baseline during the resurgence phase. In the current study, a test phase for resurgence occurred during the 0% integrity phase, during which problem behavior and compliance were placed on extinction (i.e., at 0% integrity, compliance was never followed by the delivery of a token). Resurgence of problem behavior occurred with all participants during at least one 0% integrity phase. Five participants (Greg, Kyle, Ray, Kari, and Nate) engaged in similar or higher levels of aggression during 0% integrity compared to baseline when aggression was reinforced with escape from academic tasks. One participant, James, showed a resurgence of aggression during his first and second exposure to the 0% integrity condition; however, zero levels of aggression occurred during the third and fourth exposure to 0% integrity.

The theory of resurgence supposes that higher rates of alternative reinforcement in treatment will result in quicker decreases in problem behavior, yet greater resurgence of problem behavior when treatment is discontinued. In addition, it also hypothesized that leaner rates of reinforcement result in slower decreases during treatment, but is less likely to relapse when treatment is ceased (Craig & Shahan, 2017; Smith et al., 2017). Particularly, the reinforcer rate for the alternative behavior appears to be critical in terms of treatment relapse. However, there has been discrepancy between studies that find resurgence following lean rates (e.g., Reed & Clark, 2011) and high rates (e.g., Mace et al., 2010; Nevin et al., 2016; Pritchard et al., 2014) of reinforcement in a DRA. Within the current investigation, resurgence of problem behavior occurred with similar patterns (with the exception of the third and fourth exposure for James) across continuous and variable schedules of token delivery.

One limitation of the current study is that we focused on participants who engaged in aggression maintained by escape from demands. It is possible that the current study's results may not generalize to individuals with different functions of behavior (e.g., attention-maintained behavior). Token economies can be arranged using different backup reinforcers depending on the function of the participant's problem behavior. This study incorporated an edible or tangible with a 30-s break from demands as a backup reinforcer when trading in tokens for participants with escape-maintained behavior as well as extinction for problem behavior (i.e., prompting through the demand). Yet, participants with attention-maintained behavior may earn different types of backup reinforcers, such as engaging in a preferred conversation topic or physical attention (e.g., high fives or tickles). It is possible that failing to deliver a token (omission error) during a token economy for individuals with attention-maintained behavior may be even more detrimental, because this error would involve the omission of the token as well as the omission of the attention (e.g., praise) that is paired with token delivery. Future studies could examine omission errors (e.g., failing to delivering a token, failing to deliver praise paired with a token) or blended errors (e.g., pairing a token with attention following problem behavior) with individuals with attention-maintained behavior during a token economy. In addition, it is unknown how individuals with attention-maintained behavior respond during token economy treatment relapses during work activities. Said differently, we do not know how omission errors influence compliance for individuals with attention-maintained behavior while completing academic work. Future research should continue to investigate how problem behavior function and treatment integrity errors in token economies may interact and influence appropriate behavior and problem behavior.

Another limitation of the current study is the backup reinforcer duration during the parametric analysis. The current study conducted edible and tangible preference assessments with participants to determine high and moderate-preferred items to use for backup reinforcers within the token economy. Extant research shows that delayed reinforcement schedules can be influenced by differing reinforcer dimensions, including duration and quality of reinforcement (e.g., DeLeon et al., 2014). For example, results from DeLeon et al. (2014) and Kocher et al. (2015) established that an item's (e.g., access to video games) reinforcing effectiveness is influenced by the continuity of access. Participants earned tokens to be exchanged for access to preferred leisure activities (e.g., playing Legos), with each token worth 30-s access to the leisure item. Under distributed exchange-production schedules, participants could immediately access their preferred leisure activities (i.e., a token was immediately exchanged), whereas during accumulated exchange-production schedules, participants could enter into an exchange period when they had collected 10 tokens. Participants preferred continuous access to reinforcers over discontinuous access even though accumulated exchange-production schedule was associated with delays to reinforcement. In the current study, participants exchanged 5 tokens for 30-s access to the backup reinforcer and participants could not accumulate tokens to earn more time with the backup reinforcer. Though data were not collected on participant selection of backup reinforcers for each session, anecdotally, it appeared participants would request to work for edibles rather than tangible items. Thus, the continuity of reinforcer access could be an important dimension of reinforcement that needs further investigation in token-economy procedures. In addition, future research should evaluate the influence of omission errors on the effectiveness of token economies under differing arrangements of the backup reinforcement dimension. Perhaps, omission errors influence token economy effectiveness differently if the participant exchanges for distributed or accumulated access to the backup reinforcer. Future investigations should look at this parameter of reinforcement when evaluating integrity errors during token economies.

Another potential limitation is that the rate of reinforcement was not specifically programmed and was based on the rate at which the participant complied and the current reinforcement schedule. Therefore, the reinforcement rate was not specifically programmed to be higher across the fixed or variable-schedules participant groups, meaning that there was not a specific ratio of token reinforcer difference between the two groups. Typically, investigations evaluating differences of responding under different schedules of reinforcement program a specific alternative behavior reinforcement ratio (e.g., 4:1 or 3:1; Nevin et al., 2016; Smith et al., 2017). Comparisons of responding under FR and VR schedules of token production have not been studied in clinical populations, even though these token production schedules are common in applied studies (Adcock & Cuvo, 2009; Becraft & Rolider, 2015; Breyer & Allen, 1975; Carnett et al., 2014; Everett et al., 1974; Field et al., 2004; Gilley & Ringdahl, 2014; Tarbox et al., 2002; Tarbox et al., 2006). The current investigation was a way to begin understanding how responding occurs under different schedules of reinforcement with an intervention commonly used in clinical settings, including schools, homes, and clinics. Subsequent research can program specific token-production ratios between rich and lean schedules within subjects to evaluate how omission errors are sensitive to this particular reinforcement parameter.

In summary, determination of how levels of integrity influence rates of compliance and levels of problem behavior continues to be important for improving the functioning of individuals who engage in problem behavior to escape demands. The purpose of the current investigation was to begin to understand the degree to which varying levels of treatment integrity (100%, 80%, 60%, 40%, 20%, and 0%) affect the efficacy of token economies across continuous and variable schedules of reinforcement on appropriate and problem behavior. The results provided evidence that the use of token economies can effectively maintain low rates of problem behavior, while increasing rates of compliance. In addition, the current findings demonstrated treatment relapse, or resurgence, occurred in a behavioral intervention, the token economy, that programs for delays and compliance, while prior investigations have used functional communication training. Findings suggest that for individuals with problem behavior maintained by social-negative reinforcement, a token economy intervention using a FR 1 token production must be implemented with relatively high integrity (100%). If individuals have escape-maintained problem behavior and a token economy with a VR schedule is used, it appears that omission errors are less detrimental until 40% integrity. Future research should continue to look at integrity errors and its relationship to varying parameters of reinforcement (schedules, delays to exchange, continuity of access) to ensure that the risk of treatment relapse is minimal and to help inform the levels of integrity with which interventions must be implemented by teachers, parents, and other caregivers.

References

Alstot, A. E. (2015). Effects of a peer-administered token economy on second grade physical education students' overhand throw performance. *The Physical Educator*, 72, 245-265. Retrieved from http://js.sagamorepub.com/pe/article/view/6346

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bancroft, S. L., & Bourret, J. C. (2008). Generating variable and random schedules of reinforcement using Microsoft Excel macros. *Journal of Applied Behavior Analysis*, *41*, 227–235. doi: 10.1901/jaba.2008.41–227

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1–21. doi: 10.1002/jaba.30

Becraft, J. L., & Rolider, N. U. (2015). Reinforcer variation in a token economy. *Behavioral Interventions*, *30*, 157–165. doi: 10.1002/bin.1401

Bellone, K. M., Dufrene, B. A., Tingstrom, D. H., Olmi, D. J., & Barry, C. (2014). Relative efficacy of behavioral interventions in preschool children attending Head Start. *Journal of Behavioral Education*, *23*, 378. doi: 10.1007/s10864-014-9196-6

Blumberg, S. J., Bramlett, M. D., Kogan, M. D., Schieve, L. A., Jones, J. R., & Lu, M. C. (2015). Changes in prevalence of parent–reported autism spectrum disorder in school–aged U.S. children: 2007 to 2011–2012. *National Health Statistics Reports, 65,* 1–11.

Boyle C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., Visser, S., & Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in U.S. children, 1997–2008. *Pediatrics, 127,* 1034–1042. doi: 10.1542/peds.2010–2989

Breyer, N. L., & Allen, G. J. (1975). Effects of implementing a token economy on teacher attending behavior. *Journal of Applied Behavior Analysis, 8*, 373–380. doi: 10.1901/jaba.1975.8–373

Brigham, T. A., Finfrock, S. R., Breunig, M. K., & Bushell, D. (1972). The use of programmed materials in the analysis of academic contingencies. *Journal of Applied Behavior Analysis, 5,*177–182. doi: 10.1901/jaba.1972.5–177

Capriotti, M. R., Brandt, B. C., Ricketts, E. J., Espil, F. M., & Woods, D. W. (2012). Comparing the effects of differential reinforcement of other behavior and response-cost contingencies on tics in youth with Tourette syndrome. *Journal of Applied Behavior Analysis*, *45*, 251–263. doi:10.1901/jaba.2012.45-251

Carnett, A., Raulston, T., Lang, R., Tostanoski, A., Lee, A., Sigafoos, J., & Machalicek, W. (2014). Effects of a perseverative interest–based token economy on challenging and on–task behavior in a child with autism. *Journal of Behavioral Education*, *23*, 368–377. doi: 10.1007/s10864-014-9195-7

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

Carroll, R. A., Kodak, T., & Adolf, K. J. (2016). Effect of delayed reinforcement on skill acquisition during discrete–trial instruction: Implication for treatment–integrity errors in academic settings. *Journal of Applied Behavior Analysis, 49*, 1–6. doi: 10.1002/jaba.268

Carroll, R. A., Kodak, T., & Fisher, W. W. (2013). An evaluation of program treatment–integrity during discrete–trial instruction. *Journal of Applied Behavior Analysis, 46*, 375–394. doi: 10.1002/jaba.49

Cavalari, R. S., DuBard, M., & Luiselli, J. K. (2014). Simplified habit reversal and treatment fading for chronic skin picking in an adolescent with autism. *Clinical Case Studies*, *13*, 190-198. doi:10.1177/1534650113510348

Charlop–Christy, M. H., & Haymes, L. K. (1998). Using objects of obsession as token reinforcers for children with autism. *Journal of Autism and Developmental Disorders, 28*, 189–198. doi: 10.1023/A:1026061220171

Craig, A. R., & Shahan, T. A. (2016). Behavioral momentum theory fails to account for the effects of reinforcement rate on resurgence. *Journal of the Experimental Analysis of Behavior*, *105*, 375-392. doi: 10.1002/jeab.207

Conyers, C., Miltenberger, R., Maki, A., Barenz, R., Jurgens, M., Sailer, A., Haugen, M., & Kopp, B. (2004). A comparison of response cost and differential reinforcement of other behavior to reduce disruptive behavior in a preschool classroom. *Journal of Applied Behavior Analysis, 37,* 411–415. doi: 10.1901/jaba.2004.37–411

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

Cowdery, G. E., Iwata, B. A., & Pace, G. M. (1990). Effects and side effects of DRO as treatment for self–injurious behavior. *Journal of Applied Behavior Analysis*, *23*, 497–506. doi:10.1901/jaba.1990.23–497

Dickman, S. E., Bright, C. N., Montgomery, D. H., & Miguel, C. F. (2012). The effects of response interruption and redirection (RIRD) and differential reinforcement on vocal stereotypy and appropriate vocalizations. *Behavioral Interventions, 27*,185-192. doi:10.1002/bin.1348

DiCiano, P., & Everitt, B. J. (2004). Conditioned reinforcing properties of stimuli paired with self–administered cocaine, heroin or sucrose: Implications for the persistence of addictive behavior. *Neuropharmacology, 47,* 202–213. doi: 10.1016/j.neuropharm.2004.06.005

DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students' inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis*, *40*, 447–461. doi:10.1901/jaba.2007.40-447

DiGennaro Reed F. D., Reed D. D., Baez C. N., & Maguire, H. (2011). A parametric analysis of errors of commission during discrete–trial training. *Journal of Applied Behavior Analysis, 44*, 611–615. doi:10.1901/jaba.2011.44-611

DeLeon, I. G., Chase, J. A., Frank-Crawford, M. A., Carreau-Webster, A. B., Triggs, M. M., Bullock, C. E., & Jennett, H. K. (2014). Distributed and accumulated reinforcement arrangements: Evaluations of efficacy and preference. *Journal of Applied Behavior Analysis*, *47*, 293-313. doi: 10.1002/jaba.116

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple–stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*, 519–533. doi: 10.1901/jaba.1996.29–519

Dixon, M. R., Benedict, H., & Larson, T. (2001). Functional analysis and treatment of inappropriate verbal behavior. *Journal of Applied Behavior Analysis, 34*, 361–363. doi: 10.1901/jaba.2001.34–361

Donaldson, J. M., DeLeon, I. G., Fisher, A. B., & Kahng, S. (2014). Effects of and preference for conditions of token earn versus token loss. *Journal of Applied Behavior Analysis*, *47*, 537-548. doi: 0.1002/jaba.135

Dozier, C. L., Iwata, B. A., Thomason–Sassi, J., Worsdell, A. S., & Wilson, D. M. (2012). A comparison of two pairing procedures to establish praise as a reinforcer. *Journal of Applied Behavior Analysis, 45,* 721–735. doi: 10.1901/jaba.2012.45–721

Eluri, Z., Andrade, I., Trevino, N., & Mahmoud, E. (2016). Assessment and treatment of problem behavior maintained by mand compliance. *Journal of Applied Behavior Analysis*, *49*, 383-387. doi: 10.1002/jaba.296

Everett, P. B., Hayward, S. C., & Meyers, A. W. (1974). The effects of a token reinforcement procedure on bus ridership. *Journal of Applied Behavior Analysis*, *7*, 1–9. doi: 10.1901/jaba.1974.7-1

Field, C. E., Nash, H. M., Handwerk, M. L., & Friman, P. C. (2004). A modification of the token economy for nonresponsive youth in family-style residential care. *Behavior Modification*, *28*, 438-457. doi: 10.1177/0145445503258995

Fisher, W. W., Kuhn, D. E., & Thompson, R. H. (1998). Establishing discriminative control of responding using functional and alternative reinforcers during functional communication training. *Journal of Applied Behavior Analysis, 31,* 543–560. doi: 10.1901/jaba.1998.31–543

Fiske, K. E., Isenhower, R. W., Bamond, M. J., Delmolino, L., Sloman, K. N., & LaRue, R. H. (2015). Assessing the value of token reinforcement for individuals with autism. *Journal of Applied Behavior Analysis, 48,* 448–453. doi: 10.1002/jaba.207

Flynn, S. D., & Lo, Y. Y. (2016). Teacher implementation of trial-based functional analysis and differential reinforcement of alternative behavior for students with challenging behavior. *Journal of Behavioral Education*, *25*, 1-31. doi: 10.1007/s10864-015-9231-2

Fox, D. K., Hopkins, B. L., & Anger, W. K. (1987). The long-term effects of a token economy on safety performance in open-pit mining. *Journal of Applied Behavior Analysis*, *20*, 215–224. doi:10.1901/jaba.1987.20-215

Friman, P. C. (1990). Nonaversive treatment of high-rate disruption: Child and provider effects. *Exceptional Children*, *57*, 64-69. doi: 10.1177/001440299005700108

Geiger, K. B., Carr, J. E., & LeBlanc, L. A. (2010). Function-based treatments for escape-maintained problem behavior: A treatment-selection model for practicing behavior analysts. *Behavior Analysis in Practice*, *3* (1), 22-32

Gilley, C., & Ringdahl, J. E. (2014). The effects of item preference and token reinforcement on sharing behavior exhibited by children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 8,*1425–1433. doi: 10.1016/j.rasd.2014.07.010

Gillis, J. M., & Pence, S. T. (2015). Token economy for individuals with autism spectrum disorder. In F.D. Reed & D.D. Reed (Eds.), *Autism Service Delivery* (pp. 257–277). Springer New York.

Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Effects of conditioning voices as reinforcers for listener responses on rate of learning, awareness, and preferences for listening to stories in preschoolers with autism. *Analysis of Verbal Behavior, 27*(1), 103–124.

Gresham F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review*, *18*(1), 37–50.

Hackenberg, T. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior, 91*, 257–286. doi: 10.1901.jeab.2009.91–257

Hagopian, L. P., Kuhn, S. A. 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*, 177–193. doi: 10.1901/jaba.2005.43–04

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185. doi:10.1901/jaba.2003.36–147

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis*, *4,* 17–38. doi: 1 0.1901/jaba.2001.34–17

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

Hanney, N. H. (2015). *Conditioning peers as reinforcers and the effects on mand training with preschool-aged children*. (Unpublished doctoral dissertation). Auburn University, Auburn, AL.

Himle, M. B., Woods, D. W., & Bunaciu, L. (2008). Evaluating the role of contingency in differentially reinforced tic suppression. *Journal of Applied Behavior Analysis, 41,* 285–289. doi:10.1901/jaba.2008.41–285

Hobbs, T. R., & Holt, M. M. (1976). The effects of token reinforcement on the behavior of delinquents in cottage settings. *Journal of Applied Behavior Analysis, 8*, 189–198. doi:10.1901/jaba.1976.8–189

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)

Ivy, J. W., Neef, N. A., Meindl, J. N., & Miller, N. (2016). A Preliminary Examination of Motivating Operation and Reinforcer Class Interaction. *Behavioral Interventions*, *31*, 180-194. doi: 10.1002/bin.1436

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,* 197–209. (Reprinted from *Analysis and Intervention in Developmental Disabilities*, *2,* 3–20, 1982) doi: 10.1901/jaba.1994.27-197

Iwata, B. A., & Dozier, C. L. (2008). Clinical application of functional analysis methodology. *Behavior Analysis in Practice*, *1*(1), 3.

Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis*, *27*, 131–144. doi: 10.1901/jaba.1994.

Jarmolowicz, D. P. & Lattal, K. A. (2013). Delayed reinforcement and fixed–ratio performance. *Journal of the Experimental Analysis of Behavior*, *100,* 370–395. doi: 10.1002/jeab.48

Jenkins, S. R., Hirst, J. M., & Reed, F. D. D. (2015). The effects of discrete-trial training commission errors on learner outcomes: An extension. *Journal of Behavioral Education*, *24*, 196-209. doi: 10.1007/s10864-014-9215-7

Johnston, J. M., & Johnston, G. T. (1972). Modification of consonant speech–sound articulation in young children. *Journal of Applied Behavior Analysis, 5*, 233–246. doi:10.1901/jaba.1972.5–233

Jowett Hirst, E. S., Dozier, C. L., & Payne, S. W. (2016). Efficacy of and preference for reinforcement and response cost in token economies. *Journal of Applied Behavior Analysis*, *49*, 329-345. doi: 10.1002/jaba.294

Kahng, S., Boscoe, J. H., & Byrne, S. (2003). The use of an escape contingency and a token economy to increase food acceptance. *Journal of Applied Behavior Analysis, 36*, 349–353.doi:10.1901/jaba.2003.36–349

Karraker, R. J. (1977). Self versus teacher selected reinforcers in a token economy. *Exceptional Children, 43,* 454–455. doi:10.1177/001440297704300708

Kazdin, A. E., & Bootzin, R. R. (1972). The token economy: An evaluative review. *Journal of Applied Behavior Analysis*, *5*, 343–372. doi: 10.1901/jaba.1972.5-343

Kocher, C. P., Howard, M. R., & Fienup, D. M. (2015). The effects of work-reinforcer schedules on skill acquisition for children with autism. *Behavior Modification*, *39*, 600-621. doi: 10.1177/0145445515583246

Krentz, H., Miltenberger, R., & Valbuena, D. (2016). Using token reinforcement to increase walking for adults with intellectual disabilities. *Journal of Applied Behavior Analysis*, *49*, 745-750. doi: 10.1002/jaba.326

Lahey, B. B., & Drabman, R. S. (1974). Facilitation of the acquisition and retention of sight–word vocabulary through token reinforcement. *Journal of Applied Behavior Analysis*, *7*, 307–312. doi:10.1901/jaba.1974.7–307

LeBlanc, L. A., Hagopian, L. P., & Maglieri, K. A. (2000). Use of a token economy to eliminate excessive inappropriate social behavior in an adult with developmental disabilities. *Behavioral Interventions, 15,* 135–143. doi:10.1002/(SICI)1099-078X(200004/06)15:2<135::AID-BIN51>3.0.CO;2-3

Lecavalier, L. (2006). Behavioral and emotional problems in young people with pervasive developmental disorders: Relative prevalence, effects of subject characteristics, and empirical classification. *Journal of Autism and Developmental Disorders*, *36,* 1101–1114. doi: 10.1007/s10803-006-0147-5

Lennox, D. B., Miltenberger R. G., Spengler, P., & Erfanian, N. (1988). Decelerative treatment practices with persons who have mental retardation: A review of five years of the literature. *American Journal on Mental Retardation, 92*(6), 492–501.

Leon, Y., Borrero, J. C., & DeLeon, I. G. (2016). Parametric analysis of delayed primary and conditioned reinforcers. *Journal of Applied Behavior Analysis*, *49*, 639-655. doi: 10.1002/jaba.311

LePage, J. P., DelBen, K., Pollard, S., McGhee, M., VanHorn, L., Murphy, J., & Moggie, N. (2003). Reducing assaults on an acute psychiatric unit using a token economy: A 2–year follow–up. *Behavioral Interventions, 18,* 179–190. doi:10.1002/bin.134

Lerman, D. C., & Iwata, B. A. (1993). Descriptive and experimental analyses of variables maintaining self-injurious behavior. *Journal of Applied Behavior Analysis*, *26*, 293–319. doi: 10.1901/jaba.1993.26-293

Lerman, D. C., & Iwata, B. A. (1996). Developing a technology for the use of operant extinction in clinical settings: An examination of basic and applied research. *Journal of Applied Behavior Analysis*, *29*, 345-382. doi: 10.1901/jaba.1996.29-345

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*, 349-367. doi: 10.1901/jeab.2010.93-349

Massetti, G. & Fabiano, G. (2005). Generalized conditioned reinforcer. In M. Hersen J. Rosqvist & A. M. Gross (Eds.), *Encyclopedia of behavior modification and cognitive behavior therapy* (Vol. 3, pp. 867–867). Thousand Oaks, CA: Sage Publications Ltd. doi: 10.4135/9781412950534.n2060

Matson, J. L., & Boisjoli, J. A. (2009). The token economy for children with intellectual disability and/or autism: A review. *Research in Developmental Disabilities, 30*, 240–248. doi: 10.1016/j.ridd.2008.04.001

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua T., & Saxena, S. (2011). Prevalence of intellectual disability: a meta–analysis of population–based studies. *Research in Developmental Disabilities, 32*, 419–436. doi: 10.1016/j.ridd.2010.12.018

McLaughlin, T. F., & Malaby, J. (1972). Intrinsic reinforcers in a classroom token economy. *Journal of Applied Behavior Analysis*, *5*, 263–270. doi: 10.1901/jaba.1972.5-263

Miltenberger, R. (2012). *Behavior modification: Principles & procedures* (5th ed.). Belmont, CA: Wadsworth.

Mitteer, D. R., Romani, P. W., Greer, B. D., & Fisher, W. W. (2015). Assessment and treatment of pica and destruction of holiday decorations. *Journal of Applied Behavior Analysis, 48,* 912–917. doi: 10.1002/jaba.255.

Mullen, N. W., Maxwell, H., & Bedard, M. (2015). Decreasing driver speeding with feedback and a token economy. *Transportation Research Part F: Traffic Psychology and Behaviour*, 28, 77-85. doi:10.1016/j.trf.2014.11.008

Nevin, J. A., Mace, F. C., DeLeon, I. G., Shahan, T. A., Shamlian, K. D., Lit, K., ... & Tarver, D. R. (2016). Effects of signaled and unsignaled alternative reinforcement on persistence and relapse in children and pigeons. *Journal of the experimental analysis of behavior*, *106*, 34-57. doi: 10.1002/jeab.213

Olson, S. L., & Hoza, B. (1993). Preschool developmental antecedents of conduct problems in children beginning school. *Journal of Clinical and Child Psychology, 22*, 60–67. doi: 10.1207/s15374424jccp2201_6

Pace, G. A., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18,* 249–255. doi:10.1901/jaba.1985.18–249

Pence, S. T., Roscoe, E. M., Bourret, J. C., & Ahearn, W. H. (2009). Relative contributions of three descriptive methods: Implications for behavioral assessment. *Journal of Applied Behavior Analysis*, *42*, 425–446. doi: 10.1901/jaba.2009.42-425

Pence, S. T., & St. Peter, C. C. (2015). Evaluation of treatment integrity errors on mand acquisition. *Journal of Applied Behavior Analysis, 48*, 575–589. doi: 10.1002/jaba.238

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

Phillips, E. L. (1968). Achievement place: Token reinforcement procedures in a home–style rehabilitation setting for “pre–delinquent” boys. *Journal of Applied Behavior Analysis, 1*, 213–223. doi:10.1901/jaba.1968.1–213

Piazza, C. C., Moes, D. R., & Fisher, W. W. (1996). Differential reinforcement of alternative behavior and demand fading in the treatment of escape–maintained destructive behavior. *Journal of Applied Behavior Analysis*, *29*, 569–572. doi:10.1901/jaba.1996.29–569

Plavnick, J. B., Ferreri, S. J., & Maupin, A .N. (2010). The effects of self–monitoring on the procedural integrity of behavioral intervention for young children with developmental disabilities. *Journal of Applied Behavior Analysis, 43,*315–320. doi: 10.1901/jaba.2010.43–315

Plavnick, J. B., Thompson, J. L., Englert, C. S., Mariage, T., & Johnson, K. (2016). Mediating access to Headsprout® early reading for children with Autism Spectrum Disorders. *Journal of Behavioral Education*, *3*, 357-378. doi: 10.1007/s10864-015-9244-x

Pritchard, D., Hoerger, M., Mace, F. C., Penney, H., & Harris, B. (2014). Clinical translation of animal models of treatment relapse. *Journal of the Experimental Analysis of Behavior*, *101*, 442-449. doi: 10.1002/jeab.87

Reed, P., & Clark, C. (2011). Impact of intervening learning on resurgence in humans with Autism Spectrum Disorders. *Learning & Behavior*, *39*, 163-170. doi:10.3758/s13420-010-0014-2

Roane, H. S., Lerman, D. C. & Vorndran, C. M. (2001). Assessing reinforcers under progressive schedule requirements. *Journal of Applied Behavior Analysis, 34,* 145–167. doi: 10.1901/jaba.2001.34–145

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

Rooker, G. W., Iwata, B. A., Harper, J. M., Fahmie, T. A., & Camp, E. M. (2011). False-positive tangible outcomes of functional analyses. *Journal of Applied Behavior Analysis*, *44*, 737-745. doi: 10.1901/jaba.2011.44-737

Shillingsburg, M. A., Lomas, J. E., & Bradley, D. (2012). Treatment of vocal stereotypy in an analogue and classroom setting. *Behavioral Interventions, 27,* 151-163. doi:10.1002/bin.1340

Schmidt, J. D., Bednar, M. K., Willse, L. V., Goetzel, A. L., Concepcion, A., Pincus, S. M., ... & Bowman, L. G. (2016). Evaluating treatments for functionally equivalent problem behavior maintained by adult compliance with mands during interactive play. *Journal of Behavioral Education*, *2*, 169-187. doi: 10.1007/s10864-016-9264-1

Smith, B. M., Smith, G. S., Shahan, T. A., Madden, G. J., & Twohig, M. P. (2017). Effects of differential rates of alternative reinforcement on resurgence of human behavior. *Journal of the Experimental Analysis of Behavior*, *107*, 191-202. doi: 10.1002/jeab.241

Sran, S. K., & Borrero, J. C. (2010). Assessing the value of choice in a token system. *Journal of Applied Behavior Analysis, 43*, 553–557. doi:10.1901/jaba.2010.43–553

St. Peter, C. C., Byrd, J. D., Pence, S. T., & Foreman, A. P. (2016). Effects of treatment-integrity failures on a response-cost procedure. *Journal of Applied Behavior Analysis*, *49*, 308-328 doi: 10.1002/jaba.291

St. Peter Pipkin, C. S., 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*, 47–70. doi:10.1901/jaba.2010.43-47

Tarbox, R. S. F., Ghezzi, P. M., & Wilson, G. (2002). The effects of token reinforcement on attending in a young child with autism. *Behavioral Interventions, 21,* 155–164. doi: 10.1002/bin.213

Taylor, S. A., & Mudford, O. C. (2012). Improving behavior in a residential service for youth in drug and alcohol rehabilitation. *Behavioral Interventions, 27*, 109-128. doi:10.1002/bin.1342

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.

Toussaint, K. A., & Tiger, J. H. (2012). Reducing covert self-injurious behavior maintained by automatic reinforcement through a variable momentary DRO procedure. *Journal of Applied Behavior Analysis*, *45*, 179-184. doi: 10.1901/jaba.2012.45-179

Truchlicka, M., MacLaughlin, T. F., & Swain, J. C. (1998). The effects of token reinforcement and response cost of the accuracy of spelling performance with middle–school special education students with behavior disorders. *Behavioral Interventions, 13,* 1–10. doi: 10.1002/(SICI)1099-078X(199802)13:1<1::AID-BIN1>3.0.CO;2-Z

Vargo, K. K., & Ringdahl, J. E. (2015). An evaluation of resistance to change with unconditioned and conditioned reinforcers. J*ournal of Applied Behavior Analysis, 48,* 643-662. doi:10.1002/jaba.226

Vollmer, T. R., Roane, H. S., Ringdahl, J. E., & Marcus, B. A. (1999). Evaluating treatment challenges with differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis, 32*, 9–23. doi: 10.1901/jaba.1999.32–9

Wadsworth, J. P., Hansen, B. D., & Wills, S. B. (2015). Increasing compliance in students with intellectual disabilities using functional behavioral assessment and self-monitoring. *Remedial and Special Education*, *36*, 195-207. doi:10.1177/0741932514554102

Zlomke, K., & Zlomke, L. (2003). Token economy plus self-monitoring to reduce disruptive classroom behaviors. *The Behavior Analyst Today*, *4*, 177-182. doi: 10.1037/h0

Table 1

Aggregated Aggression Rate and Percentage of Compliance for Participants in Fixed-Ratio Token Economy Group

Participant	Greg		Kyle		Ray	
Condition	AGG (rpm)	COMP (%)	AGG (rpm)	COMP (%)	AGG (rpm)	COMP (%)
BL	1.9	26.1	2.0	19.3	1.4	11.7
100	0.02	85.3	0.03	76.8	0.01	51.7
80	0.4	80.7	0.2	68.0	0.6	38.9
60	0.2	44.4	0.5	35.3	0.3	31.8
40	1.0	30.0	1.0	20.4	0.6	25.6
20	1.2	12.8	0.4	27.5	0.3	26.3
0	1.5	25.5	2.4	20.7	0.7	19.2

Note. AGG indicates aggression, COMP indicates compliance with adult directives, rpm stands for responses per minute, and % stands for the percentage of directives with compliance.

Table 2

Aggregated Aggression Rate and Percentage of Compliance for Participants in Variable-Ratio Token Economy Group

Participant	Kari		Nate		James	
Condition	AGG (rpm)	COMP (%)	AGG (rpm)	COMP (%)	AGG (rpm)	COMP (%)
BL	1.8	7.1	2.5	19.4	3.3	6.9
100	0.2	83.8	0.3	69.4	0.0	97.5
80	0.1	89.4	0.01	92.7	0.0	96.8
60	0.0	90.3	0.05	94.0	0.6	83.7
40	0.9	61.2	0.3	78.6	0.9	87.2
20	5.0	27.4	0.5	86.1	0.1	70.0
0	10.5	10.2	3.0	37.0	1.3	59.1

Note. AGG indicates aggression, COMP indicates compliance with adult directives, rpm stands for responses per minute, and % stands for the percentage of directives with compliance.

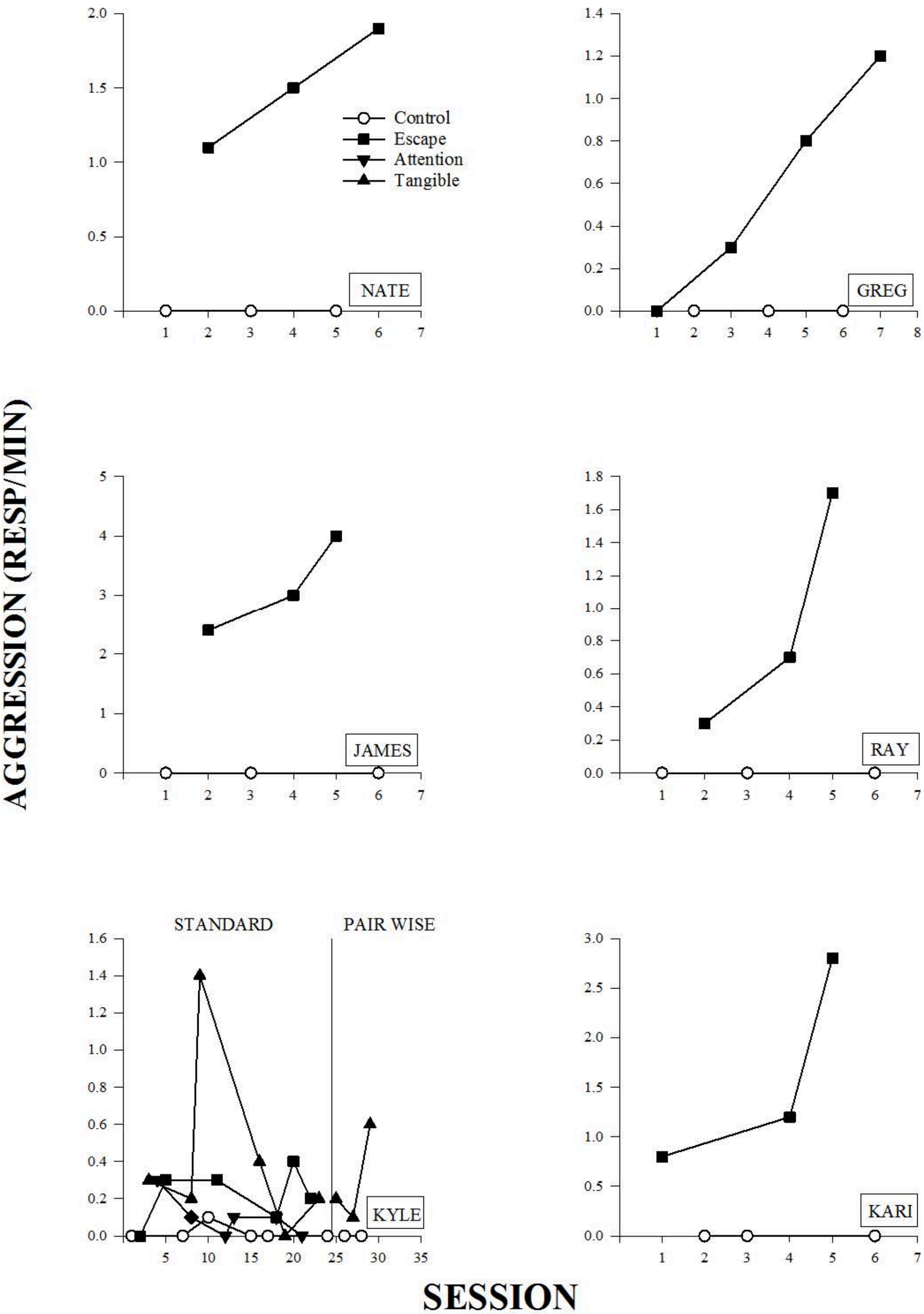

Figure 1. Rates (resp/min) of aggression across sessions during functional analyses for Nate, Greg, James, Ray, Kyle, and Kari

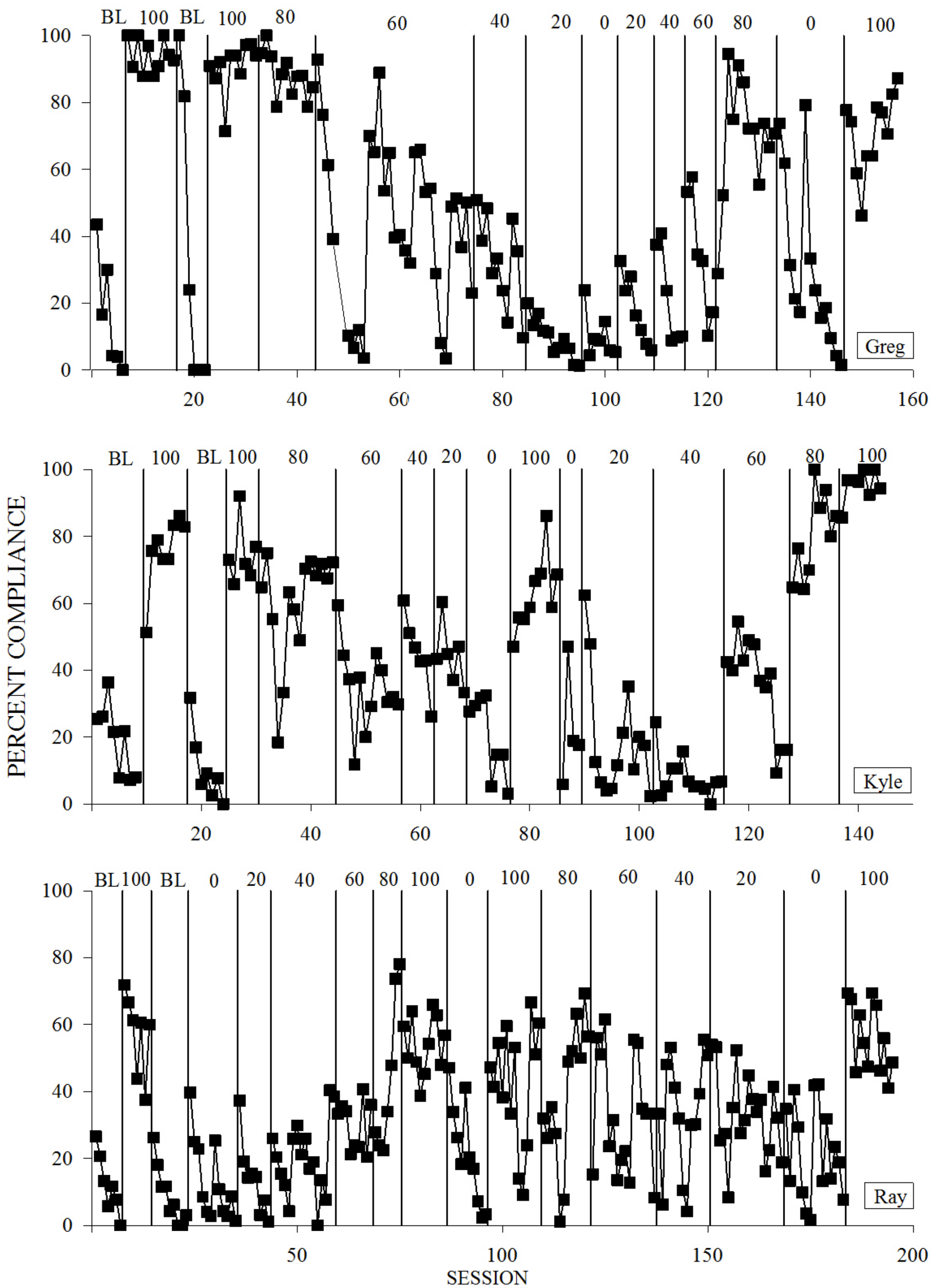

Figure 2. Percentage of compliance across sessions during the parametric treatment analysis for Greg, Kyle, and Ray.

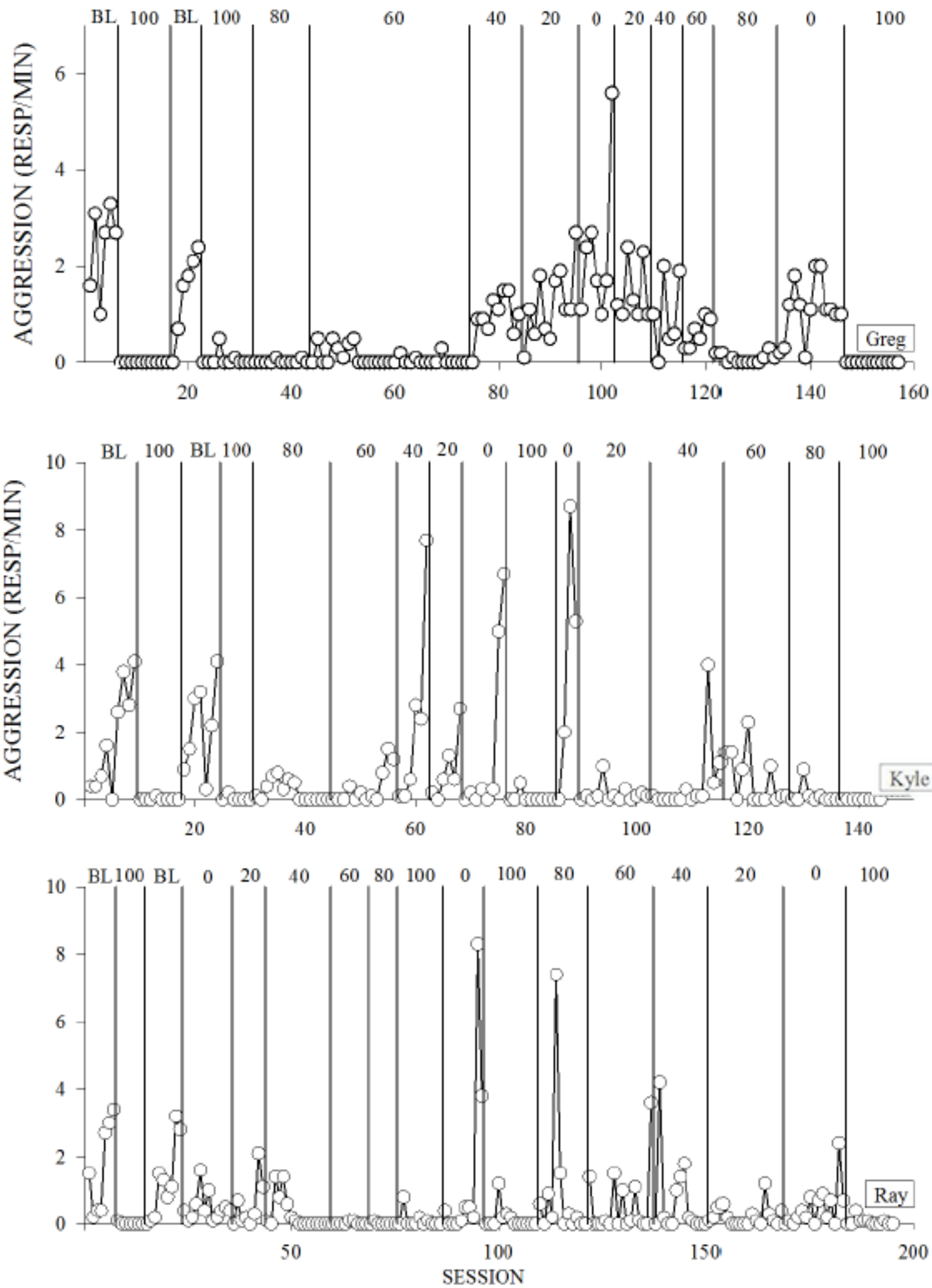

Figure 3. Rate (resp/min) of aggression across sessions during the parametric treatment analysis for Greg, Kyle, and Ray.

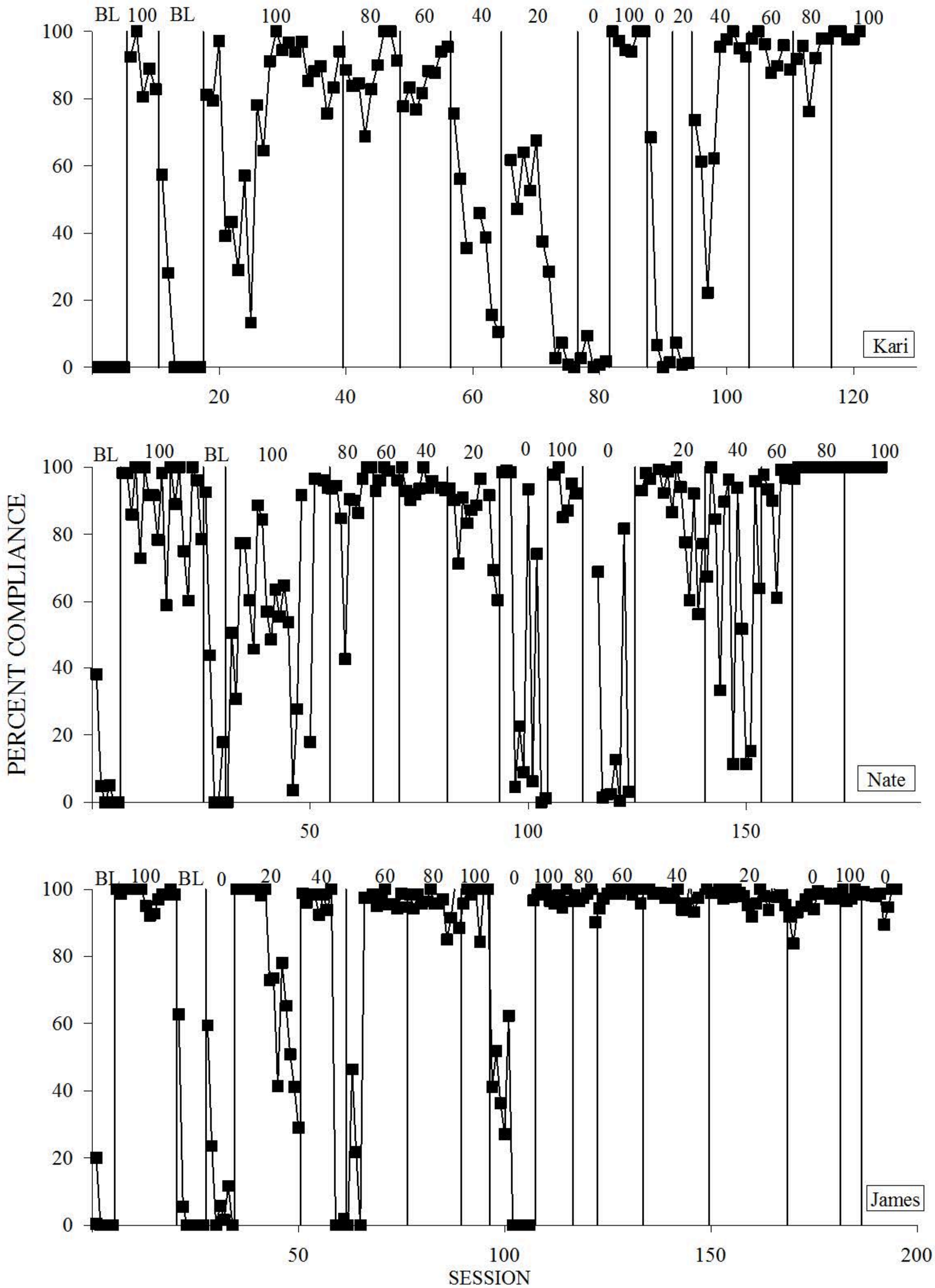

Figure 4. Percentage of compliance across sessions during the parametric treamtent analysis for Kari, Nate, and James.

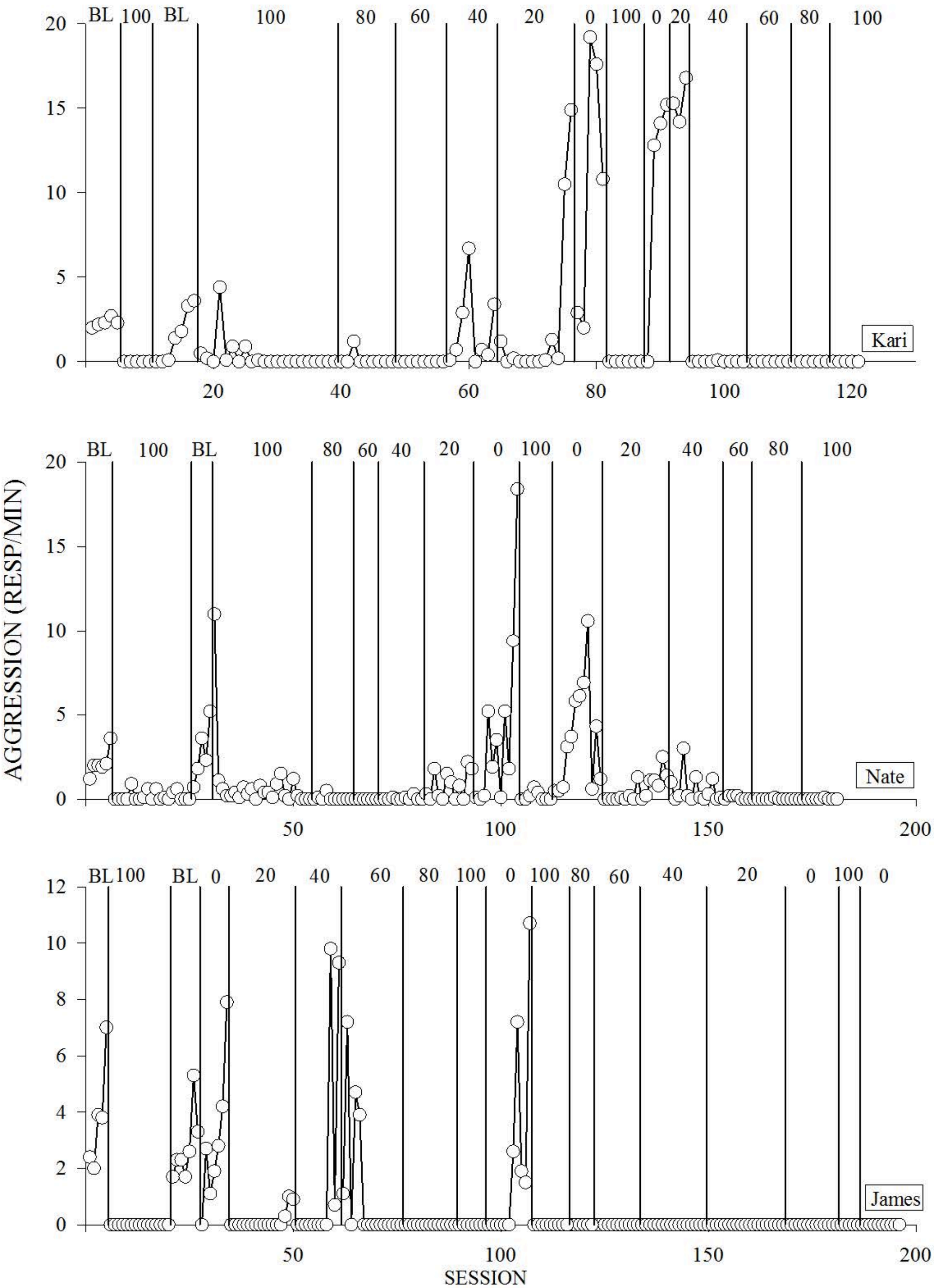

Figure 5. Rates (resp/min) of aggression during the parametric treatment analysis for Kari, Nate, and James.