USING PROGRESSIVE RATIO SCHEDULES TO EVALUATE EDIBLE, LEISURE, AND TOKEN REINFORCEMENT

Danielle M. Russell, B.S.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 2013

APPROVED:

Einar T. Ingvarsson, Major Professor
Manish Vaidya, Committee Member
Richard Smith, Committee Member and Chair of the Department of Behavior Analysis
Thomas Evenson, Dean of the College of Public Affairs and Community Service
Mark Wardell, Dean of the Toulouse Graduate School

Russell, Danielle M. Using Progressive Ratio Schedules to Evaluate Edible, Leisure, and Token Reinforcement. Master of Science (Behavior Analysis), May 2013, 55 pp., 15 figures, references, 51 titles.

The general purpose of the current study was to evaluate the potency of different categories of reinforcers with young children diagnosed with developmental delays. The participants were two boys and one girl who were between the ages of 7 and 8. In Phase 1, the reinforcing potency of tokens, edible items, and leisure items was evaluated by using a progressive ratio (PR) schedule. For two participants, tokens resulted in the highest PR break points. For one participant, edibles resulted in the highest break points (tokens were found to have the lowest break points). In Phase 2, the effects of presession access on the break points of edibles and tokens were examined. This manipulation served as a preliminary analysis of the extent to which tokens might function as generalized conditioned reinforcers. During Phase 2, presession access altered the break points of edibles, but not tokens. The findings of the current study suggest that PR schedules may be useful as a means to better assess certain dimensions of tasks and how they affect reinforcer effectiveness (e.g., amount of effort the client is willing to exert, the duration at which the client willing to work, how many responses the client will emit, etc.), and to evaluate to what extent tokens actually function as generalized conditioned reinforcers.

ACKNOWLEDGEMENTS

Foremost, I would like to thank Dr. Einar Ingvarsson for taking me on as his student. You are a fantastic advisor and I truly appreciate all of your guidance and support. I would also like to thank Dr. Richard Smith for allowing me to work under him as a part of BARC. You have provided me with such great learning opportunities and I have learned so much from working with you these past four years. To Dr. Manish Vaidya, I really appreciate you being on my thesis committee; you have such a remarkable perspective. I would also like to recognize all the staff at the Child Study Center. To my friends: Audrey Shivers, I could not have made it through graduate school without you. You are always there to listen, let me vent, and give me support and advice, even when the day is a -42. Nicole Quirk, all I can say is you are the wifey I never knew I needed but can never live without. You are so caring and thoughtful and without all your wonderful stories and fun mail surprises, I would live such a sad existence. Katy Atcheson, over the years we have become such great friends and I really appreciate all of your advice, especially our three hour phone conversations. Kimberly James-Kelly, you are the best IOA data collector a person could ask for and this thesis would have never been possible without all your help. We became such close friends on our long drives to CSC and I really miss those times. Lloyd Roberson, I am so grateful for your eternal optimism. You are an excellent listener and never fail to put a smile on my face and help me forget about all my troubles. Derrick Spear, you were my proofreading guru. I truly appreciate all the support you provided those first two years. Finally, I would like to acknowledge Ronald and Donna Russell. Dad, you are such a wise man and I can always count on you to give me perspective and advice. Mom, you are my emotional rock; you are always there to share my excitement or my tears. I want to thank-you from the bottom of my heart because without the two of you, I would not be the person I am today.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Considerable research efforts have been dedicated to developing methods to identify stimuli that will function as reinforcers (e.g., DeLeon & Iwata, 1996; Fisher et al., 1992; Roane, Vollmer, Ringdahl, & Marcus, 1998). Numerous studies have demonstrated that an effective reinforcer can expedite learning and maintain appropriate behavior (Baer, Peterson, & Sherman, 1967; Karsten & Carr, 2009; Olenick & Pear, 1980). Additionally, reinforcement is a widely used and accepted practice to increase behavior in applied settings. Therefore, the ability to efficiently identify and determine whether a stimulus is an effective reinforcer is of great clinical importance.

Typically, in order to identify whether a stimulus will function as a reinforcer, applied researchers and practitioners conduct a preference assessment to determine a relative ranking of stimulus preferences. When conducting a preference assessment, stimuli can be presented individually, in pairs, or in multiple stimulus formats (DeLeon & Iwata, 1996; Fisher et al., 1992; Pace, Ivancic, Edwards, Iwata, & Page, 1985; Roane, Vollmer, Ringdahl, & Marcus, 1998; Windsor, Piche, & Locke, 1994). The ranking is determined either by the amount of time spent engaging with the stimulus or by the number of times it was selected relative to the other stimuli. After preference is determined, a reinforcer assessment is usually conducted with mastered responses, during which each response produces access to a preferred stimulus (Fisher & Mazur, 1997; Piazza, Fisher, Hagopian, Bowman, & Toole, 1996). One limitation of using continuous reinforcement schedules during reinforcer assessment is that practitioners do not always use fixed ratio (FR) 1 schedules in applied settings; schedules of reinforcement in those contexts are often much leaner.

Several studies have directly evaluated various parameters that may influence the outcomes of preference and reinforcer assessments (DeLeon, Iwata, Goh, & Worsdell, 1997; Tustin, 1994). These studies identified schedule requirements as a factor that impacts reinforcer effectiveness. In a study with three adult males diagnosed with intellectual disabilities, Tustin (1994) found that a stimulus that was both preferred and effective at maintaining responding when schedule requirements were low was not necessarily equally preferred or effective when schedule requirements were increased. Tustin concluded that reinforcer potency is not static and is influenced by several variables such as the schedule requirements, the availability of alternative reinforcers, and the nature of alternative reinforcers. Likewise, DeLeon, Iwata, Goh, and Worsdell (1997) discovered that when two similar stimuli, previously found to be roughly equivalent in preference ranking, were available under concurrent FR 1 schedules, both participants (adults diagnosed with developmental disabilities) showed little or no differences in preference for either stimulus. However, when the schedule requirements increased to FR 5, a clear preference for one of the stimuli developed for both participants. This difference also maintained when the schedule requirements increased to FR 10.

Given that schedule requirements influence reinforcer potency, and the terminal goal of most applied efforts is to identify preferred stimuli that will maintain and increase responding over time, Roane, Lerman, and Vorndran (2001) suggested using progressive ratio (PR) schedules as a better means to assess reinforcer capability. In a PR schedule, the number of responses required to access reinforcement systematically increases within session. Escalation of PR schedules is typically based on prior responding (i.e., as more responses are emitted and/or more reinforcers are delivered, the number of responses required to produce reinforcement increases). In most cases, the session is terminated and the break point is obtained when there is

no responding for some period of time (2-5 min). With PR schedules, relative reinforcement effects can be identified by comparing the obtained break point for each stimulus (Hodos, 1961). One advantage of this method is that it allows for a relatively rapid evaluation of reinforcing efficacy under increasing response requirements.

In two experiments, Roane et al. (2001) evaluated the effectiveness of preferred stimuli by using PR schedules. Four individuals who were diagnosed with developmental disabilities participated in the first study and three of those same individuals participated in the second study. Roane et al. found that while response patterns of higher and lower preference stimuli failed to differentiate with FR 1 schedules, PR schedules produced differences in break points between various stimuli. They also discovered that stimuli with higher break points were more effective in the treatment of destructive behavior than those with lower break points. Jerome and Sturmey (2008) evaluated the interactions of adults diagnosed with intellectual disabilities and preferred versus non-preferred direct care staff in order to determine the extent to which these interactions were actually reinforcing. Using PR schedules, these authors found that interactions with preferred staff maintained responding under higher schedule requirements for all three participants. Penrod, Wallace, and Dyer (2008) used PR schedules to evaluate the reinforcing effects of high versus low preference stimuli with four children diagnosed with developmental disabilities. For three out of four participants, they found that higher preference stimuli generated higher break points. These combined outcomes suggest that PR schedules may be a better match for identifying effective reinforcers given certain tasks (e.g., if response maintenance under relatively lean ratio schedules is the goal).

Although direct delivery of preferred stimuli as reinforcers appears to be the most common method of reinforcing responses in skill acquisition programs and other behavioral

interventions, practitioners also use token economies with clients in many cases. Token economies are reinforcement systems that involve the contingent delivery of tokens following the occurrence of the targeted behavior (completing tasks, pro-social behavior, etc.). Accumulated tokens can later be exchanged for other items (e.g., food and toys); these are usually referred to as back-up reinforcers. By pairing tokens with items with known reinforcing effects (back-up reinforcers), the tokens may acquire conditioned reinforcing properties (Cooper, Heron, & Heward, 2007). Early investigations conducted in the laboratory evaluated token systems and the extent to which tokens functioned as conditioned reinforcers (Cowles, 1937; Kelleher, 1956; Kelleher, 1958; Wolfe, 1936). These studies demonstrated that poker chips could be established as conditioned reinforcers for chimpanzees.

If stimuli (e.g., tokens) are paired with a large variety of back-up reinforcers they may acquire generalized conditioned reinforcement properties (Cooper, Heron, & Heward, 2007). There are a number of advantages to using generalized conditioned reinforcers (Kazdin, 1972): (a) they bridge the delay between responses provided by the client and accessing a preferred item or activity, (b) they are less susceptible to the effects of satiation, (c) they maintain performance and responding over extended periods of time, (d) they provide minimal interruptions and allow for sequences of responses to be reinforced easily, and (e) they have the ability to be more potent than any single primary reinforcer; provided that the primary reinforcer(s) are in the array. Additionally, Ayllon and Azrin (1968a) described several advantages of using tangible conditioned reinforcers, such as tokens, instead of other generalized conditioned reinforcers (e.g., praise statements). Some of these benefits were (a) the number of tokens provided can relate easily and quantitatively to the amount of reinforcement earned, (b) tokens are portable, can remain with the client, and can be delivered easily across a variety of contexts and environments,

(c) there is no limit to the number of tokens one can earn, (d) tokens are durable and allow for consistency and standardization, and (e) tokens offer a visual representation of improvement and progress.

Dating back to the 1800s, researchers and clinicians have effectively used token systems as behavior management and motivational tools (Kazdin, 1978). In applied settings, verbal instructions that specify the prevailing contingencies and explain token procedures are often sufficient to condition tokens as reinforcers. However, when instructions alone are not effective, tokens can be established as conditioned reinforcers by providing clients with repeated opportunities to exchange them for the back-up reinforcers. Typically, a small number of tokens are given to the client, who is immediately given the opportunity to exchange them. This process is meant to establish the value of tokens. Once this value is conditioned, the client is instructed that he or she can earn tokens for engaging in certain behaviors. Usually, tokens are considered generalized conditioned reinforcers when they can effectively reinforce behavior across a wide range of motivating operation conditions (e.g., satiation) (Cooper, Heron, & Heward, 2007).

Most of the research on token economies to date has been conducted in the laboratory rather than in applied settings. We currently do not have a large literature base to fully understand how potent tokens are as reinforcers or to what extent tokens actually function as generalized conditioned reinforcers in the context of application (Hackenberg, 2009). Likewise, while much research has been conducted on the reinforcing effects of sensory, edible, and tangible stimuli; there have been few, if any, studies comparing and evaluating the reinforcing potency of edibles, leisure items, and tokens. Given that token economies are so commonly used and tokens are considered to be conditioned reinforcers, it is logical to compare the relative effectiveness of tokens with other items commonly used as reinforcement, such as edible and

leisure items. This line of research is especially important given the various costs and benefits of different reinforcement formats. Edibles are among the most common stimuli utilized due to the ease of delivery and quick consumption time. Nonetheless, these stimuli are not without their limitations. Given growing concerns about childhood obesity, diabetes, and related health concerns, many parents and physicians are apprehensive about the use of children's preferred edibles, which often consist of unhealthy treats, to be used as reinforcers.

The use of leisure items may be a desirable alternative to edibles, because they do not pose the same health concerns. However, using leisure items as reinforcers can be difficult. DeLeon, Iwata, and Roscoe (1997) found that in preference assessments with both edible and leisure items in the array of options, twelve out of fourteen individuals, who were diagnosed with mental retardation, showed a preference for edible items. Another limitation with leisure items is the magnitude required for the item to function as effective reinforcers. For many children, brief access to a movie is not sufficient to maintain effortful responding. Relatedly, it may be inconvenient or impossible to deliver access to leisure items or activities in applied contexts.

A common concern with edible and leisure reinforcers is the deleterious effects of satiation/habituation on the reinforcing potency of these stimuli (McAdam et al., 2005; Murphy, McSweeney, Smith, & McComas, 2003; Rincover & Newsom, 1985). One of the major advantages of tokens is that they should not be subject to these same effects of satiation, if they truly function as generalized conditioned reinforcers. However, in order for token systems to be resistant to satiation and habituation effects, a variety of items should be available as back-up reinforcers. An ability to identify a number of effective reinforcers would therefore pose a challenge to successful implementation of token economies. PR schedules can serve as one way to assess the extent to which conditioning procedures are successful in creating new reinforcers

or increasing the value of existing reinforcers, potentially facilitating the use of token economies. Given this, one purpose of the current study is to evaluate and compare the reinforcing potency of tokens, edibles, and leisure items by using a PR schedule.

Due to the importance of identifying effective and durable reinforcers for purposes of intervention and treatment, it is vital to examine the various factors, such as presession access, that may affect how stimuli function as reinforcers. Skinner suggested that through deprivation or satiation, behavior can be increased or decreased, respectively. There is a line of research suggesting that motivating operations (e.g., satiation and deprivation) play a critical role in the three term contingency. Vollmer and Iwata (1991) evaluated the influence of establishing operations on behavior with five adult males diagnosed with mental retardation. These researchers demonstrated that presession access (satiation) and no access (deprivation) to music, food, or social praise decreased (satiation) and increased (deprivation) responding maintained by these stimuli. Smith, Iwata, Goh, and Shore (1995) demonstrated a method for identifying establishing operations for problem behaviors maintained by escape. Nine adults with developmental disabilities and self-injury participated in the studies. The researchers manipulated various dimensions of tasks such as novelty, session duration, and the rate of trial presentation while keeping other antecedents and consequences constant. Smith et al. found that these variables can mediate the evocative effects of tasks for behavior maintained by negative reinforcement. Several studies also emphasized the importance of considering these variables when developing and interpreting behavioral assessments such as functional analyses of problem behavior (Iwata et al., 1994; Smith & Iwata, 1997; Worsdell, Iwata, Conners, Kahng, & Thompson, 2000) and preference assessments (Hagopian, Long, & Rush, 2004; McAdam et al., 2005).

The effects of establishing operations have been evaluated by examining how free access versus restricted access to a reinforcing stimulus immediately prior to sessions impacts responding (e.g., McComas, Thompson, & Johnson, 2003; Vollmer & Iwata, 1991). Interestingly, studies on this topic have produced conflicting results. Some found that presession access to a stimulus produced reduced levels of responding. For example, O'Reilly et al. (2007) assessed the effects of presession levels of attention during alone (no therapist present) and attention-extinction (therapist present but no attention delivered) conditions. One adult male diagnosed with autism served as the participant for this study. The researchers found much higher levels of problem behavior during those conditions in which no presession attention was delivered. A subsequent study conducted by O'Reilly et al. (2009) examined the effects of three different presession conditions - brief access, no access, and satiation - on the problem behavior of two children diagnosed with autism. Their results revealed the highest levels of problem behavior following brief access and no access presession conditions and the lowest levels following the satiation condition. Conversely, other studies found that presession access either did not produce any clear differentiations in responding (e.g., Gutierrez et al., 2007; O'Reilly et al., 2008; Vollmer & Iwata, 1991; Worsdell et al., 2000) or that it resulted in increased levels of responding (Roantree & Kennedy, 2006; Sy & Borrero, 2009). For instance, Sy and Borrero (2009) evaluated the impact short, medium, and long periods of presession access to edible and nonedible reinforcers had on response rates with three boys diagnosed with developmental disabilities. For one participant, they found that any presession access to an edible reinforcer decreased response rates. However, short and medium periods of presession access to nonedible reinforcers resulted in similar or increased response rates for two participants. While inconsistent results have been obtained with regards to the effects of presession access on edible and leisure

items, behavior analysts generally agree that the reinforcing effectiveness of these kinds of stimuli can be directly affected by motivational variables. By contrast, it is generally held that generalized conditioned reinforcers are not nearly as sensitive to these same types of motivational influences. Thus, if tokens function as generalized conditioned reinforcers they should show resilience to potential motivating influences such as satiation. Given this, a second purpose of this study is to evaluate the effects of presession access on the break points of edibles and tokens using PR schedules. This manipulation will serve as a preliminary analysis of the extent to which tokens might function as generalized conditioned reinforcers for the current participants.

CHAPTER 2

PHASE 1: METHOD

Participants and Setting

Three children (2 boys and 1 girl) who attended a school for children with developmental disabilities served as participants for this study. Participants were identified by the school's director and were recruited through a flyer sent home to their parents. In order for a participant to be considered eligible for this study he or she had to (a) have a guardian who gave permission to participate in the study and (b) have experience with token economies. Generally speaking, token training for these participants consisted of creating a themed token board and tokens specialized to the child's interests and then having the child earn one token which was immediately exchanged for any item the child selected from the classroom. The number of tokens to be earned was then gradually increased based on performance.

Carmendy was an 8 year old girl who was diagnosed with developmental delays and functioned on a kindergarten grade level, according to the Woodcock-Johnson III Test of Achievement (Woodcock, McGrew, & Mather, 2001). Prior to the study, Carmendy had experience using tokens for about fourteen months. On average, she earned about ten tokens before accessing reinforcement in the classroom. With difficult tasks, only one response was required per token, but with mastered tasks, two to three responses were required before earning a token. According to teacher report, Carmendy typically chose to exchange her tokens for small toys (i.e., Polly Pockets, Legos, etc.), board games, and candy.

Zane was a 7 year old boy who had a diagnosis of attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder. According to the Woodcock-Johnson III Test, he

functioned on a first grade level. Prior to the study, Zane had experience using tokens for about two years. The number of responses required to earn a single token varied from one to five, depending on task difficulty. According to teacher report, Zane typically earned about ten tokens before accessing reinforcement and he chose to exchange his tokens for Legos, Wii, iPad, and jelly beans.

Damien was a 7 year old boy and was diagnosed with developmental delays and autism spectrum disorder. According to the Woodcock-Johnson III Test, he functioned on an end of first grade/beginning of second grade level. Prior to the study, Damien had experience using tokens for about eighteen months. On average, Damien completed two to three responses before earning each token and he earned about ten tokens before accessing reinforcement. According to teacher report, Damien typically chose to exchange his tokens for video games, board games, and occasionally more lessons.

All conditions and sessions were conducted in a room (approximately 1.83 m x 3.05 m) containing a bookshelf, two desks, four chairs, a tripod with video camera, and a three tier plastic storage bin. For all participants, one session was conducted per day, 3 to 5 days per week, depending on participant availability.

Tasks and Materials

During sessions, the participants worked on previously mastered academic skills. The experimenter interviewed the teachers to identify skills that had previously been mastered and then selected tasks on the basis of each participant's individual level of development and academic ability, as reported by their teachers and reflected in standardized test results. Math facts (addition) were chosen for all participants because (a) all of them had mastered some

number of addition math facts and (b) math facts involve discrete responses that do not require continuous instructions or prompts from the therapist. Figure 1 is an illustration of one of the math facts used in this study. Each math fact was written on a 107.95 mm x 139.7 mm piece of laminated paper (215.9 mm x 279.4 mm computer paper, cut in half). During all conditions, Carmendy worked on addition math facts zero through one up to plus eleven and Zane and Damien worked on addition math facts zero through five up to plus eleven. There was an aluminum pan (228.6 mm x 279.4 mm) with a sign attached that said "I'm Done" located on the far left hand corner of the desk. The math fact sheets were assembled in one large stack located at the far right hand corner of the desk. After taking a math fact from the stack and writing the correct answer on the sheet under the equal sign, the participant placed the completed math fact in the pan. The participant was given a dry-erase marker to write the answers and a facial tissue to erase mistakes. The token board consisted of one side of a green file folder that was laminated with the words "Token Board" written at the top and had three long strips of Velcro centered beneath the words. The tokens were brown paper circles, about the size of a quarter, that were laminated with the word "Token" printed across the center and Velcro affixed to the back. In the token condition, the therapist placed the token board in between the pan and the sheets at the far center of the desk. In all sessions and conditions, the participants were seated at a metal desk (approximately 609.6 mm x 457.2 mm) and the therapist sat adjacent to the participants' right side.

Measurement and Interobserver Agreement

The experimenter and trained observers collected data on correct and incorrect responses, reinforcer access, total session time, off-task behavior, and the progressive ratio (PR) break

point. Observers scored a response each time the child placed the completed task (math problem) in the finished pan. When the child wrote the correct answer to the math problem on the sheet, observers scored that response as correct. Observers counted a response as incorrect when either (a) the answer was not written clearly enough for the therapist to read or (b) when the child wrote the wrong answer on the sheet. Reinforcer access began when the therapist placed her hand on the math sheets or when the child took the item from the therapist and it ended when the child started working again or when the therapist returned the work materials to the desk. Session time began after delivering the initial task instructions and ended when the child asked to be finished or when the break point criterion was reached. Off-task behavior was scored 5 s after the child began engaging in an alternative activity, such as playing with the marker, and ended when the child began working on the task. Observers circled the PR break point when either the child was off-task for 2 min, no responses occurred for 2 min, or when the child said "I'm done."

All sessions were videotaped and observers scored responses using a data collection app on an iPod and using pencil and paper for interobserver agreement**.** Interobserver agreement was assessed for at least 30% of sessions by having a second observer independently watch the videos and collect data on all the above described behaviors. Interobserver agreement for incorrect and correct responses and PR break point was calculated by using point by point agreements versus disagreements. For example, the therapist and the IOA observer would score whether the participant answered math fact 1 incorrectly or correctly; an agreement was scored on math fact 1 if both observers scored the same response (both marked as incorrect or both marked as correct). This process continued until each math fact the child completed during the session was scored. IOA was then calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. Interobserver agreement

for reinforcer access, total session time, and off-task behavior was calculated by dividing the smaller number of seconds by the larger number of seconds and multiplying by 100. Average interobserver agreement (IOA) for Carmendy was 100% for incorrect and correct responses, off-task behavior, and PR break point. Average IOA for total session duration and reinforcer access was 97% (range, 92% to 99%), and 94% respectively (range, 80% to 100%). Average interobserver agreement for Zane was 100% for incorrect responses, off-task behavior, and PR break point. Average IOA for correct responses and total session duration was 99% (range, 98% to 100%), and average IOA for reinforcer access was 91% (range, 85% to 97%). Average interobserver agreement for Damien was 100% for incorrect and correct responses, off task behavior, and PR break point. Average IOA for total session duration and reinforcer access was 99% (range, 95% to 100%), and 96% respectively (range, 93% to 99%).

Procedure

Preference assessment. After receiving consent, the experimenter interviewed the participants' teachers to identify preferred edible and leisure items. The experimenter also asked each participant to name his or her favorite toys and snacks. After identifying stimuli, two paired-choice preference assessments, following the procedures described by Fisher et al. (1992), were conducted with each child; one assessment evaluated preference among edible items and a second assessment evaluated preference among leisure items. During this assessment, five to eight items were paired once with every other item in a counterbalanced order. The experimenter placed two stimuli approximately 127 mm apart, on the table in front of the participant. The participant had access to the food/toy he or she selected for one minute. The stimulus that was not selected was removed. The experimenter blocked all attempts to reach for both stimuli and

instructed the child to “pick just one.” The top three edible and the top three leisure items, identified with the above assessments, were used as reinforcers in this study.

PR conditions. There were three conditions in this study; edible, leisure, and token. The top three edibles from the preference assessment were used in the edible condition, the top three leisure items were used in the leisure condition and the top three edibles and the top three leisure items (six items total) were used in the token condition. Prior to each session, the experimenter placed the items on the table and the child selected what he or she wanted to earn during that session. The experimenter delivered the edible item that the child selected prior to the session as the reinforcer in the edible condition, the leisure item the child selected as the reinforcer in the leisure condition, and a token as the reinforcer in the token condition. Immediately after the child selected an item to earn, the experimenter removed all other items. Session materials were then placed on the child’s desk (marker, tissue for erasing, math facts, and “finished” pan) and instructions were provided to the child, i.e., “Today you will be working on math facts. You will receive _____ for completing math problems. Just grab a sheet from the stack, answer the problem and then put it in the finished bin. If at any point you want to be all done working, you can tell me I’m done” (experimenter points to all done sign). “What do you say when you want to be all done?” The session began after the experimenter delivered the instructions and the child answered the question by saying “I’m done.”

During sessions, the children worked on the previously described math fact tasks by writing the answer to each problem in the designated area on the laminated sheet. The participant then placed the completed sheet in the “finished” pan, took the next sheet from the stack, and continued. When the participant completed the PR schedule requirement, the therapist placed a hand on the math sheets and delivered the reinforcer. The PR requirements increased

arithmetically (step size of 1); beginning with one response to access reinforcement, then PR 2, PR 3, etc. In the edible and leisure conditions, the participant was given one minute access to consume the edible or play with the leisure item after completing each schedule requirement. If the participant made an error, the therapist said "That's not right" and allowed the child a few seconds to correct the error on his or her own. If the participant did not know the answer, the experimenter vocally prompted the correct answer. Any incorrect or prompted answers were counted toward the PR schedule requirement after the correct answer was written on the sheet and placed in the pan. If the participant tried to engage in conversation during the task or reinforcer time, the experimenter responded by saying "We can talk about that later if you want" and ignored all additional initiations within 30 s of the verbal redirection.

The PR break point was achieved and the session was terminated when the child asked to be finished or when no responding or off-task behavior occurred for two minutes. In the edible and leisure conditions, the experimenter immediately walked the child back to his or her classroom. In the token condition, the experimenter and child counted the number of tokens and the child was given access to the item he or she selected prior to the session. Each token was worth one edible or one minute access to the leisure item. For example, if the break point was PR 10, the child had been awarded ten tokens that were exchanged for ten minutes of play time with the leisure item or ten pieces of the edible.

Experimental Design

The first condition was randomly assigned to the participants, such that each participant started with a different condition. Subsequent conditions were counterbalanced across participants. The experimental conditions described above were arranged in an A-B-C design.

CHAPTER 3

PHASE 1: RESULTS AND DISCUSSION

Figure 2 depicts the results for the preference assessments for Carmendy. The preference assessments identified Jelly Bean, Oreo, and Kit Kat as the top three edibles and iPad, bowling, and an Angry Birds Game as the top three leisure items. As shown in Figure 3, the token condition had the highest break points (average of last 3 data points was PR 6), followed by the edible condition (average of last 3 data points was PR 4), and leisure condition (average of last 3 data points was PR 3). The open (exchange for leisure items) and closed (exchange for edible items) circles in Figures 3-5 depict the items for which Carmendy exchanged her tokens during the token condition. For 4 of 13 sessions, Carmendy exchanged her tokens for leisure items and for 9 sessions she exchanged her tokens for edibles. Session times ranged from 15 s to 15 min, with an average of 4.1 min. Figure 4 displays the rate of responding and the number of responses completed per session. Rate of responding was highest in the edible condition averaging 5.7 responses per minute (rpm), followed by 5.3 rpm in the leisure condition, and 4.3 rpm in the token condition. It should be noted that rate per minute for all participants, sessions, and conditions was calculated by taking the total number of responses for that session and dividing it by the total session duration minus reinforcer access time. The average number of responses completed during the edible, leisure, and token conditions was 12, 11, and 18 responses per session, respectively. Figure 5 shows the percentage of correct responses and the percentage of off-task behavior per session for Carmendy. The average percentage of correct responses was 92% for the edible condition, 97% for the leisure condition, and 94% for the token condition. Carmendy was on task for all sessions except for a few seconds during two sessions in the token

condition (11 s and 5 s, specifically) and all sessions were terminated when she stated “I’m done.”

Figure 6 depicts the results for the preference assessments for Zane. The preference assessments identified Doritos, Oreo, and Kit Kat as the top three edibles and iPad, bowling, and Legos as the top three leisure items. As shown in Figure 7, the token condition had the highest break points (average of last 3 data points was PR 14), followed by the edible condition (average of last 3 data points was PR 11), and leisure condition (average of last 3 data points was PR 6). The open (leisure items) and closed (edible items) circles in Figures 7-9 depict the items for which Zane exchanged his tokens during the token condition. For 8 of 13 sessions, Zane exchanged his tokens for leisure items and for 5 sessions he exchanged his tokens for edibles. Session times ranged from 5 to 35 min, with an average of 11.9 min. Figure 8 displays the rate of responding and the number of responses completed per session. Rate of responding was highest in the edible condition, averaging 11 rpm, followed by 10 rpm in the token condition, and 6 rpm in the leisure condition. The average number of responses completed during the leisure, token, and edible conditions was 36, 122, and 91 responses per session, respectively. Figure 9 shows the percentage of correct responses and percentage of off-task behavior per session for Zane. The average percentage of correct responses was 95% for the leisure condition, and 96% for the token and edible conditions. Zane was on task during all sessions and conditions and all sessions were terminated when he stated “I’m done.”

Figure 10 depicts the results for the preference assessments for Damien. The preference assessments identified Three Musketeers Bar, M&M, and chocolate chip cookie as the top three edibles and iPad, Legos, and an Angry Birds Game as the top three leisure items. As shown in Figure 11, the edible condition had the highest PR break points (average of last 3 data points was

PR 11), followed by the leisure condition (average of last 3 data points was PR 8), and token condition (average of last 3 data points was PR 7). The open (leisure items) and closed (edible items) circles in Figures 11-13 depict the items for which Damien exchanged his tokens during the token condition. For all 8 sessions, Damien exchanged his tokens for edibles. Session times ranged from 3 to 33 min, with an average of 12.2 min. Figure 12 shows the rate of responding and the number of responses completed per session. Rate of responding was highest in the edible and leisure conditions, averaging 6.3 and 6.2 rpm respectively, followed by 5.3 rpm in the token condition. The average number of responses completed during the token, edible, and leisure conditions was 34, 62, and 60 responses per session. Figure 13 displays the percentage of correct responses and the percentage of off-task behavior per session for Damien. The average percentage of correct responses was 96% for the token condition and 99% for the leisure and edible conditions. Damien was on task during all sessions and conditions and all sessions were terminated when he stated "I'm done." Figures 14 and 15 display the PR break points and the cumulative number of responses across sessions for all three conditions and participants, respectively.

The results of Phase 1 indicate that tokens may function as generalized conditioned reinforcers because the token condition had the highest PR break points for both Carmendy and Zane. It should be noted that Damien chose to exchange his tokens only for edibles in Phase 1, which may account for the token PR break points being lower. All three participants made very few errors (percent correct ranged from 92%-99%) and all participants were consistently on task, suggesting that the items used in the conditions were reinforcers. The edible condition had the highest rate of responding for all three participants, which is consistent with the benefits of using edibles as reinforcers. These results suggest that edibles may be the best stimuli to use if the goal

is to achieve a high rate of responding. Given that tokens were found to be effective reinforcers, Phase 2 was conducted to further evaluate to what extent tokens functioned as generalized conditioned reinforcers. In this phase, presession access was manipulated to examine the PR break points of edible items and tokens and to test the reinforcing nature of tokens.

CHAPTER 4

PHASE 2: METHOD

Participant and Setting

Damien from Phase 1 served as the participant for this study. The setting and materials were the same as in Phase 1, except for the addition of a food scale which was used to weigh consumption amount.

Data collection and Interobserver Agreement

Observers collected data on the same behaviors and in the same manner as in Phase 1. The experimenter also weighed the edibles prior to and after each presession access period to determine the total amount of grams consumed. Interobserver agreement was assessed during 33% of sessions and was calculated using the same method as Phase 1. Average interobserver agreement (IOA) was 100% for incorrect responses and PR break point. Average IOA for total session duration and correct responses was 99% (range, 86% to 100%). Average IOA for off-task behavior was 97% (range, 82% to 100%) and 95% for reinforcer access (range, 81% to 100%).

Procedure

Baseline. There were two conditions in this phase; edible and token. Prior to conducting the presession manipulations, a multielement baseline (Baseline 1) was conducted in which the edible and token conditions were alternated semi-randomly. Each condition could not occur more than two times consecutively. Baseline 1 sessions were conducted using similar procedures to Phase 1. The only difference in procedures was in the token condition, in which the participant

selected the item to earn after the session was terminated instead of prior to the session. This procedural change was implemented for two reasons (a) to allow for a better assessment of the effects of satiation in presession access comparisons and (b) so the participant was not limited to a particular stimulus he selected at the beginning of the session. Additionally, after presession manipulations were completed, a reversal to baseline (Baseline 2) was conducted. In Baseline 2, sessions were identical to Baseline 1 with the exception that Damien was allowed to exchange his tokens for broader array of leisure items. This change was made to evaluate whether having a larger selection of leisure items from which to choose, as was the case in presession comparison 2, would drastically alter baseline rates.

Presession access comparison. After Baseline 1, the participant entered the presession access comparison (PAC). The procedures were identical to the multielement Baseline 1 condition, except that prior to each session, the child was seated at the desk and given access to the top three edibles from the preference assessment conducted in Phase 1. In PAC 1, the participant had five minutes to consume as much of the three edibles as he chose. The instructions given to Damien were as follows: "You can eat as much of the snacks as you want. If at any point you are full, you can just tell me 'I'm full'." The experimenter then started a five minute timer. When the timer beeped or when the participant stated that he was full or wanted to stop eating, the session for that day was initiated.

In PAC 2, the presession consumption time increased to ten minutes. Additionally, the token condition was altered to include a larger variety of leisure items (i.e., the child could select whatever toy, game, activity that was available in the classroom). However, the edible items in the token condition were still limited to the top three edibles from the preference assessment. These manipulations were made for two reasons: (a) the experimenters discovered during PAC 1

that five minutes was not enough time for Damien to become satiated on the edibles and (b) during the token condition, in PAC 1, Damien repeatedly refused to exchange the tokens for any of the available items (the top three edible and top three leisure items from the preference assessments conducted in Phase 1).

Experimental Design

For the baseline conditions and presession access comparisons, a multielement design was used in which the edible and token conditions were alternated semi-randomly. Additionally, each condition could not occur more than two times consecutively.

CHAPTER 5

PHASE 2: RESULTS AND DISCUSSION

As shown in Figure 11, during Baseline 1 the token condition had the highest break points (average of last 3 data points was PR 11) followed by the edible condition (average of last 3 data points was PR 10). The open (leisure items) and closed (edible items) circles in Figures 11-13 depict the items for which Damien exchanged his tokens during token sessions in Baseline 1. For all sessions in Baseline 1, Damien exchanged his tokens for edibles. Session times ranged from 10 to 50 min, with an average of 21.8 min. Figure 12 displays the rate of responding and total number of responses per session during Baseline 1. Rate of responding was highest in the token condition, averaging 7.1 rpm followed by 6.3 rpm in the edible condition. The average number of responses completed during the edible and token conditions in Baseline 1 was 124 and 164 responses per session, respectively. Figure 13 shows the percentage of correct responses and percentage of off-task behavior per session in Baseline 1. The average percentage of correct responses was 99% for the edible condition and 100% for the token condition. Damien was on task for most sessions in Baseline 1; the only times he was off-task was when he was sick and he stopped to blow his nose. All sessions were terminated when he stated “I’m done.”

As shown in Figure 11, during presession access comparison (PAC) 1, the token and edible conditions had similar break points (average of last 3 data points was PR 9). The open (leisure items), closed (edible items) and grey (no selection) circles in Figures 11-13 depict the items for which Damien exchanged his tokens during token sessions in PAC 1. For all token condition sessions during PAC 1, Damien refused to exchange his tokens for the any of the edible or leisure items. Session times ranged from 8 to 12 min, with an average of 9 min. Figure 12 depicts the rate of responding and total number of responses per session during PAC 1. Rate

of responding was similar in both the token and edible conditions with 6.5 rpm in the token condition and 6.4 rpm in the edible condition. The average number of responses completed during the edible and token conditions in PAC 1 was 54 and 48 responses per session, respectively. Figure 13 shows the percentage of correct responses and percentage of off-task behavior per session in PAC 1. The average percentage of correct responses for the edible and token conditions was 100%. Damien was on task for most sessions during PAC 1, except for when he was sick, and all sessions were terminated when he stated "I'm done."

As shown in Figure 11, during PAC 2 the token condition had higher PR break points (average of last 3 data points was PR 11) than the edible condition (average of last 3 data points was PR 0). The open (leisure items) and closed (edible items) circles in Figures 11-13 depict the items for which Damien exchanged his tokens during token sessions in PAC 2. For all sessions in PAC 2, Damien exchanged his tokens for leisure items. Session times ranged from 0 to 16 min, with an average of 7.5 min. Given that Damien did not want to work for edibles in the edible condition, it was not possible to collect data on the percentage of correct responses or rate per minute during PAC 2. Figure 12 displays the rate of responding and total number of responses per session during PAC 2. The average rate of responding was 6.6 rpm in the token condition. The average number of responses completed during the edible and token conditions in PAC 2 was 0 and 56 responses per session, respectively. Figure 13 shows the percentage of correct responses and percentage of off-task behavior per session during PAC 2. The average percentage of correct responses was 100% in the token condition. Damien was on task for most sessions in PAC 2 and all sessions were terminated when he stated "I'm done."

As shown in Figure 11, during Baseline 2 the token and edible conditions had similar break points (average of last 3 data points was PR 7). The open (leisure items) and closed (edible

items) circles in Figures 11-13 depict the items for which Damien exchanged his tokens during token sessions in Baseline 2. For all sessions in Baseline 2, Damien exchanged his tokens for leisure items. Session times ranged from 0 to 9 min, with an average of 4.6 min. Figure 12 displays the rate of responding and total number of responses per session during Baseline 2. Rate of responding was highest in the token condition, averaging 7.8 rpm followed by 6 rpm in the edible condition. The average number of responses completed during the edible and token conditions in Baseline 2 was 25 and 31 responses per session, respectively. Figure 13 depicts the percentage of correct responses and percentage of off-task behavior per session during Baseline 2. The average percentage of correct responses was 99% for the token condition and 98% for the edible condition. Damien was on task for all sessions during Baseline 2, except for two when he decided to make-up a song and proceeded to sing it; this lasted approximately 6 s and 10 s in each case. All sessions were terminated when he stated "I'm done."

Interestingly, during the multielement Baseline 1 the PR break points for the edible and token condition reversed; tokens had both higher break points and response rates. It should also be noted that during Phase 1 and Baseline 1, Damien only exchanged his tokens for edibles. During presession access comparison (PAC) 1, PR break points and responding remained at baseline levels for both edibles and tokens. Additionally, Damien refused to exchange his tokens for any item. However, after increasing the consumption time and adding in an exchange for a broader array of leisure items during PAC 2, PR break points for edibles dropped to zero while break points for tokens remained unchanged. Also during PAC 2, Damien only exchanged his tokens for leisure items. These results suggest that tokens may be generalized conditioned reinforcers, because they were not affected by satiation. Lastly, during Baseline 2 the PR break

points for edibles were recovered, while break points for tokens remained around Baseline 1 levels.

CHAPTER 6

GENERAL DISCUSSION

During Phase 1, the PR break points for tokens, edibles, and leisure items were evaluated and compared with two boys and one girl to determine the reinforcing potency of these stimuli. Similar trends were found for two out of three participants; tokens had the highest break points followed by edibles and leisure items. For Damien, edibles had the highest break points followed by leisure items and tokens. However, it should be noted that Damien only exchanged his tokens for edibles in Phase 1, which may account for the lower break points in the token condition. For all sessions and conditions during Phase 1, average percentage of correct responses was above 90% and average off-task behavior occurred for less than 1% of any given session for all participants. Additionally, the number of responses per minute was highest in the edible condition for all three participants.

During Phase 2, the effects of presession access on edibles and tokens were examined with one participant (Damien) from Phase 1 as a test to see if tokens functioned as generalized conditioned reinforcers. In Phase 2, presession access altered the reinforcing potency of edibles, but not tokens. Damien also switched from exchanging his tokens exclusively for 3 Musketeers Bars in Phase 1 to exchanging the tokens for a variety of leisure items during PAC 2 and Baseline 2. For all sessions and conditions in Phase 2, average percentage of correct responses was 100% and average off-task behavior occurred for less than 2% of any given session. Additionally, the number of responses per minute was highest in the token condition.

Because there have been few, if any, studies comparing and evaluating the reinforcing potency of categories of reinforcers, the results of Phase 1 extend the previous research by examining break points with edibles, leisure items, and tokens. The current findings have

implications for the evaluation of different classes of reinforcers, which may be important for children that have a limited number of items and activities that function as reinforcers. For example, PR schedules may be useful in determining whether newly conditioned stimuli or novel items actually function as reinforcers and how hard the learners will work to earn these items. Furthermore, the results of Phase 1 corroborate previous findings (Jerome & Sturmey, 2008; Penrod, Wallace, & Dyer, 2008; Roane, Lerman, & Vorndran, 2001) by suggesting that progressive ratio schedules may be a promising technology for identifying potent reinforcers.

The results of this study are similar to those shown by Sy and Borrero (2009) with regards to the effects of presession access on responding. In the current study, we found that although five minutes of consumption time did not greatly affect the PR break points of edibles, it decreased rate of responding. Increasing consumption time to ten minutes eliminated responding for edibles completely. This study extends previous literature by incorporating an evaluation of the effects of presession access on tokens. Additionally, this study highlights the importance of considering motivating operations when programming reinforcement delivery. For example, it may be more effective to use tokens rather than edibles as reinforcement soon after meals. Sy and Borrero (2009) did not find that presession access reduced responding with nonedibles; therefore, studies comparing the effects of presession access to leisure items may be helpful to determine if similar findings with PR break points would be achieved.

As stated previously, stimuli are considered generalized conditioned reinforcers when they can effectively reinforce behavior across a wide range of motivating operation conditions (Cooper, Heron, & Heward, 2007). Two advantages of generalized conditioned reinforcers are that they maintain performance over extended periods of time and they are not as vulnerable to the effects of satiation (Kazdin, 1972). Results of Phase 1 indicate that tokens may function as

generalized conditioned reinforcers because they had the highest break points for two participants, and maintained a substantial rate of responding even for the participant for whom this was not the case. Additionally, during Phase 2 for Damien, presession access altered the reinforcing potency of edibles, but not tokens. Damien also switched from exchanging his tokens exclusively for one specific edible in Phase 1 to exchanging the tokens for a range of leisure items in Phase 2; further suggesting that tokens may have functioned as generalized conditioned reinforcers.

The current findings suggest that it may be useful for clinicians to use PR schedules as a way to verify token program mastery. Through the use of PR schedules, the clinician may discover that although a client may work to earn tokens, the PR break points may be low; suggesting that tokens are not functioning as intended and additional conditioning and token training for that client is needed.

One limitation of this study is that all participants had extensive and largely uncontrolled histories with tokens. It is possible that this history resulted in higher PR break points and the apparent generalized conditioned reinforcement of tokens. It is unclear which elements, if any, of the token training and history with tokens may have been necessary or sufficient to achieve the current results. Given this, one cannot be sure that similar results would be obtained if the participants had a different training history and/or a more limited experience with using tokens. Future research evaluating the effects of different histories of token training on PR breakpoints is warranted.

Vollmer and Hackenberg (2001) identified the need for a more fundamental understanding of reinforcement contingencies, especially those involving the class of generalized conditioned reinforcers. Using PR schedules as a means to examine these contingencies may

prove to be a fruitful avenue worthy of exploration. This line of research is important, especially from an applied perspective, as it is advantageous to not be constrained to having to deliver primary or tangible reinforcers directly all the time. As described by Ayllon and Azrin (1968a), there are countless benefits to using tokens as generalized conditioned reinforcers, such as portability and durability, ease of delivery, and the visual representation of improvement and progress. Evaluation of how factors such as training history, exposure, and number of back-up reinforcers affect PR break points may assist in development of models for teaching token systems and facilitating a generalized conditioned reinforcement function for tokens. Additionally, research could evaluate how participant characteristics may alter the extent to which tokens can be conditioned to function as generalized conditioned reinforcers. It is possible that a certain performance level is desirable or a specific number of back-up reinforcers is necessary to expedite acquisition.

One aspect of Damien's data that warrants discussion is the results of the multielement comparison (particularly the two baseline conditions) when compared with the results of Phase 1. The finding that break points reversed during Baseline 1 could be due to the differences in procedural arrangement; in Phase 1 the participants selected the stimulus to earn prior to the session, but in Phase 2 the participant selected the stimulus after session termination. It is possible that making the choice up front made the token solely represent that specific stimulus, thereby making the token less valuable then earning the stimulus directly. Conversely, the results may suggest that presenting the conditions sequentially may not be the best method for assessing reinforcer potency. Conducting conditions sequentially allows for a rigorous evaluation of the effects of one variable on responding at a time and eliminates potential confounding influence between conditions. This was the rationale for initially conducting the conditions sequentially.

The advantages of a multielement analysis are that it allows any confounds with history to impact both conditions equally, it incorporates multiple short reversals, and it allows one to superimpose a true reversal design, as was done in Phase 2 of the current study. Further examination is warranted to determine the best experimental design to evaluate reinforcer strength. Specifically, research might compare the current approach in conjunction with a concurrent-chains arrangement and/or a reversal design. Studies of this nature would allow researchers to see what the PR break points are for these stimuli with each arrangement in order to determine the most efficient and effective design.

Roane et al. (2001) found that stimuli with higher break points were more effective in decreasing problem behavior and increasing appropriate behavior. These findings, combined with the outcomes of this study, suggest that tokens may be the most effective stimuli to use when trying to increase appropriate behavior. Additionally, Penrod et al. (2008) found that higher preference stimuli resulted in higher break points. All the items used in this study were ranked as the top three preferred items according to the initial preference assessments. An interesting extension of the current study would be to replicate the procedures with the bottom three items from the preference assessments and compare the break points. In addition, it might be useful to evaluate the break points of tokens if they are only exchangeable for lesser preferred items. Studies of this nature would not only expand the findings of Penrod et al. (2008) but would also further highlight the importance of carefully considering which preference assessment to use when identifying preferred stimuli. For example, certain preference assessments (SS) are known to identify several highly preferred stimuli, which may be suitable for when schedule requirements are low (mastered or easy tasks). However, other types of

preference assessments (paired choice) provide a better differentiation and gradient of stimuli, which may be advantageous when teaching new skills or working on more difficult tasks.

It might also be useful to evaluate child preference for the reinforcement conditions as a measure of concurrent validity for the PR schedule comparison. This could be achieved by presenting the conditions in a concurrent chains arrangement in which the participant makes a response in order to select which condition he or she would like to experience (Sran & Borrero, 2010; Tiger, Hanley, & Hernandez, 2006). It may be the case that while tokens maintain responding longer and have higher break points; the child would actually prefer to earn edibles directly. In addition, it would be interesting to evaluate whether choice affects the PR break points in order to determine if higher break points would be obtained for the same conditions and stimuli if the child was able to select the condition for that session or task. Research evaluating choice suggests that performance on tasks is greater (more on task behavior and less inappropriate behavior) when the individuals were provided with self-selected rather than experimenter-selected tasks (Bambara, Ager, & Koger, 1994; Parsons, Reid, Reynolds, & Bumgardner, 1990). This line of research would have great implications for the applied field as it would allow clinicians to be more effective in programming skill acquisition.

It is clear from this study that PR schedules provide valuable information on the number of responses an individual will emit to obtain a particular reinforcer. This information is especially beneficial when clinicians are trying to thin the schedule of reinforcement for a particular response. To further expand upon these findings, studies should examine the same stimuli used in this study with different tasks to evaluate the extent to which task difficulty affects PR break points and reinforcer potency. Research has demonstrated that task difficulty affects not only the amount of inappropriate behavior emitted but also the efficacy of some

reinforcers (Lannie & Martens, 2004; Neef, Shade, & Miller, 1994; Weeks & Gaylord-Ross, 1981). Having information as to what extent task difficulty effects the potency of tokens, edibles, and leisure items as reinforcers will also allow clinicians to more effectively thin schedules without clients experiencing ratio strain.

It should be noted that this study is not without its limitations. First, because there were only three participants in Phase 1 and one participant in Phase 2, replications of this study are needed with additional participants to validate these findings. Second, the number of items available for the participant to choose from may have impacted the PR break points, especially in the token condition. When Damien was allowed to exchange his tokens for a larger selection of leisure items in Phase 2, the average PR break points were 11 as opposed to 7 in Phase 1. Additional research investigating the optimal number of items available to the client is warranted to determine how this could affect break points. Examining the effects of magnitude on PR break points with these stimuli might be another avenue for future studies. It could be the case that the limited amount of reinforcer access time (one minute) was responsible for leisure items having the lowest break points for two out of the three participants. Furthermore, it may be useful to evaluate praise and other social consequences in addition to the reinforcers included in this study. Notwithstanding, these findings suggest that PR schedules may be useful as a means to better assess certain dimensions of tasks and how they affect reinforcer effectiveness (e.g., amount of effort the client is willing to exert, the duration at which the client willing to work, how many responses the client will emit, etc.), and to evaluate to what extent tokens actually function as generalized conditioned reinforcers.

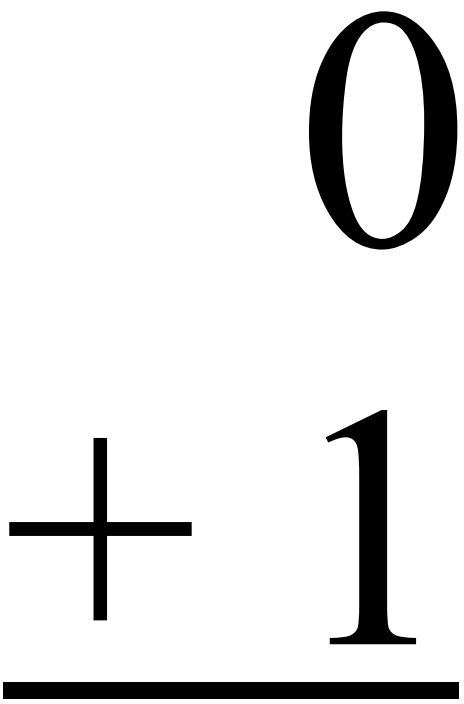

Figure 1. An illustration of the math tasks.

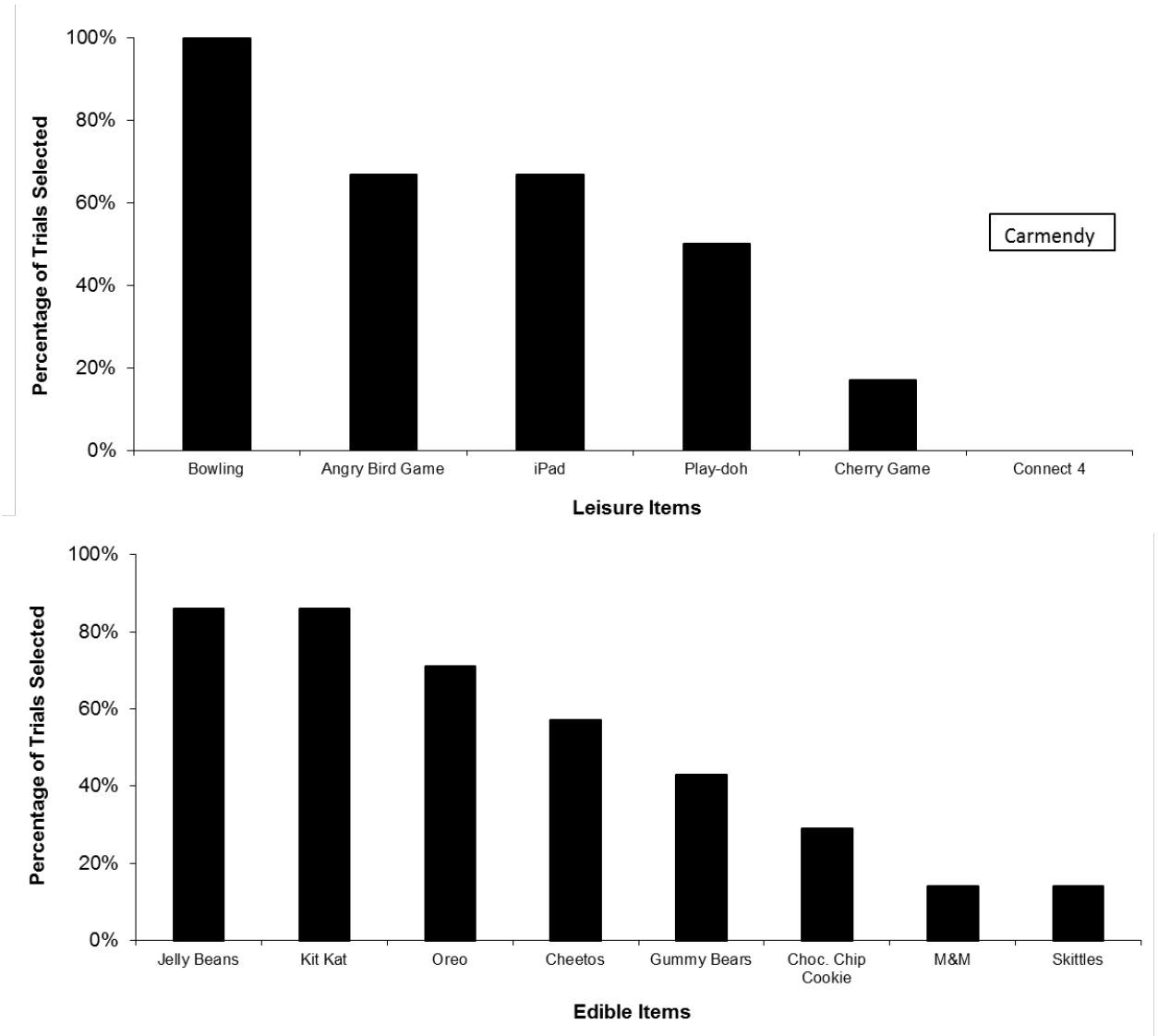

Figure 2. Percentage of trials each leisure item (top panel) and each edible item (bottom panel) was selected.

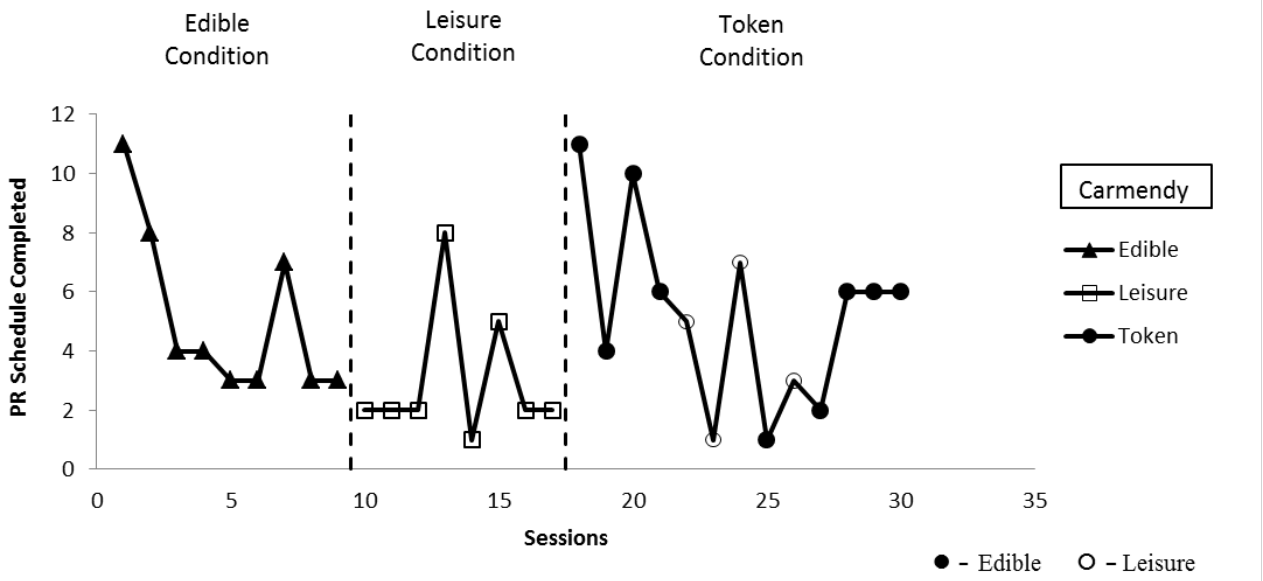

Figure 3. Progressive ratio break point completed for each session. Edible condition is first, followed by leisure and token. In the token condition, closed circles indicate sessions in which Carmendy exchanged tokens for edible items and open circles indicate selection of leisure items.

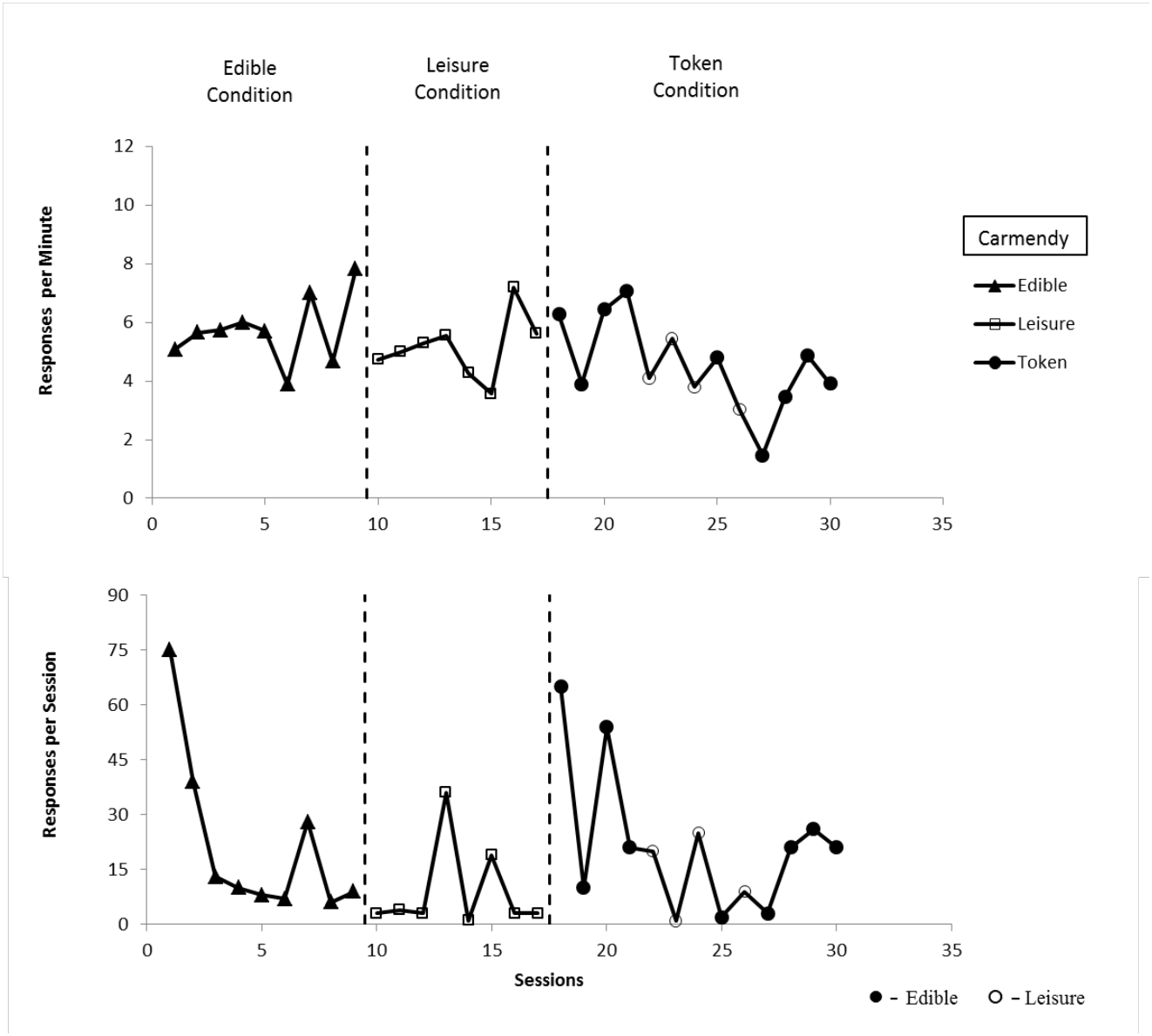

Figure 4. Responses per minute (top panel) and total number of responses per session (bottom panel). In the token condition, closed circles indicate sessions in which Carmendy exchanged tokens for edible items and open circles indicate selection of leisure items.

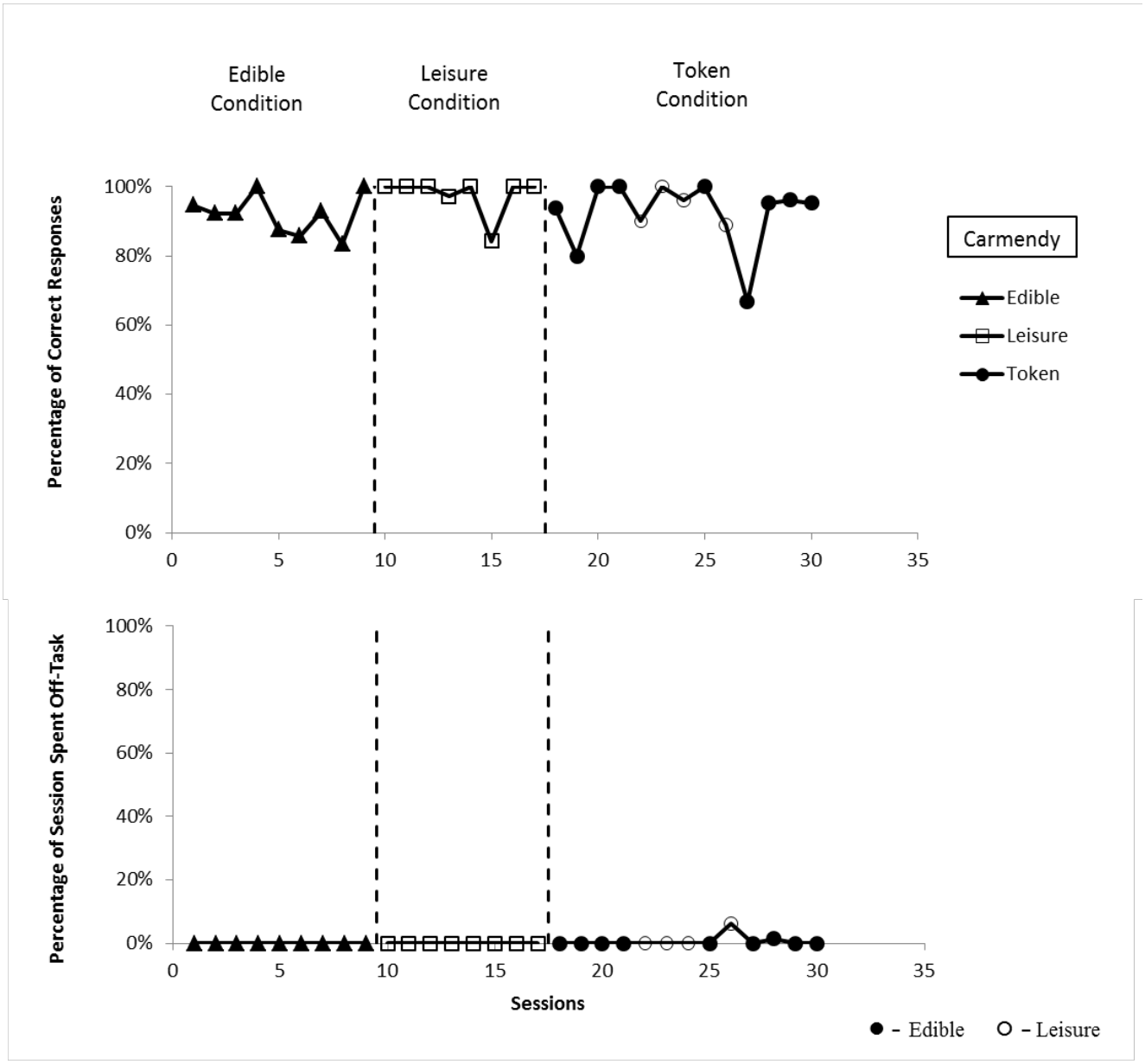

Figure 5. Percentage of correct responses per session (top panel) and percentage of session spent off-task (bottom panel). In the token condition, closed circles indicate sessions in which Carmendy exchanged tokens for edible items and open circles indicate selection of leisure items.

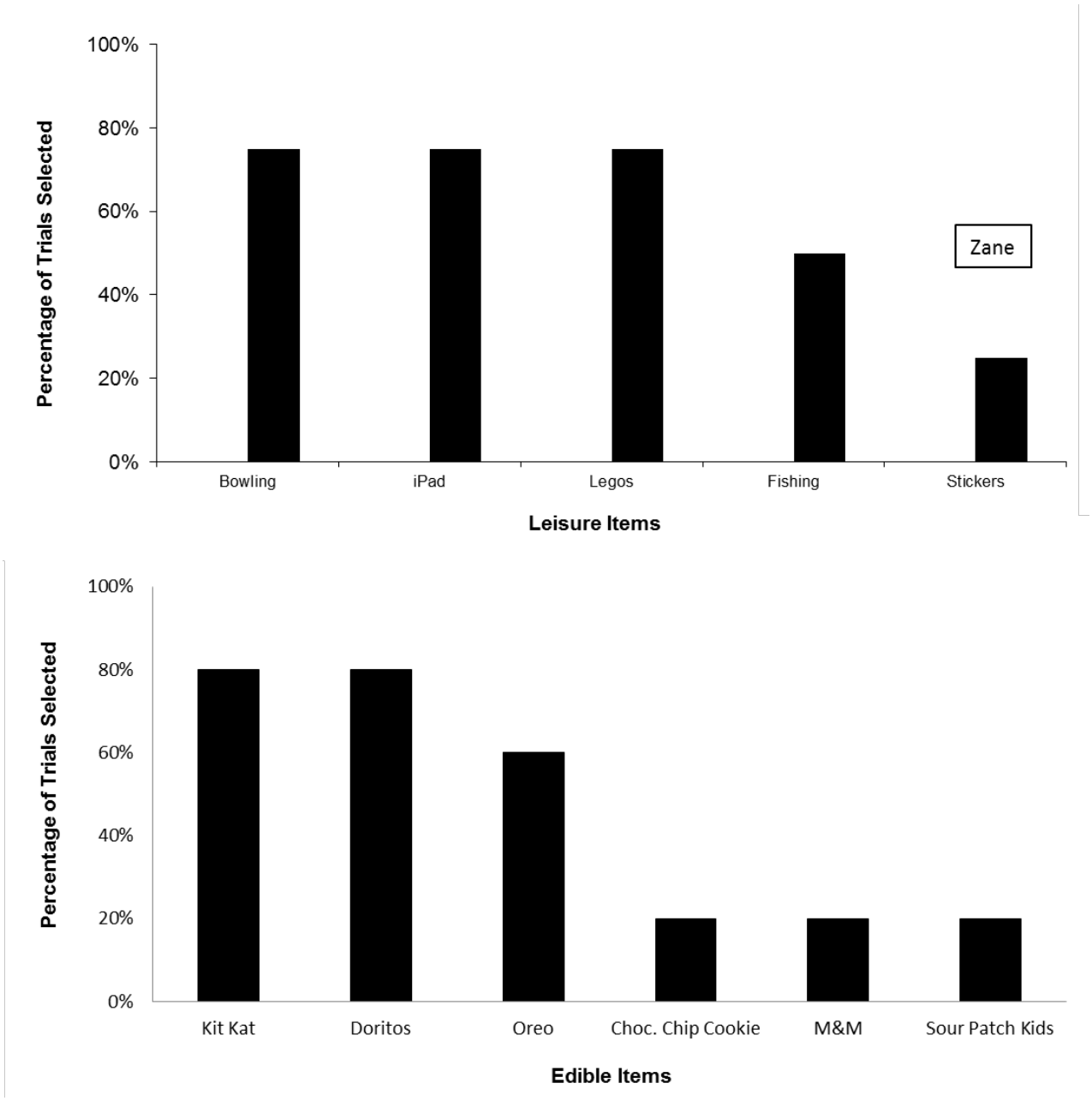

Figure 6. Percentage of trials each leisure item (top panel) and each edible item (bottom panel) was selected.

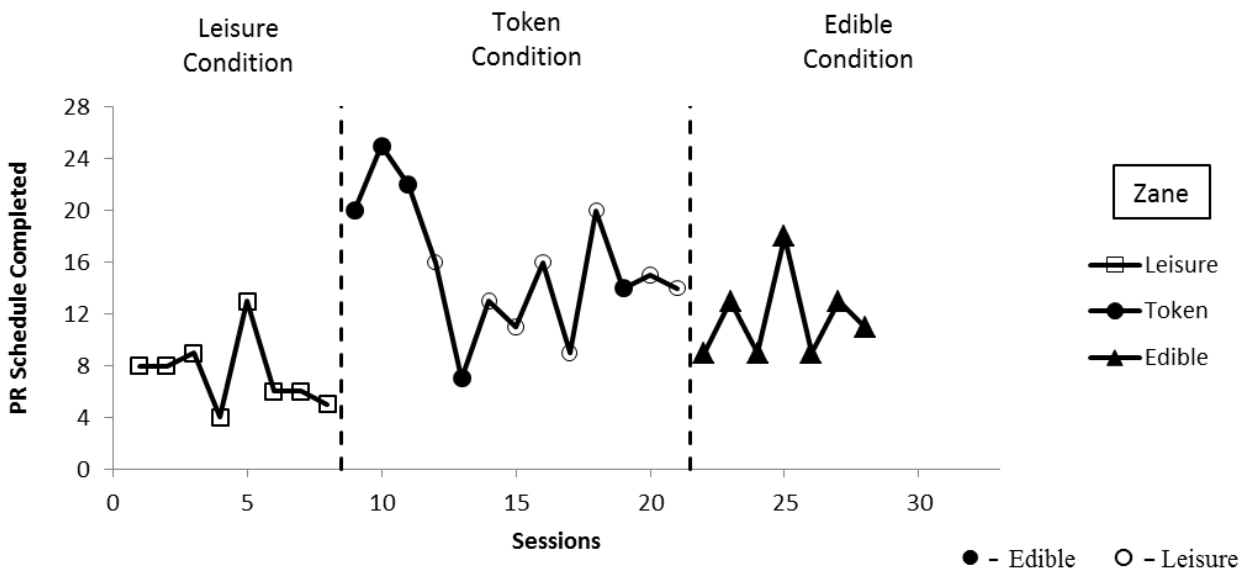

Figure 7. Progressive ratio break point completed for each session. Leisure condition is first, followed by token and edible. In the token condition, closed circles indicate sessions in which Zane exchanged tokens for edible items and open circles indicate selection of leisure items.

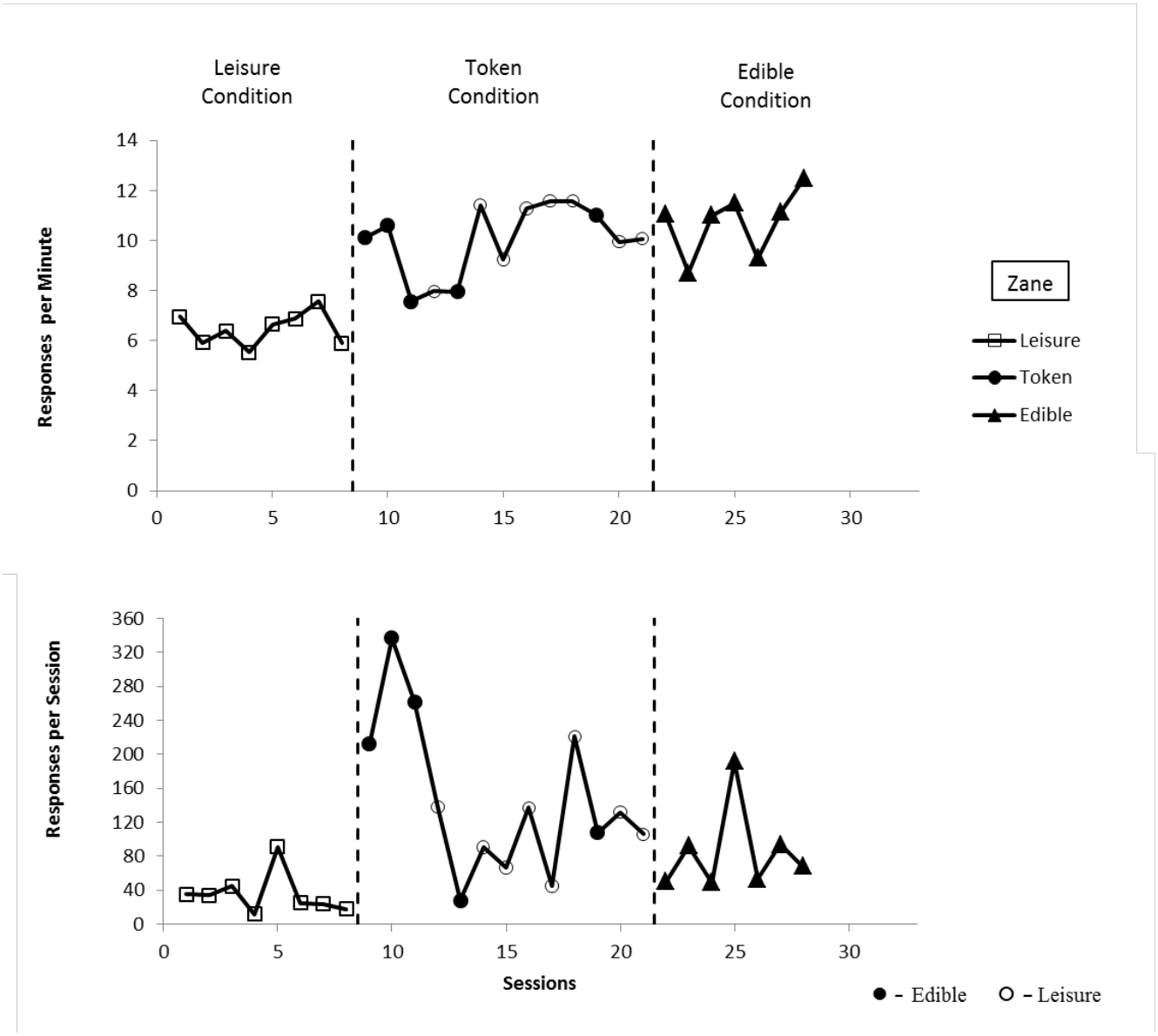

Figure 8. Responses per minute (top panel) and total number of responses per session (bottom panel). In the token condition, closed circles indicate sessions in which Zane exchanged tokens for edible items and open circles indicate selection of leisure items.

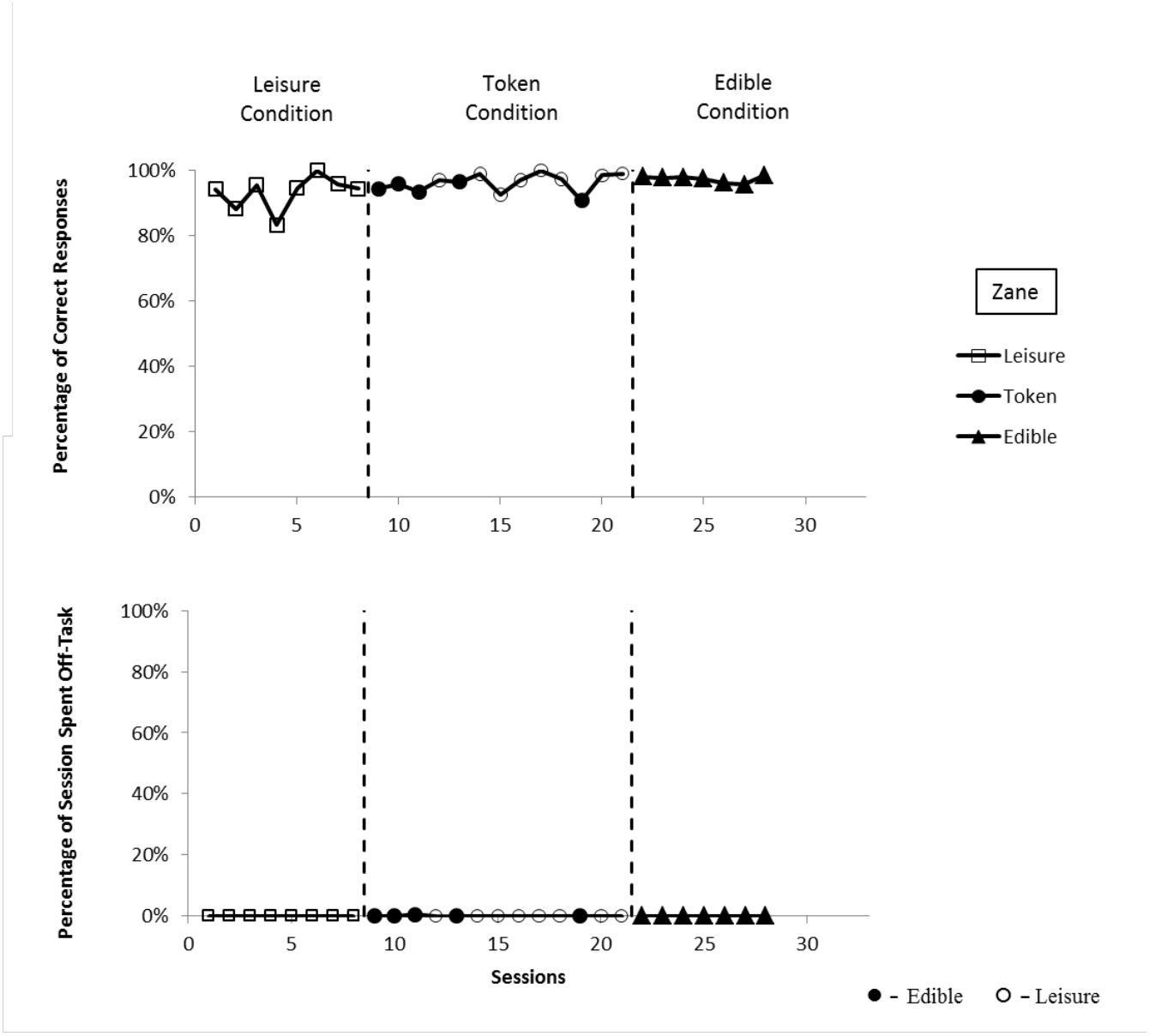

Figure 9. Percentage of correct responses per session (top panel) and percentage of session spent off-task (bottom panel). In the token condition, closed circles indicate sessions in which Zane exchanged tokens for edible items and open circles indicate selection of leisure items.

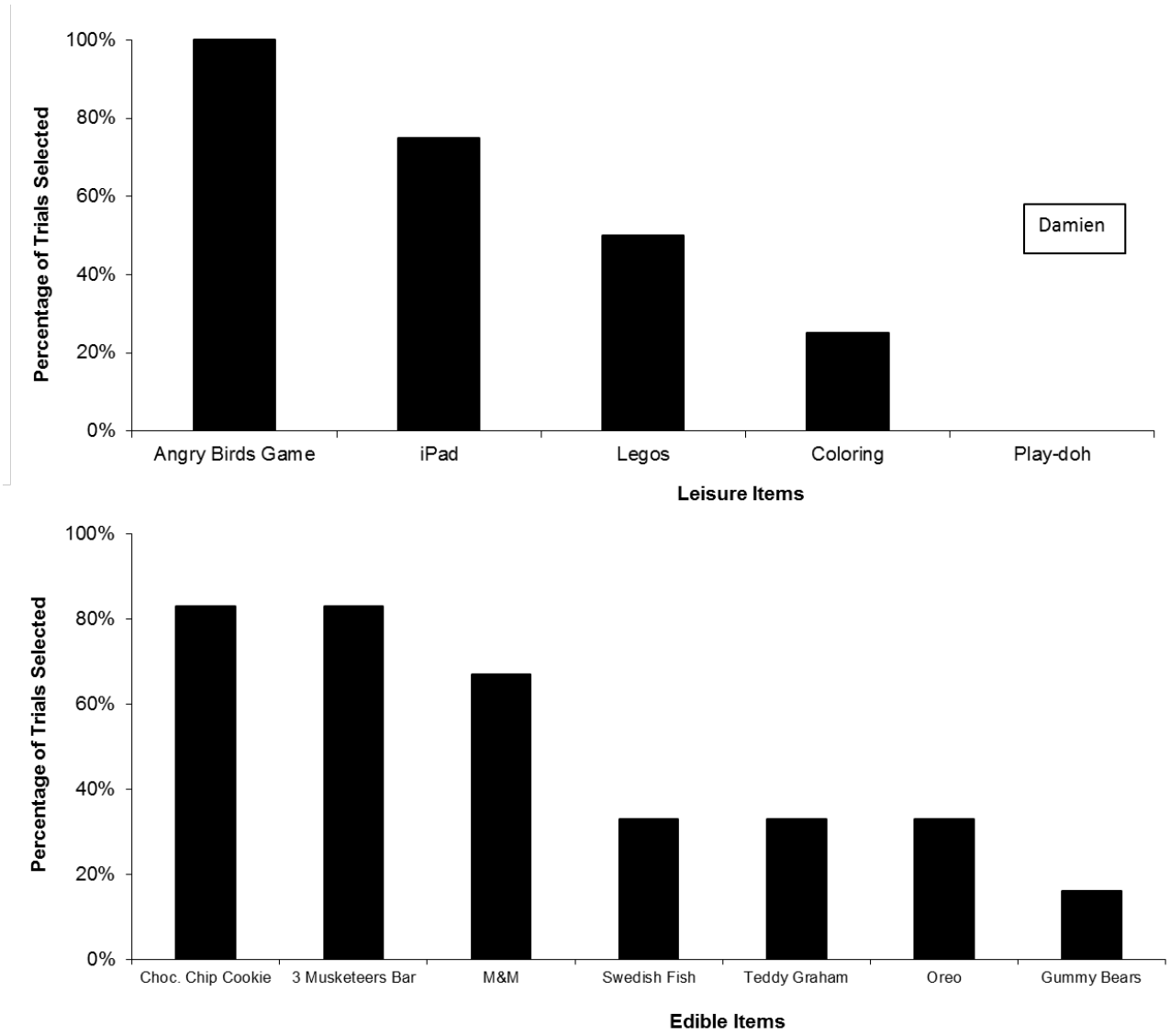

Figure 10. Percentage of trials each leisure item (top panel) and each edible item (bottom panel) was selected.

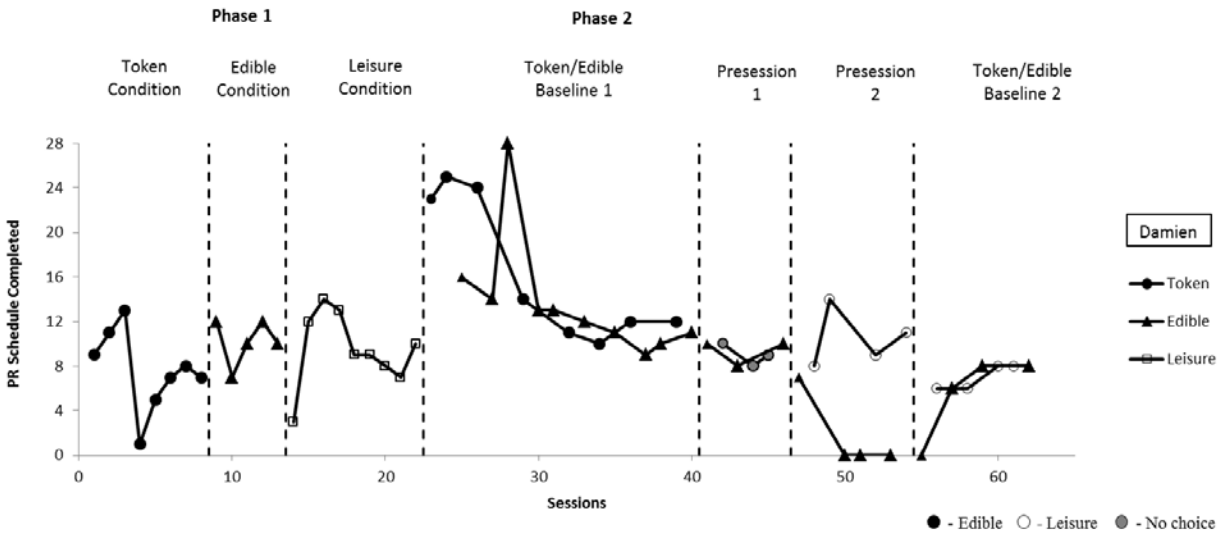

Figure 11. Progressive ratio break point completed for each session. Phase 1 consists of token condition first, followed by edibles and leisure items. Phase 2 consists of Baseline 1 first, followed by presession 1, presession 2, and Baseline 2. In the token condition, closed circles indicate sessions in which Damien exchanged tokens for edible items, open circles indicate leisure items, and gray circles indicate that he chose not to exchange his tokens for any item.

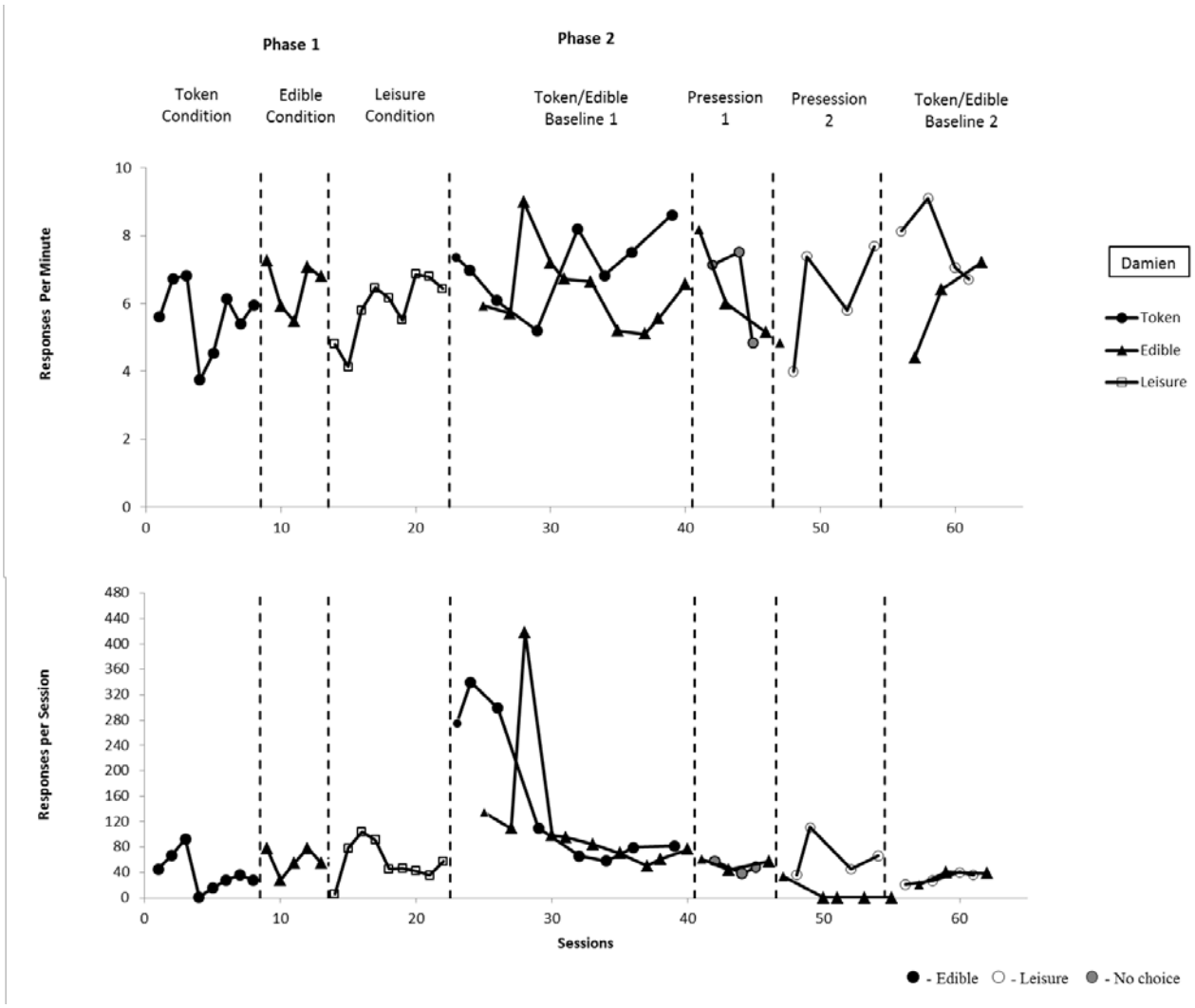

Figure 12. Responses per minute (top panel) and total number of responses per session (bottom panel). In the token condition, closed circles indicate sessions in which Damien exchanged tokens for edible items, open circles indicate leisure items, and gray circles indicate that he chose not to exchange his tokens for any item.

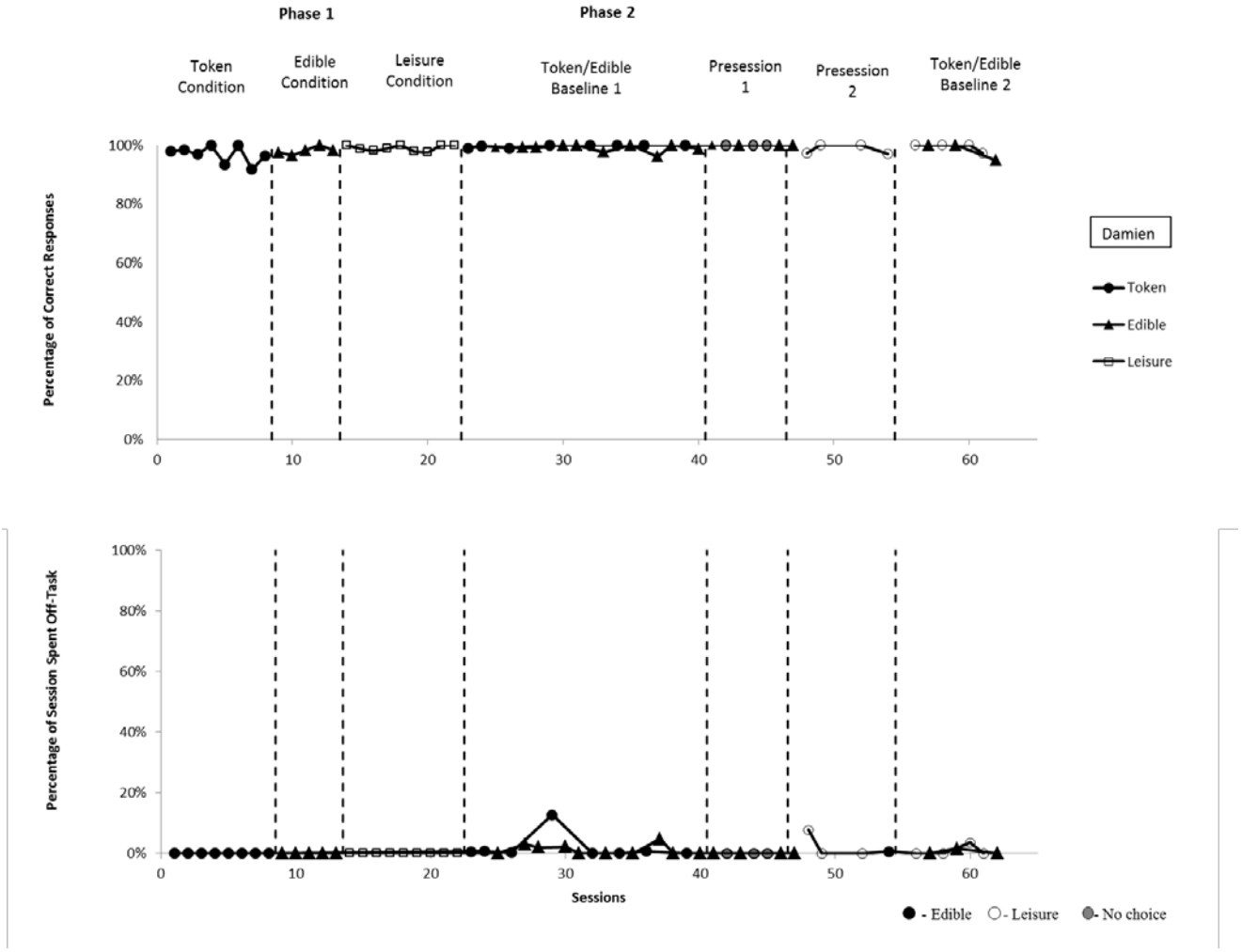

Figure 13. Percentage of correct responses per session (top panel) and percentage of session spent off-task (bottom panel). In the token condition, closed circles indicate sessions in which Damien exchanged tokens for edible items, open circles indicate leisure items, and gray circles indicate that he chose not to exchange his tokens for any item.

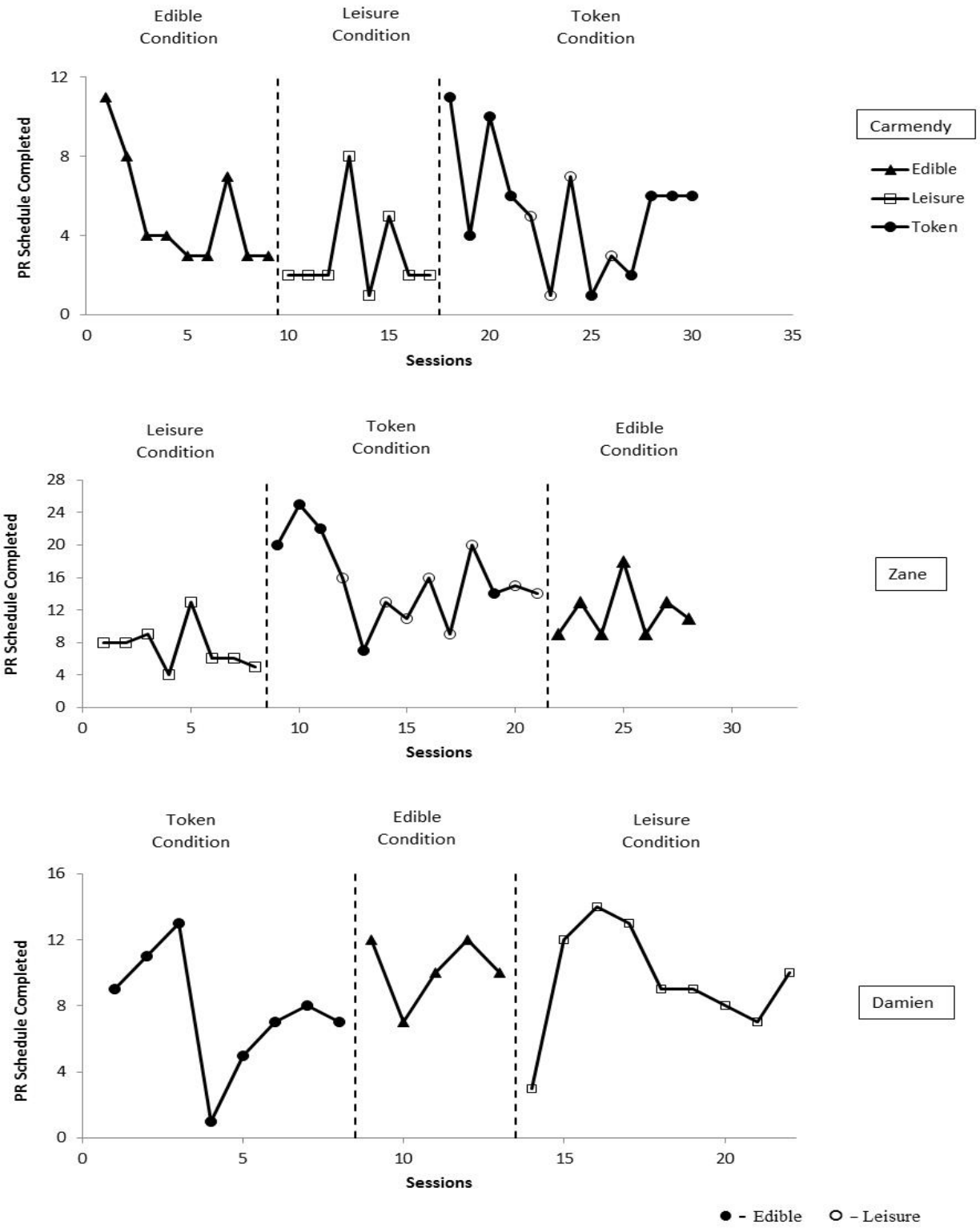

Figure 14. PR break points during Phase 1 for Carmendy (top panel), Zane (middle panel), and Damien (bottom panel). In the token condition, closed circles indicate sessions that tokens were exchanged for edible items and open circles indicate selection of leisure items.

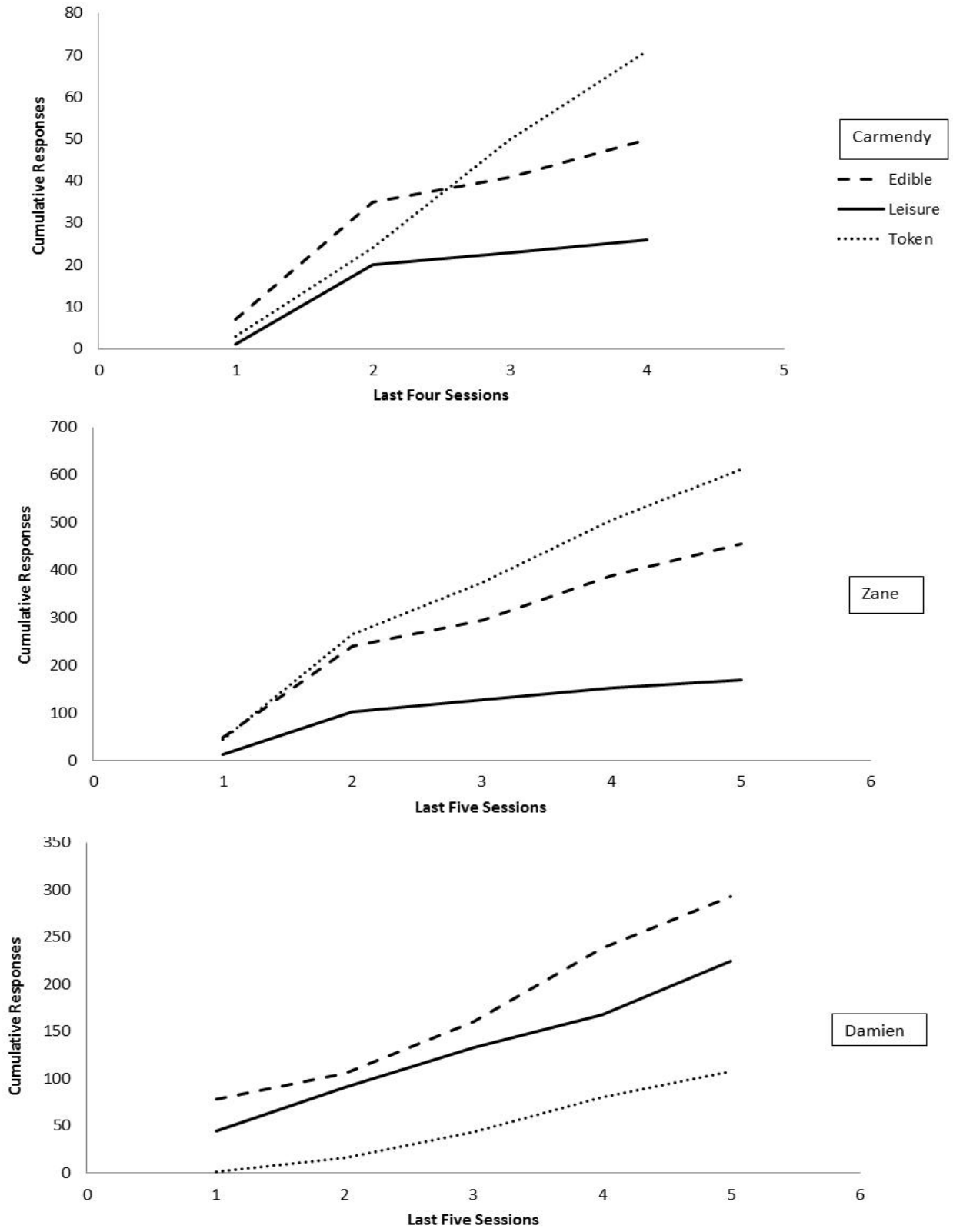

Figure 15. Cumulative number of responses during Phase 1 for the last four sessions for Carmendy (top panel) and last five sessions for Zane (middle panel) and Damien (bottom panel).

REFERENCES

Ayllon, T., & Azrin, N. H. (1968a). *The token economy: A motivational system for therapy and rehabilitation.* New York: Appleton-Century-Crofts.

Baer, D. M., Peterson, R. F., & Sherman, J. A. (1967). The development of imitation by reinforcing behavioral similarity to a model. *Journal of the Experimental Analysis of Behavior, 10,* 405-416.

Bambara, L. M., Ager, C., & Koger, F. (1994). The effects of choice and task preference on the work performance of adults with severe disabilities. *Journal of Applied Behavior Analysis, 27,* 555–556.

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

Cowles, J. T. (1937). Food-tokens as incentives for learning by chimpanzees. *Comparative Psychological Monographs, 12,* 1-96.

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.

DeLeon, I. G., Iwata, B. A., Goh, H., & Worsdell, A. S. (1997). Emergence of reinforcer preference as a function of schedule requirements and stimulus similarity. *Journal of Applied Behavior Analysis, 30,* 439-449.

DeLeon, I. G., Iwata, B. A., & Roscoe, E. M. (1997). Displacement of leisure reinforcers by food during preference assessments. *Journal of Applied Behavior Analysis, 30,* 475-484.

Fisher, W. W., & Mazur, J. E. (1997). Basic and applied research on choice responding. *Journal of Applied Behavior Analysis, 30,* 387-410.

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

Gutierrez, A., Vollmer, T. R., Dozier, C. L., Borrero, J. C., Rapp, J. T., Bourret, J. C., et al. (2007). Manipulating establishing operations to verify and establish stimulus control during mand training. *Journal of Applied Behavior Analysis, 40,* 645-658.

Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior, 91,* 257-286

Hagopian, L. P., Long, E. S., & Rush, K. S. (2004). Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28,* 668-677.

Hodos, W. (1961). Progressive ratio as a measure of reward strength. *Science, 134,* 943-944.

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., et al. (1994). The functions of self-injurious behavior: An experimental epidemiological analysis. *Journal of Applied Behavior Analysis, 27,* 215-240.

Jerome, J., & Sturmey, P. (2008). Reinforcing efficacy of interactions with preferred and nonpreferred staff under progressive-ratio schedules. *Journal of Applied Behavior Analysis, 41,* 221-225.

Karsten, A. M., & Carr, J. E. (2009). The effects of differential reinforcement of unprompted responding on the skill acquisition of children with autism. *Journal of Applied Behavior Analysis, 42,* 327–334.

Kazdin, A. E. (1978). *History of behavior modification: Experimental foundations of contemporary research.* Baltimore, MD: University Park Press.

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

Kelleher, R. T. (1956). Intermittent conditioned reinforcement in chimpanzees. *Science, 124,* 679-680.

Kelleher, R. T. (1958). Fixed-ratio schedules of conditioned reinforcement with chimpanzees. *Journal of the Experimental Analysis of Behavior, 1,* 281-289.

Lannie, A.L., & Martens, B. K. (2004). Effects of task difficulty and type of contingency on students' allocation of responding to math worksheets. *Journal of Applied Behavior Analysis,* 37, 53-65.

McAdam, D. B., Klatt, K. P., Koffarnus, M., Dicesare, A., Solberg, K., Welch, C., et al. (2005). The effects of establishing operations on preferences for tangible items. *Journal of Applied Behavior Analysis, 38,* 107-110.

McComas, J. J., Thompson, A., & Johnson, L. (2003). The effects of presession attention on problem behavior maintained by different reinforcers. *Journal of Applied Behavior Analysis, 36,* 297-307.

Murphy, E. S., McSweeney, F. K., Smith, R. G., & McComas, J. J. (2003). Dynamic changes in reinforcer effectiveness: Theoretical, methodological, and practical implications for applied research. *Journal of Applied Behavior Analysis, 36,* 421-438.

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

Olenick, D. L., & Pear, J. J. (1980). The differential reinforcement of correct responses to probes and prompts in picture-name training with severely retarded children. *Journal of Applied Behavior Analysis, 13*, 77-89.

O'Reilly, M., Edrisinha, C., Sigafoos, J., Lancioni, G., Cannella, H., Machalicek, W., et al. (2007). Manipulating the evocative and abative effects of an establishing operation: Influences on challenging behavior during classroom instruction. *Behavioral Interventions, 22,* 137-145.

O'Reilly, M., Lang, R., Davis, T., Rispoli, M., Machalicek, W., Sigafoos, J., et al. (2009). A systematic examination of different parameters of presession exposure to tangible stimuli that maintain problem behavior. *Journal of Applied Behavior Analysis, 42,* 773-783

O'Reilly, M., Sigafoos, J., Lancioni, G., Rispoli, M., Lang, R., Chan, J., et al. (2008). Manipulating the behavior-altering effect of the motivating operation: Examination of the influence on challenging behavior during leisure activities. *Research in Developmental Disabilities, 29,* 333-340.

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

Parsons, M. B., Reid, D. H., Reynolds, J., & Bumgardner, M. (1990). Effects of chosen versus assigned jobs on the work performance of persons with severe handicaps. *Journal of Applied Behavior Analysis, 23,* 253-258.

Penrod, B., Wallace, M. D., & Dyer, E. J. (2008). Assessing reinforcer potency of high preference and low-preference reinforcers with respect to response rate and response patterns. *Journal of Applied Behavior Analysis, 41,* 177-188.

Piazza, C. C., Fisher, W. W., Hagopian, L. P., Bowman, L. G., & Toole, L. (1996). Using a choice assessment to predict reinforcer effectiveness. *Journal of Applied Behavior Analysis, 29,* 1-9.

Rincover, A., & Newsom, C. D. (1985). The relative motivational properties of sensory and edible reinforcers in teaching autistic children. *Journal of Applied Behavior Analysis, 18,* 237-248.

Roane, H. S., Lerman, D. C., & Vorndran, C. M. (2001). Assessing reinforcers under progressive schedule requirements. *Journal of Applied Behavior Analysis, 34,* 145-167.

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.

Roantree, C. F., & Kennedy, C. H. (2006). The paradoxical effect of presession attention on stereotypy: Antecedent attention as an establishing, not an abolishing, operation. *Journal of Applied Behavior Analysis, 39,* 381-384.

Smith, R. G., & Iwata, B. A. (1997). Antecedent influences on behavior disorders. *Journal of Applied Behavior Analysis, 30,* 343-375.

Smith, R. G., Iwata, B. A., Goh, H. L., & Shore, B. A. (1995). Analysis of establishing operations for self-injury maintained by escape. *Journal of Applied Behavior Analysis, 28,* 433-445.

Sran, S. K., & Borrero, J. C. (2010). Assessing the value of choice in a token system. *Journal of Applied Behavior Analysis, 43,* 553-557.

Sy. J. R., & Borrero, J. C. (2009). Parametric analysis of presession exposure to edible and nonedible stimuli. *Journal of Applied Behavior Analysis, 42,* 833-837.

Tiger, J. H., Hanley, G. P., & Hernandez, E. (2006). An evaluation of the value of choice with preschool children. *Journal of Applied Behavior Analysis, 39,* 1-16.

Tustin, R. D. (1994). Preference for reinforcers under varying schedule arrangements: A behavioral economic analysis. *Journal of Applied Behavior Analysis, 27,* 597-606.

Vollmer, T. R., & Hackenberg, T. D. (2001). Reinforcement contingencies and social reinforcement: Some reciprocal relations between basic and applied research. *Journal of Applied Behavior Analysis, 34,* 241-253.

Vollmer, T. R., & Iwata, B. A. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis, 24,* 279-291.

Weeks, M., & Gaylord-Ross, R. (1981). Task difficulty and aberrant behavior in severely handicapped students. *Journal of Applied Behavior Analysis, 14,* 449-463.

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

Wolfe, J. B. (1936). Effectiveness of token rewards for chimpanzees. *Comparative Psychological Monographs, 12,* 1-72.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III.* Itasca, IL: Riverside.

Worsdell, A. S., Iwata, B. A., Conners, J., Kahng, S., & Thompson, R. H. (2000). Relative influences of establishing operations and reinforcement contingencies on self-injurious behavior during functional analyses. *Journal of Applied Behavior Analysis, 33,* 451-461.