

January 01, 2011

Comparison of second-order token and tandem schedules using clinically common production and exchange values

Julie M. Stine
Northeastern University

Recommended Citation

Stine, Julie M., "Comparison of second-order token and tandem schedules using clinically common production and exchange values" (2011). *Applied Behavioral Analysis Master's Theses*. Paper 84. <http://hdl.handle.net/2047/d20001211>

**Comparison of Second-Order Token and Tandem Schedules Using
Clinically Common Production and Exchange Values**

A Thesis Presented

By

Julie M. Stine

The Department of Counseling and Applied Educational Psychology

In partial fulfillment of the requirements for the degree of

Master of Science

In the field of

Applied Behavior Analysis

Northeastern University

Boston, Massachusetts

August 2011

NORTHEASTERN UNIVERSITY

Bouvé College of Health Sciences Graduate School

Thesis Title: Comparison of Second-Order Token and Tandem Schedules Using Clinically Common
Production and Exchange Values

Author: Julie M. Stine

Department: Counseling and Applied Educational Psychology

Approved for Thesis Requirements of Master of Science Degree

Dr. Jason Bourret

Date

Dr. William Ahearn

Date

Dr. Chata Dickson

Date

Comparison of Second-Order Token and Tandem Schedules Using
Clinically Common Production and Exchange Values

Table of Contents

1.	Abstract	P. 4
2.	Introduction	
	a. Benefits of token economies.....	P. 5
	b. Basic Research.....	P. 5
	c. Applied Research.....	P. 11
	d. Summary and Purpose.....	P. 12
3.	Method	
	a. Participant.....	P. 13
	b. Setting and Materials.....	P. 13
	c. Response Measurement	P. 14
	d. Interobserver Agreement.....	P. 15
	e. Procedure.....	P. 16
4.	Results	
	a. Paired Stimulus Preference Assessment.....	P. 17
	b. Conditioned Reinforcer Assessment.....	P. 17
	c. Comparison of Second-Order Token and Tandem Schedules.....	P.18
	d. Generality Probe.....	P. 20
5.	Discussion.....	P. 21
6.	References.....	P. 24
7.	Figures.....	P. 26

Abstract

Token economies have potential application for use in the facilitation and maintenance of behavior change, yet have not been demonstrated experimentally to maximize rate of responding. Findings from basic research suggest second-order chained schedules generate slower response rates and longer pre-ratio pauses when compared with second-order variable or tandem schedules. Applied research supports the use of token economies in increasing appropriate behavior and reducing problem behavior, but has not provided an analogue to basic research. The purpose of this study is to compare responding under second-order token and tandem schedules using clinically common production and exchange values with a participant diagnosed with autism. Rate of responding on a free operant task (i.e., target touching) and an applied task (i.e., sorting clothing) were compared; for both tasks tokens were presented contingently following responses during the token condition, but were absent during the tandem condition. Results from one participant indicate longer pre-ratio pauses and duration per component under the second-order token schedule than the tandem schedule. These findings replicate basic research findings and suggest the use of tokens is contraindicated.

Keywords: token economies, pre-ratio pauses, second-order token schedules, tandem schedules

Comparison of Second-Order Token and Tandem Schedules Using
Clinically Common Production and Exchange Values

Applied behavior analysis offers a wide range of operant technology with which the maladaptive behavior of individuals with disabilities can be reduced and opportunities to acquire functional behavior may be enhanced; the use of token economies has been reviewed in literature as one application of operant methodology that may maintain behavioral change (e.g., Kazdin & Bootzin, 1972; Kazdin, 1982; Hackenberg, 2009). A token economy is a system by which generalized conditioned reinforcers, such as tokens, are established and exchanged following instances of a target behavior for a variety of back-up reinforcers (Catania, 2007).

Token economies have several benefits. Tokens can be provided at any time, without interrupting sequences of responding, and provide immediate reinforcement by reducing the latency between the target response and delivery of a back up reinforcer; additionally, the use of tokens may provide access to many reinforcing stimuli through exchange for back-up reinforcers, and has been demonstrated to maintain responding over time (Kazdin & Bootzin, 1972). These findings suggest that token economies may offer many benefits when used in clinical settings, but little applied research has evaluated this claim.

Basic research studies have evaluated a number of schedules of reinforcement in conjunction with conditioned reinforcement. Second-order schedules involve a production schedule, by which tokens or generalized conditioned reinforcers are earned, and an exchange schedule, by which accumulated tokens are exchanged for a back-up reinforcer (Catania, 2007). For instance, an FR 20 (FR 10) second-order schedule would involve a fixed ratio 10 schedule in which 10 responses would yield one token and a fixed ratio 20 schedule in which 20 tokens would yield

access to a back-up reinforcer. Second-order schedules allow for the evaluation of each element of a complex schedule of reinforcement.

Kelleher (1958) evaluated the effects of second-order, fixed ratio schedules of reinforcement with two male chimpanzees. Subjects earned poker chips by pressing a telephone key in the presence of a white light. When the white light was extinguished and a red light illuminated, depositing the poker chips into a slot yielded edibles. Responding began at FR 30 and gradually increased to FR100 before reaching a final schedule of FR 50 (FR 125). As schedule requirements increased, post-reinforcement pauses (PRP) increased with the longest pauses occurring in the earlier components of the schedule. Alternately, response rates decreased as pauses interrupted running rates following the accumulation of several poker chips; a pattern of responding was observed in which rates were high and stable or no responding occurred (i.e., pauses). Kelleher obtained the highest running rates during lower fixed ratio requirements (e.g., FR 30) as compared to higher fixed ratio schedules (e.g., FR 100).

In addition to second-order schedules, chained schedules have been used to evaluate conditioned reinforcement. Chained schedules are compound schedules that involve the arrangement of two or more schedules of reinforcement in which each component or link is signaled by a specific stimulus (Catania, 2007). The use of chained schedules in research is beneficial as the specific stimuli associated with each link of a chained schedule permit analysis of responding following each component (Kelleher & Gollub, 1962). Tandem schedules have also been used in comparisons of schedules of conditioned reinforcement. Like chained schedules, tandem schedules are compound schedules of reinforcement involving the arrangement of two or more schedules; unlike chained schedules, components or links are not associated with any discriminative stimuli (Catania). Chained and tandem schedules, when

yoked to require the same number of responses per completion, differ only in regard to the presence or absence of signaling stimuli, allowing direct comparison of the functions of these stimuli.

To assess the effects of discriminative stimuli, Kelleher and Fry (1962) evaluated the effects of chained fixed interval schedules and tandem fixed interval schedules with two male pigeons. Under chained schedules, each of the three fixed interval components was correlated with a particular key light color; a chained condition was compared with a variable chained condition that differed only in sequence of colors of light associated with each component. Under the tandem schedule, one color key light was in place until each the reinforcer was delivered following completion of the third component. Prolonged pauses were observed during the first component with positive acceleration of response rate in the second and third components. When schedules transitioned from chained to variable chained, responding resembled the pattern and rate of responding observed under the previously correlated component and gradually increased in rate with decreasing pauses. Responding under tandem schedules indicated occasional pauses during the first component and increased responding over subsequent components to higher rates than those observed during either chained schedule. Under each of the schedules, lower response rates were obtained during the first component with an increasing trend in which the highest response rates were observed in the final component.

In an additional assessment, Thomas (1964) compared three-component chained and tandem schedules of reinforcement with five pigeons using a key-pecking response. When the chained schedule was in effect, the response key was illuminated with a specific color of light per component (i.e., red, white or blue); when the tandem schedule was in effect, the response key was illuminated by the same color of light (i.e., yellow) until the third component had been

completed. These schedules were presented as components of a multiple schedule in which either the chained or tandem schedule was in effect, and was signaled when a different schedule began by the color of the key light. Response rates were higher during the tandem schedule with shorter post-reinforcement pauses than found under the chained components.

Findley and Brady (1965) compared a chained schedule to a tandem schedule of reinforcement with two adult male chimpanzees. The tandem schedule was presented first in which a red light was illuminated and 4,000 button presses were required to access a food hopper. Following extinction, the chained schedule was presented while a green light was illuminated; 4,000 button presses were required before access to a food hopper was provided, but following every 400 responses a light flashed for 0.5 s. The red and green lights were reversed per condition before returning to original settings. The participant was then given an opportunity to choose the chained or tandem schedule. Results indicated a preference for the chained schedule whenever available, regardless of the color associated with that schedule, and shorter post-reinforcement pauses following the first 30 schedule completions with no pauses during the last 60 schedules. The second participant completed 4,000 presses per light with unlimited access to the food hopper following 120,000 responses; post-reinforcement pauses were observed following every 4,000 responses. These findings suggest a preference for chained schedules over tandem schedules, and the assertion that a high volume of work can be maintained by conditioned reinforcement.

Second-order schedules of reinforcement may also be compared to simple schedules; Jwaideh (1973) compared second-order chained and tandem schedules of reinforcement with pigeons. Each component of the chained schedule featured the illumination of a response key with a differently-colored light, and the tandem schedule was conducted with the same color

response key. A total of 120 responses were required per schedule and included variation of the number of components and responses required to access reinforcement (i.e., an FR 3 [FR 40] schedule as well as an FR 5 [FR 24] schedule) to evaluate differences in running rate, post-reinforcement pauses, and interresponse time (IRT) between schedules. Results indicated a lower running rate during the chained schedule, with highest response rates occurring during the first chain and toward the middle of the three- and five-component schedules. Lower rates were also identified for the first component of each schedule. PRP values were longer during chained schedules than tandem, with exception of schedules at the lowest value in which PRP duration was equivalent; this difference was greatest at higher ratio values even though the same number of responses were required for each schedule. Time to ratio completion was higher during chained schedules as compared with tandem schedules, and higher for the five-component than the three-component chained schedule.

Jwaideh (1973) hypothesized that the stimuli present during the early links of the chained schedule may have been under weaker stimulus control and acted as an S-delta (S^Δ) in which the hues of light were more closely associated with extinction than reinforcement. These stimuli were not present during the tandem schedule, suggesting conditions under which responding was followed by reinforcement (i.e., the single color of light) were more similar throughout the tandem condition than the chained schedule. Additionally, the final link of the chain was reversed with the primary link; responding followed the same patterns as observed with the original order of presentation, with shorter PRP rates in the presence of the stimulus once associated with the final link that returned to previously observed rate durations following exposure to the reordered stimuli. When stimuli associated with the last link were presented

during the first and last links of the chained schedule, PRP rates decreased again before returning to previously observed rates.

Webbe and Malagodi (1978) evaluated second-order schedules of reinforcement, comparing fixed ratio and variable ratio components, with three male rats. During baseline, an FR 6 (FR 20) schedule was in place in which 20 lever presses resulted in the presentation of a token (i.e., a dark glass marble) with no consequences for depositing tokens into the token receptacle. Once responding was stable, a VR 6 (FR 20) was implemented, with a range of 1-14 tokens earned per exchange. Once stable, the schedule returned to FR 6 (FR 20). Response rates were higher during the variable ratio schedules, and pauses occurred earlier under the variable ratio condition than under the fixed ratio schedule, with longer pauses during the first and second components. These findings were consistent with those of Kelleher (1958).

Foster, Hackenberg and Vaidya (2001) conducted an additional review of the effects of second-order schedules of reinforcement with fixed ratio and variable ratio schedules of reinforcement with four pigeons. During the first experiment, pigeons were training to key-peck under a gradually-increased fixed ratio schedule until responding was stable under an FR 50 schedule. Response rates decreased as the fixed ratio exchange ratio increased while higher, more stable response rates were obtained when the tandem schedule was in place. Post-reinforcement pauses were longer during initial components for both schedules, but to a greater extent for the fixed ratio schedules. During the second experiment, a multiple schedule was introduced in which a different color key light signaled a VR (FR) schedule or a FR (VR) schedule. As the fixed ratio exchange schedule increased, response rates decreased and post-reinforcement pauses increased. Higher response rates were observed under the variable ratio exchange schedule with post-reinforcement pauses that were lower than those observed under the

fixed ratio schedule; this effect was more dramatic at higher values. These findings suggest response rates obtained during second-order schedules depend on the token exchange schedule in place rather than the production schedule.

Findings from basic research suggest the use of token economies in conjunction with fixed ratio schedules of reinforcement may yield slower response rates and longer post-reinforcement pauses when compared with variable ratio schedules (Kelleher, 1958; Kelleher & Fry, 1962; Webbe & Malagodi, 1978) or tandem schedules (Thomas, 1964). When compared with tandem schedules of reinforcement, increasing response requirements may have a greater effect on responding on chained schedules (Findley & Brady, 1965); lower running rates and longer post-reinforcement pauses may be obtained under chained schedules and stimuli associated with links in a chained schedule may maintain response rates when order of presentation is varied (Jwaideh, 1973). These laboratory studies suggest caution when implementing token economies for therapeutic purposes, but have not been demonstrated in clinical application with human participants.

Phillips (1968) conducted one of the first studies evaluating a token system in an applied setting. Three adolescent boys enrolled in a family-style, residential behavior modification program participated in this study in which a token system was used to increase target behavior across social, self-care and academic domains. The token system involved points provided contingently upon the occurrence of specified appropriate behavior and removed contingently upon the occurrence of specified inappropriate behavior. Following accumulation, points could be exchanged for privileges at the end of the week including leisure activities and the opportunity to engage in infrequent activities. Across all three participants, rates of inappropriate

behavior (e.g., aggressive verbal behavior) decreased and rates of appropriate behavior (e.g., bathroom tidiness and punctuality) increased.

In a subsequent study, Phillips, Phillips, Fixsen and Wolf (1971) evaluated the point-system token economy in the same applied setting – the family-style, residential behavior modification program. Six boys participated in this study in which points were awarded following instances of appropriate behavior and removed contingently upon instances of inappropriate behavior. Once accumulated, points could be exchanged for a variety of back-up reinforcers (e.g., weekend outings or spending money). Results indicated an increase in appropriate behavior with all six participants across several contexts (e.g., punctuality, cleaning tasks and saving money). The study further indicated the utility of a token economy in an applied setting to modify the behavior of participants.

Token economies have also been implemented in classroom settings; Miller and Schneider (1970) compared the acquisition of writing skills with 22 underprivileged children. Participants were divided into two groups, one that earned tokens for completing activities and a control group that did not earn tokens, and given a writing pretest. The children in the token group were gradually introduced to the token system such that the children first experienced each trade-in activity, were required to earn a token before trading in for an activity and then only allowed to trade in tokens during a designated leisure time. A free token condition was included as well, to allow the children to access the leisure activities without completing the writing task. Results indicated participants in the token group completed the writing practice sheets when tokens could be earned. Additionally, though participants in the token and control groups performed at a similar level during the pretest (i.e., about 35% accuracy), the token group demonstrated large gains in writing skills (i.e., 100% accuracy) over the control group (i.e., 38%).

Applied research findings have demonstrated the potential benefits of using token economies, including increases in appropriate behavior and skill acquisition with decreases in inappropriate behavior (Phillips, 1968; Phillips et al., 1971; Miller & Schneider, 1971). In many clinical settings the implementation of token economies has become standard practice, yet it is unclear if economies are maximizing performance. Research conducted within applied settings has not approximated the methodology used in basic research, such as the use of an arbitrary response to evaluate response rate under different schedules of reinforcement, yet these findings may be relevant to the use of token economies in applied settings. Further, token economies are frequently used in settings with participants with developmental disabilities; little research has evaluated the effects of tokens with this population. The purpose of the current study is to evaluate the effects of second-order, fixed-ratio token and fixed ratio tandem schedules using naturally occurring production and exchange values with a participant diagnosed with autism.

Method

Participant

John was a 17 year-old male diagnosed with autism and mental retardation; he was enrolled in a residential program at a school for individuals diagnosed with autism and other developmental disabilities. He communicated through the use of static pictures and an adaptive electronic device in addition to gestures or the use of several modified American Sign Language manual signs. John had used a token economy as part of his clinical programming for several years prior to this study.

Setting and Materials

All sessions were conducted in a classroom with limited distractions at the participant's school. During sessions, John was seated across from the experimenter in a room with a table and one chair. A cardboard response board measuring 89.5 cm by 28 cm was used to record a target-touching response. The response board was covered with green construction paper and at each end of the response board, a rectangle of laminated white paper measuring 21.5 cm by 14 cm, or half a sheet of standard-sized paper,

was attached with Velcro®. These white rectangles were used as targets. A simple token board constructed of black laminated construction paper, measuring 21.5 cm by 14 cm, was present during token sessions. Small strips of Velcro® attached to the token board denoted the number of tokens required per schedule (i.e., five strips under the FR 5 [FR 1] schedule and two strips under the FR 2 [FR 1] schedule). Red poker chips with a diameter of 4 cm were used as tokens. A small digital timer and a transparent plastic cup containing a small piece of an edible were present during all sessions. All sessions were videotaped.

Response Measurement

The dependent variable was *target touching*, defined as contact between any part of the participant's hand and one targets followed by contact between any part of the participant's hand and the opposite target without touching the table, chair, or green-papered response board between the targets. Contact between the participant's hand and any part of the response board outside of the targets was not followed by any consequences, but the participant was required to touch each target once before additional responses were accepted. All sessions were scored electronically by the primary therapist using data collection software and a laptop computer. Data were collected for response board duration, frequency of responses, frequency of token delivery during token sessions, frequency of reinforcer delivery, and duration of consumption time. PRP were calculated by subtracting the time at which the first target-touching response occurred from the time at which the response board was presented. Consumption time was calculated by subtracting the time at which the reinforcer was delivered and the response board removed from the time at which the response board was represented; consumption time was removed to allow for the calculation of rate of responding using only session time in which the response board was present.

An additional response was selected for generality probe sessions. The dependent variable was *sorting clothing* and was defined as picking up one article of clothing from a pile and placing the item into a plastic bin with a corresponding article of clothing (e.g., placing a sock into the bin with a sock in it). The researcher blocked attempts to pick up more than one article of clothing. If an incorrect response

occurred (i.e., if an item was placed into the incorrect bin), the item was removed and placed at the bottom of the pile of clothing to be sorted. Data were collected as denoted above with one difference – data were collected for the duration of time the bins were available rather than response board duration. Data were analyzed and entered using the procedure described above.

Response rate for each session was calculated by removing consumption time from session time and dividing the total number of responses by the remaining session time. Component duration was calculated by averaging the time, measured in seconds, to complete each component within the FR 5 or FR 2 schedule across all sessions within a phase. For each phase, the same number of responses were required before a back up reinforcer was delivered, allowing for comparison of the time to component completion. Component duration for the last six schedule completions within each phase was calculated by average the time, in seconds, to complete each component within the FR 5 or FR 2 schedule across the final session within a phase.

Interobserver Agreement

Interobserver agreement (IOA) data were collected for 34.6% of sessions using the target-touching response. Prior to collecting IOA, a secondary observer scored a minimum of three sessions with the primary therapist and was required to score 90% or greater agreement for all three sessions. The trained, secondary observer used the same data collection software to score sessions electronically. To calculate IOA, session data scored by the primary therapist were compared with session data scored by the secondary observer using IOA software (i.e., Instant IOA). This program divided session time into 10-s bins to compare data entry in intervals, yielding an occurrence, nonoccurrence, and proportional agreement percentage for each session scored. Average occurrence agreement was 92.0% (range: 67.4 – 100%), average nonoccurrence agreement was 90.8% (range: 75 – 100%), and average proportional agreement was 91.6% (range: 78.8 – 98.9%).

IOA data were also collected for the generality probe sessions using the method described above. Occurrence, nonoccurrence, and proportional agreement were calculated for 33.3% of sessions by a secondary observer; average occurrence agreement was 92.8% (range: 85.3 – 100%), average

nonoccurrence agreement was 94.1% (range: 84.4 – 100%), and average proportional agreement was 93.4% (range: 86.6 – 96.9 %).

Procedure

Paired stimulus preference assessment. A paired stimulus preference assessment (Fisher et al., 1992) was conducted to identify a highly preferred edible. Eight healthy edible items were compared including the following foods: green olive, orange, green apple, chick pea, banana, black olive, popcorn, and baked chips. During assessment, a small piece of an edible was placed at the center of a Styrofoam plate and two plates were placed in front of the participant per trial.

Conditioned reinforcer assessment. A conditioned reinforcer assessment was conducted to evaluate whether an existing token economy functioned as a conditioned reinforcer; during this assessment, target touching was evaluated. Sessions were terminated after five min had elapsed. A multielement design was used in which a second-order token schedule (i.e., FR 5 [FR 1]) was alternated with extinction sessions. During token sessions, the response board was placed in front of the participant. Following the first four responses, a token was placed on the token board; after the fourth response, the response board was removed and a one-minute pause elapsed. This pause was designed to interrupt the temporal contiguity between the target-touching response and the delivery of a token. After one min, the participant was prompted to engage in an alternate response (i.e., hand clapping) and the last token was placed on the board. The tokens were then exchanged for an edible and the session ended. During extinction sessions, the response board was placed in front of the participant but no consequences were delivered for target touching.

Schedule comparison with target-touching response. Following these assessments, sessions were conducted in a multielement design with alternating second-order token and tandem schedules. During token sessions, the response board was placed in front of the participant with a token board and plastic cup containing an edible. Across all token sessions, an FR 1 production schedule was used such that following every target-touching response a token was placed on the token board. Two exchange schedules were evaluated – FR 2 and FR 5 – in which the participant was required to accumulate two or five tokens

before exchanging tokens for an edible. After this exchange, the response board was removed until the participant had finished consuming the edible. During tandem sessions, the response board was again placed in front of the participant with a plastic cup containing an edible. The number of responses required to access the edible was yoked to the token schedule. For instance, if the token schedule was FR 5 (FR 1) in which five tokens were required to exchange for an edible, an FR 5 schedule was used for the tandem schedule in which five responses were required to access the edible. The only difference between token and tandem sessions was the presence of the tokens and token board – the response requirement was the same across schedules.

Generality probe. Finally, a generality probe was conducted using an applied task in which the participant sorted clothing into bins. Sessions were conducted in a multielement design with alternating second-order token and tandem schedules. During token sessions, three plastic bins were placed in front of the participant with a token board and plastic cup containing an edible. An FR 5 exchange schedule was evaluated in which the participant was required to accumulate five tokens before exchanging token for an edible. Following an exchange, the plastic bins were removed until the participant finished consuming the edible. During tandem sessions, the plastic bins were again placed in front of the participant with a plastic cup containing an edible. A schedule of FR 5 was used in which five responses were required to access the edible. As with the assessment described above, the only difference between token and tandem sessions was the presence of the token and token board with the same response requirement for both sessions.

Results

The results from the paired stimulus preference assessment are displayed in Figure 1. Green olives were the only edible selected 100% of the time it was offered and thus were provided throughout the duration of this study following schedule completion. All other edible items were selected less than 75% of offered opportunities including: oranges (71.4%), green apples (57.1%), chick peas (57.1%), bananas (50.0%), black olives (42.9%), popcorn (21.4%), and baked chips (7.1%); these items were selected as healthy alternatives due to participant weight concerns.

Figure 2 displays the results from the conditioned reinforcer assessment. The participant responded at a higher rate during token sessions (range: 30.7 – 41.4 responses per minute) than extinction sessions (range: 0.4 – 16.7 responses per minute), indicating that tokens functioned as a conditioned reinforcer for target touching.

The results from the comparison of second-order token and tandem schedules are displayed in Figure 3 as a rate of responses per minute. A second-order token schedule of FR 5 (FR 1) was selected as the participant was using this schedule of reinforcement in clinical programming at the time of this assessment; a schedule of FR 5 was in place during tandem sessions such that the same number of responses was required in each condition to access an edible. The rate of responding was similar across the first 27 sessions with a slightly higher rate of responding obtained during the tandem condition. Sessions 28-40 yielded differentiated responding in which the participant engaged in a higher rate of target touching during the tandem condition (range: 45.5 – 62.9 responses per minute), in the absence of tokens, than during the token condition (range: 39.1 – 46.0 responses per minute). Stability criteria were reached for both the token and tandem schedules after session 40. All stability measures were calculated using response rates from the last six sessions. The last three sessions met all criteria measures with all three data points falling within 20% of the mean, without all data points falling above or below the mean.

At session 41, a second-order token schedule of FR 2 (FR 1) and a tandem schedule of FR 2 were introduced to determine whether a smaller number of responses required to access an edible would yield different response rates. The participant responded at a higher rate during the tandem condition in the absence of tokens (range: 29.1 – 40.4 response per minute) with lower rates obtained during the token condition (range: 26.9 – 31.1 responses per minute). Stability criteria were met and the schedules were changed to FR 5 (FR 1) for the second-order token condition and FR 5 for the tandem condition. Again, a higher rate of responding was obtained for the tandem condition (range: 43.9 – 65.7 responses per minute) than the token condition (range: 38.1 – 52.8 responses per minute), with a similar rate of responding as obtained in the first presentation of this schedule. Stability criteria were met by session 85 for both conditions.

A replication of the FR 2 (FR 1) second-order token schedule and the FR 2 tandem schedule began at session 86. Sessions were not run for a few weeks due to an increase in the participant's problem behavior; when continued, a decrease in rate of responding occurred for both conditions. Rate of responding was slightly higher in the token condition (range: 12.7 – 22.6 responses per minute) compared to the tandem condition (range: 8.9 – 19.5 responses per minute) for sessions 86-93 – a finding not previously observed. Sessions were paused for one week due to participant behavior. At session 94, sessions recommenced and response rates rose to levels observed during the previous FR 2 (FR 1) token and FR 2 tandem schedules. A higher rate of responding was obtained during the tandem condition (range: 31.9 – 36.2 responses per minute) than the token condition (range: 22.9 – 28.7 responses per minute). Stability criteria were met for the final three sessions of each schedule at session 107.

In addition to rate of responding, the average duration per component was analyzed. Figure 4 displays these results as an average duration of time, measured in seconds, to complete each component for all sessions within that phase. A longer duration per component was obtained during the first FR 5 (FR 1) second-order token schedule (range: 1.1 – 3.1 s) as compared to the duration per component in the first FR 5 tandem schedule (range: 0.9 – 2.4 s). A longer PRP was obtained for the first component of both token and tandem schedules, with a longer duration in the first component. A longer duration was also obtained for the first FR 2 (FR 1) token schedule (1.0 s and 1.9 s, respectively) and the FR2 tandem schedule (1.3 s and 0.8 s, respectively); again, a longer PRP was obtained for the first component as compared to the second component.

The replication of the FR 5 (FR 1) token and FR 5 tandem schedule yielded slightly longer component duration in the token schedule (range: 0.7 – 1.8 s) as compared to the tandem schedule (range: 0.5 – 1.5 s). A longer PRP was again obtained for the first component as compared with later components, and was larger for the token schedule than the tandem schedule. Finally, the replication of the FR 2 (FR 1) token and FR 2 tandem schedule resulted in longer duration during the token schedule (1.5 s and 2.7 s, respectively) than the tandem schedule (2.5 and 0.9 s, respectively). A longer PRP was

again noted in the first component for each schedule and was longer than PRP rates obtained in the first FR 2 (FR 1) and replication of the FR 5 (FR 1) schedules.

Component duration was then analyzed for the last six sessions of each phase as these data were used to evaluate stability criteria before changing phases and can be seen in Figure 5. During the first FR 5 (FR 1) token schedule and the FR 5 tandem schedule, the average component duration for the first component of the token condition was longer (range: 0.9 – 2.3 s) than the tandem condition (range: 0.5 – 1.7 s) across all components. When the first FR 2 (FR 1) token schedule and FR 2 tandem schedule were in place, the token condition again yielded a longer duration per component (range: 1.0 – 2.1 s) than the tandem condition (range: 0.8 – 1.3 s). The replication of the FR 5 (FR 1) token and FR 5 tandem schedule yielded a higher duration per component under the token condition (range: 0.7 – 1.9s) than the tandem condition (range: 0.5 – 1.3 s). Finally, a replication of the FR 2 (FR 1) token condition and FR 2 tandem condition indicated a higher duration per component under the token condition (range: 1.3 – 20.5 s) than the tandem (range: 0.8 – 1.4 s). The longest component durations were obtained during the final phase – the replication of the FR 2 (FR 1) token and FR 2 tandem conditions.

The results from the generality probe are displayed in Figure 6, as a rate of responses per minute. A second-order token schedule of FR 5 (FR 1) and a tandem schedule of FR 5 were compared. A total of five responses were required to access the cup with the edible and the only difference between schedules was the presence of the token board and tokens during the token condition only. Response rates were higher during the tandem condition (range: 19.8 – 25.8) for all but session 10 when compared to the token condition (range: 15.7 – 21.8). The same stability criteria used above was evaluated for the generality probe, and met for both conditions by session 18.

The results from the average component duration obtained during the generality probe are displayed in Figure 7. Longer component durations were obtained during the token condition in which tokens and the token board were present (range: 2.3 – 4.1 s) as compared to the tandem condition when tokens were not present (range: 2.1 – 2.9 s). The longer durations were observed during the first component for both

the token and tandem conditions with the longest PRP occurring during the first component of the token condition.

Finally, the results from the average component duration for the last six schedule completions are displayed in Figure 8. A longer PRP can be observed in the first component for both the token and tandem conditions as compared with subsequent components. A longer component duration was obtained for all token components (range: 2.3 – 3.8 s) relative to tandem components (range: 2.0 – 3.1 s) but rates were similar across both conditions.

Discussion

This study evaluated the use of token economies in a clinical setting, with a participant diagnosed with autism. Token economies have been widely used with individuals with autism, but have not been demonstrated to function beneficially. If the use of tokens is advantageous, it should yield higher response rates when tokens and token boards are present than when the same task is presented without tokens; similarly, the time required to complete a response should be of a smaller duration when token economies are in use than when absent. The findings from the current study, however, suggest another conclusion. After an initial undifferentiated pattern of responding, a higher rate of responses per minute was obtained under the tandem schedule in which tokens and a token board were absent. Up to 20 more responses per minute occurred during the tandem condition than the token condition. This pattern of responding was obtained under the FR 5 (FR 1) second-order token condition and FR 5 tandem schedule, and replicated, as well as the FR 2 (FR 1) second-order token and FR 2 tandem schedule. Additionally, the duration of time to complete each component indicated a similar pattern of responding in which a longer period of time was required to complete each component when tokens and a token board were present when compared to the tandem condition in which they were absent.

The results from the generality probe yielded similar patterns of responding. A higher overall rate of responding was obtained during the tandem condition as compared to the token condition, and a shorter average component duration was obtained during the tandem condition as compared with the token condition. These findings indicated a slower overall response rate and longer component duration when

tokens were present, suggesting the use of tokens is contraindicated for an arbitrary response (i.e., target touching) and an applied task (i.e., sorting clothing).

These findings are consistent with previous research such that overall response rates were higher under the tandem schedule than the token schedule (Foster, Hackenberg, & Vaidya, 2001; Jwaideh, 1973). The tandem schedule may have functioned as a variable ratio, or VR, schedule in which after a specified number of responses, a reinforcer is delivered. Unlike the token schedule, there are no stimuli presented to signal the proximity of reinforcer delivery. The participant may respond at a higher rate as the faster responding occurs, the faster a reinforcer is delivered. Under the token schedule, the delivery of initial tokens may function as an S-delta, signaling that more responses must occur before reinforcement is forthcoming; as a result, longer PRP may occur as earlier components are completed as compared with later components (Jwaideh).

The finding that the PRP was greater during initial components as compared with subsequent components is also consistent with previous research (Foster, Hackenberg, & Vaidya, 2001; Jwaideh, 1978; Thomas, 1968; Webbe & Malagodi, 1978). A higher PRP was obtained during the FR token schedule than the tandem schedule under two schedule manipulations. Under both schedules, the PRP decreased as additional components were completed and as the proximity to the delivery of a reinforcer decreased. This finding was replicated across both the FR 5 (FR 1) token, FR 5 tandem and FR 2 (FR 1) token, FR 2 tandem schedules.

The current study extends previous research in its application in a clinical setting with a participant diagnosed with autism. Schedules values were selected to approximate the token economy in place at the time of this study. Tokens were demonstrated to function as a conditioned reinforcer, increasing response rates during assessment, yet decreased response rates when compared with a tandem schedule of reinforcement. This pattern of responding was observed when a target-touching response was evaluated as well as the clothes-sorting response; the addition of this assessment extends previous research by assessing patterns of responding with an applied task, demonstrating the effects of tokens with more than

one task. These results suggest that the use of token economies should not be assumed to function as a conditioned reinforcer, but should be assessed to determine the specific effects of implementation.

One limitation to this study is the scope of generality of the results. A comparison of responding under token and tandem schedules of reinforcement was compared using an arbitrary task and an applied task that was in mastery. It is possible that token economies function differently when used with various tasks, such as a task in acquisition as opposed to a task in mastery. Another limitation involves the manipulation of the schedules of reinforcement. This study evaluated two exchange schedules – FR 5 and FR 2 – but only evaluated one production schedule (i.e., FR 1). It is possible that different production schedules may yield different patterns of responding. Future research should conduct this assessment with additional participants and evaluate additional tasks. Different production schedules should be evaluated to determine the effects of manipulating response requirement to access reinforcement.

References

- Catania, C. A. (2007). *Learning, Interim (Fourth) Edition*. Cornwall-on-Hudson, NY; Sloane Publishing, LLC.
- Findley, J. D., & Brady, J. V. (1965). Facilitation of large ratio performance by use of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 8, 125-129.
- 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.
- Jwaideh, A. R. (1973). Responding under chained and tandem fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior*, 19, 259-267.
- Kazdin, A.E. (1982). The token economy: a decade later. *Journal of Applied Behavior Analysis*, 15, 431-445.
- Kazdin, A. E., & Bootzin, R. R. (1972). The token economy: An evaluative review. *Journal of Applied Behavior Analysis*, 5, 343-372.
- Kelleher, R. T. (1958). Fixed-ratio schedules of conditioned reinforcement with chimpanzees. *Journal of Comparative Physiological Psychology*, 49, 571-575.
- Kelleher, R. T., & Fry, W. T. (1962). Stimulus functions in chained fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 5, 167-173.
- Kelleher, R. T. and Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 5, 543-597.
- Miller, L. K. & Schneider, R. (1970). The use of a token system in project head start. *Journal of Applied Behavior Analysis*, 3, 191-197.
- 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.

- Phillips, E. L., Phillips, E. A., Fixsen, D. L., & Wolf, M. M. (1971). Achievement place: Modification of the behaviors of pre-delinquent boys within a token economy. *Journal of Applied Behavior Analysis, 4*, 45-59.
- Thomas, J. R. (1964). Multiple baseline investigation of stimulus functions in an FR chained schedule. *Journal of the Experimental Analysis of Behavior, 7*, 241-246.
- Webbe, F. M. & Malagodi, E. F. (1978). Second-order schedules of token reinforcement: comparisons of performance under fixed ratio and variable ratio exchange schedules. *Journal of the Experimental Analysis of Behavior, 30*, 219-224.

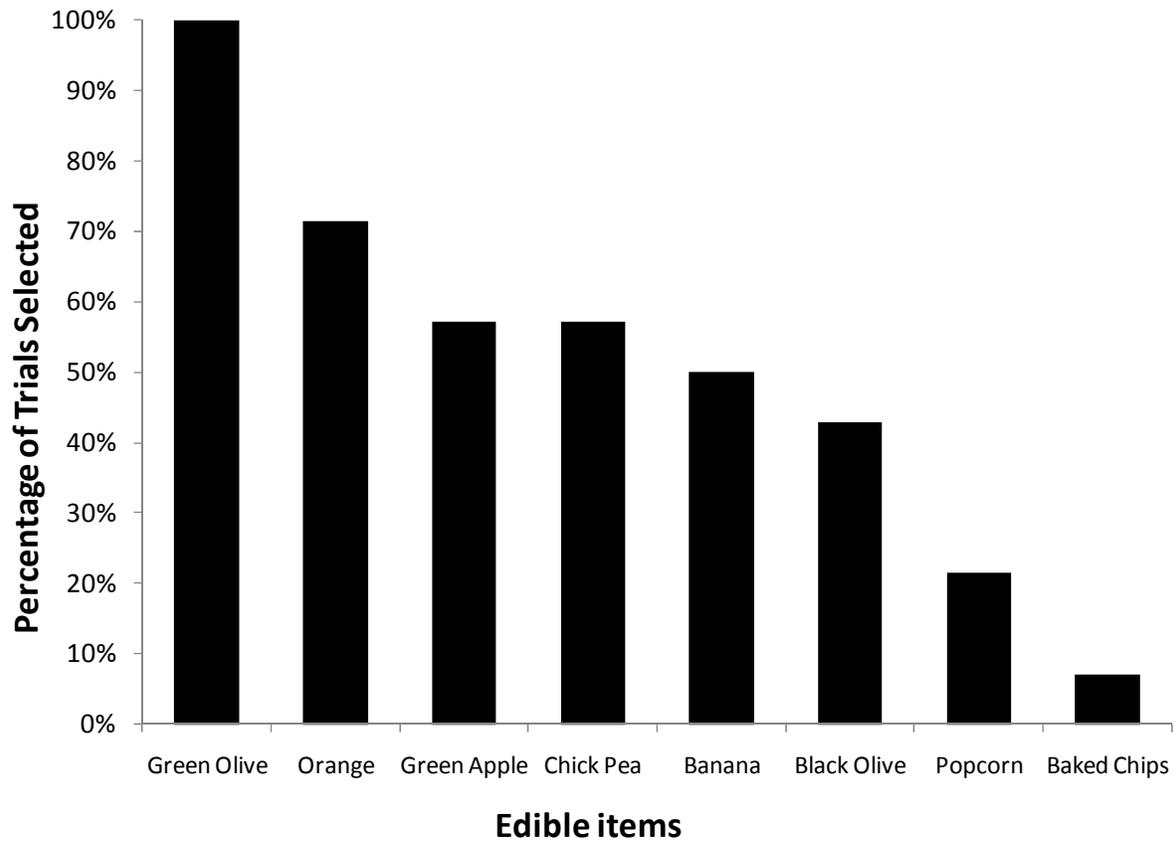


Figure 1. Displays the results from a paired stimulus preference assessment. Edible items are displayed across the x-axis with percentage of trials selected on the y-axis.

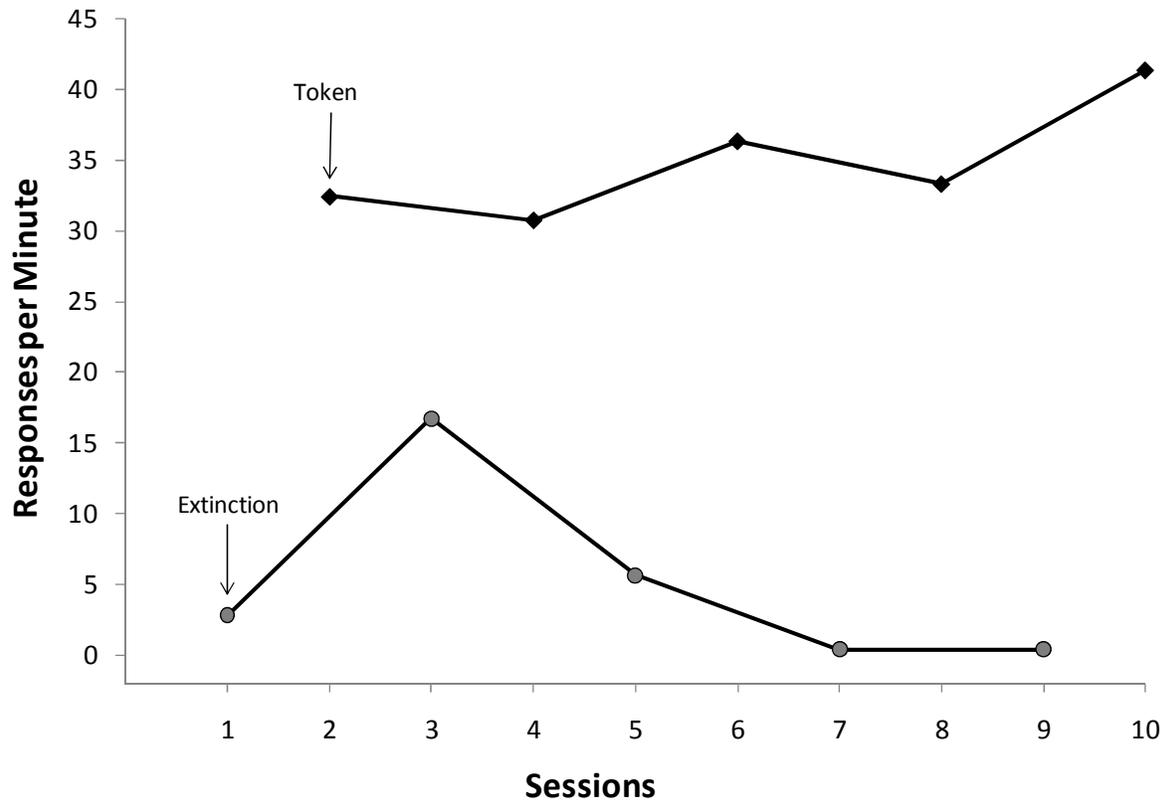


Figure 2. Displays the results from the conditioned reinforcer assessment. Sessions are displayed on the x-axis with responses per minutes on the y-axis. Token sessions are represented by the black, filled diamonds and extinction sessions by the gray circles.

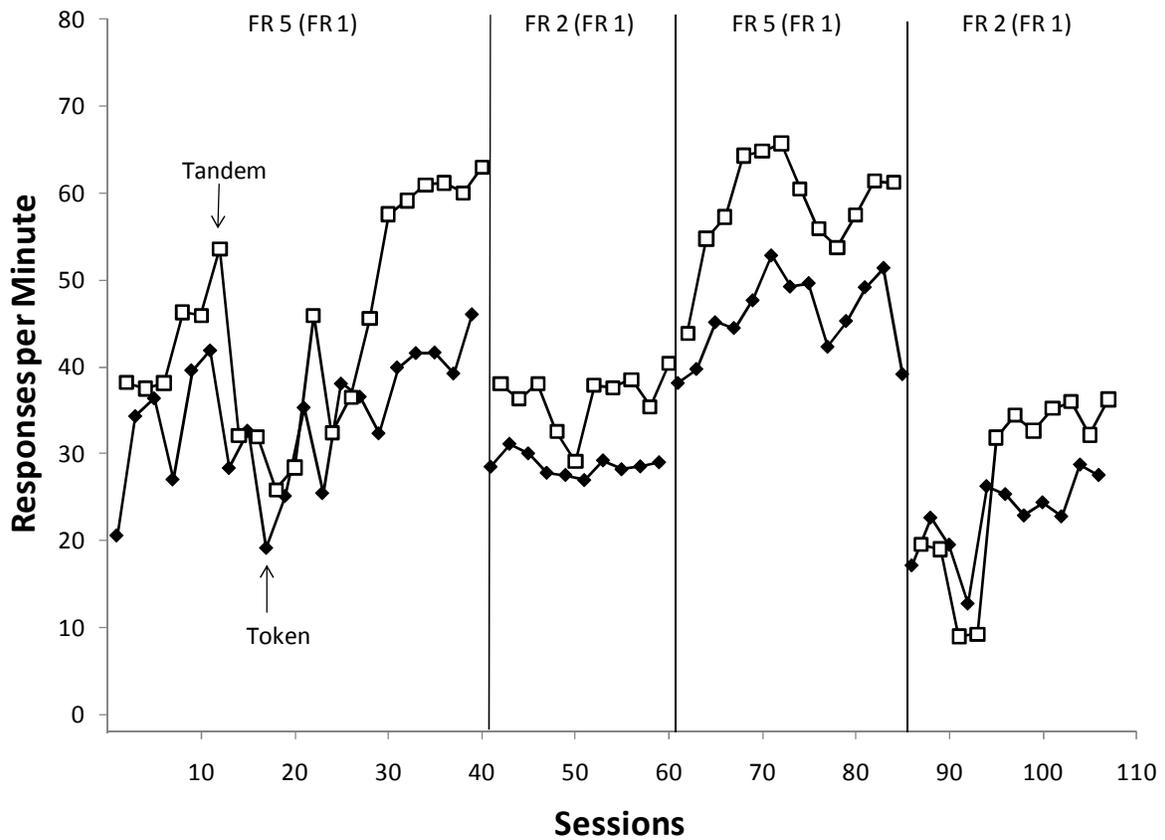


Figure 3. Displays response rates from the comparison of second-order token and tandem schedules. Sessions are displayed on the x-axis with responses per minutes on the y-axis. Token sessions are represented by the black, filled diamonds and tandem sessions are represented by the white squares.

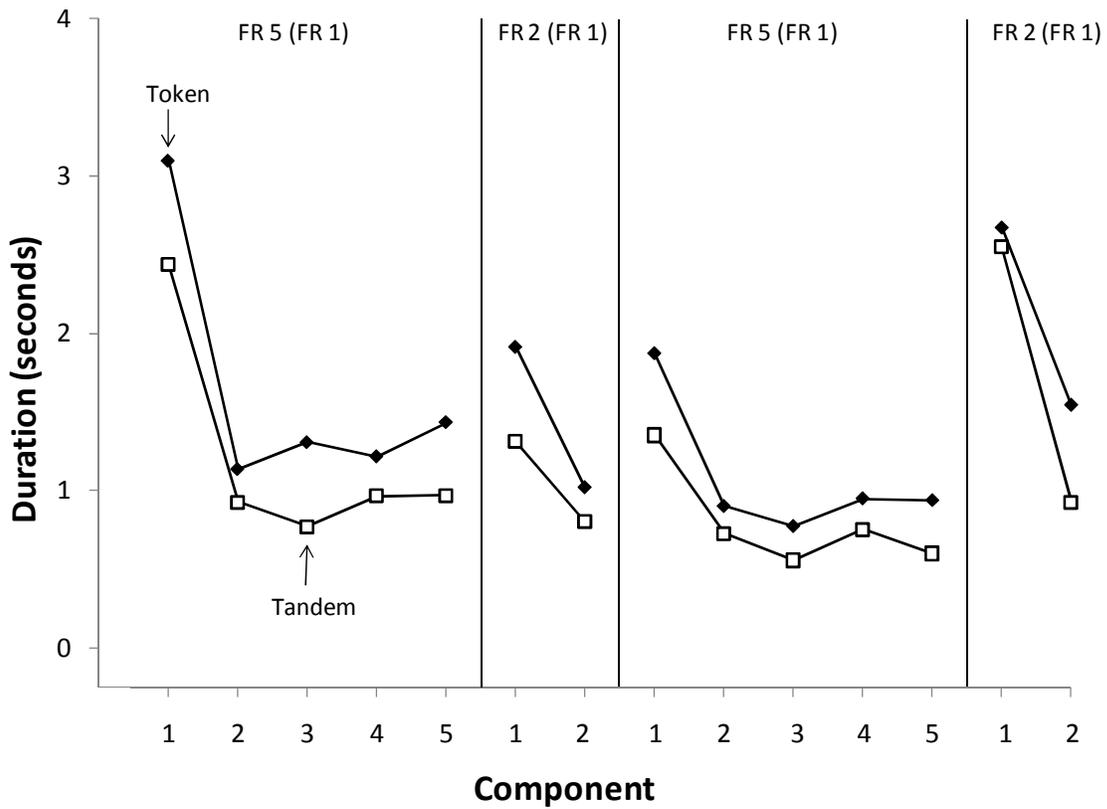


Figure 4. Displays the results from the average component duration obtained during the comparison of second-order token and tandem schedules. Components per schedule are displayed on the x-axis with average duration to component completion, measured in seconds, across the y-axis. Token sessions are represented by the black, filled diamonds and tandem sessions are represented by the white squares.

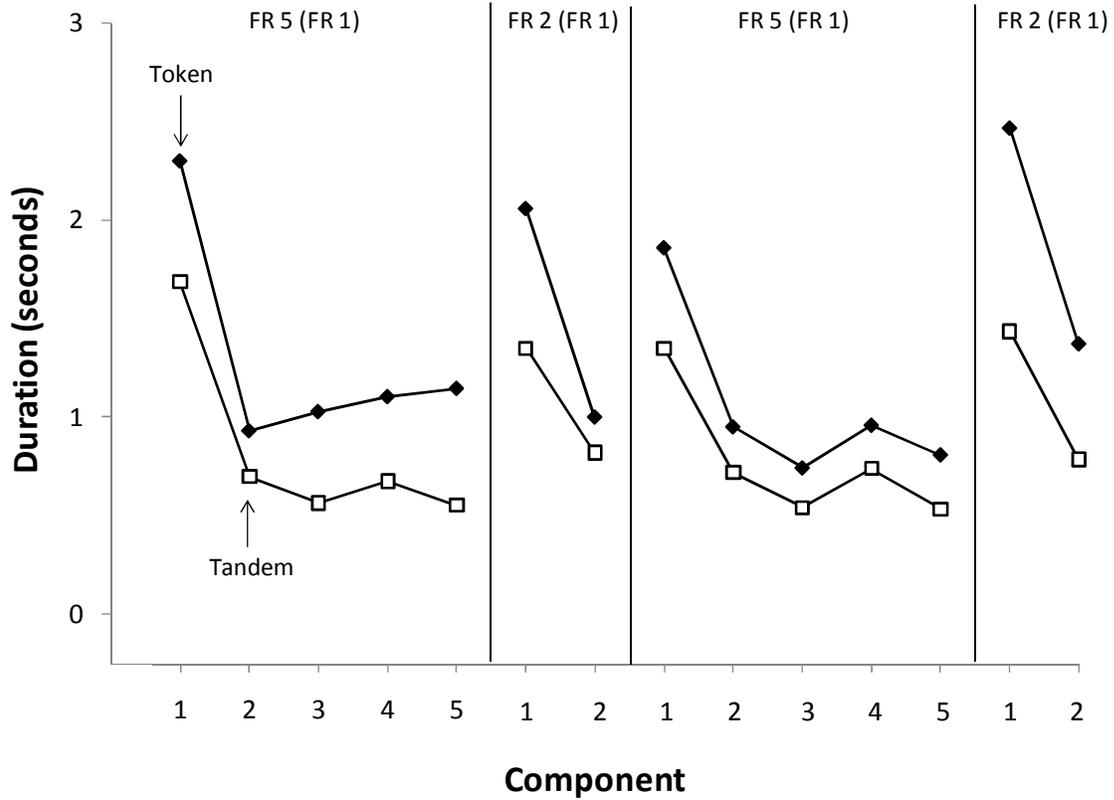


Figure 5. Displays the results from the component duration of the last six schedule completion obtained during the comparison of second-order token and tandem schedules. Components per schedule are displayed on the x-axis with the average duration to component completion for the last six schedule completions, measured in seconds, across the y-axis. Token sessions are represented by the black, filled diamonds and tandem sessions are represented by the white squares.

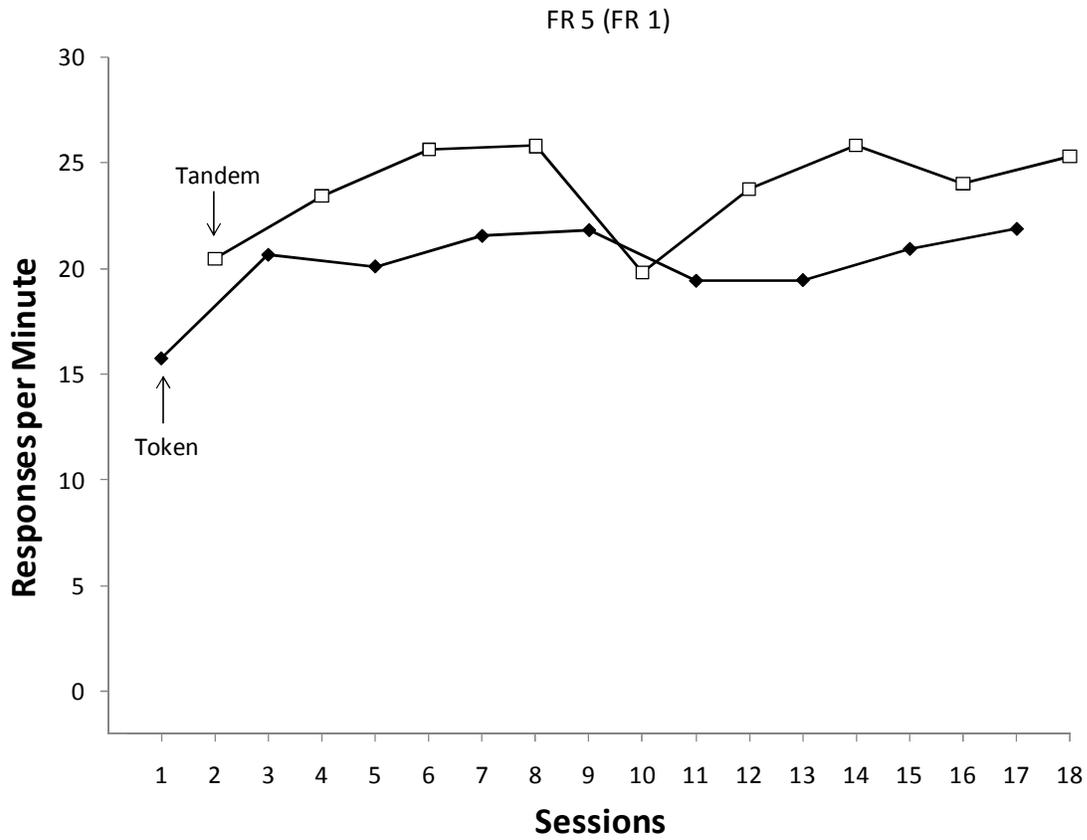


Figure 6. Displays the results from the comparison of second-order token and tandem schedules using an applied task. Sessions are displayed on the x-axis with responses per minutes on the y-axis. Token sessions are represented by the black, filled diamonds and tandem sessions are represented by the white squares.

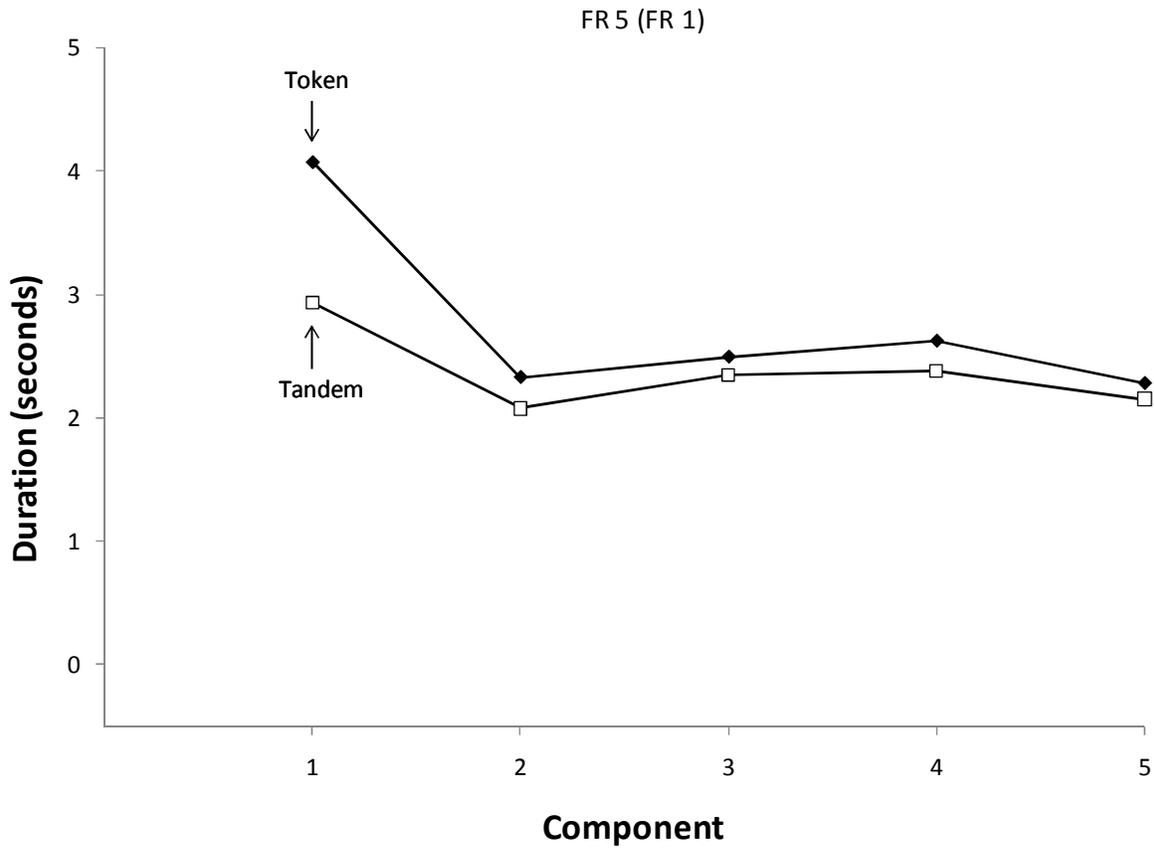


Figure 7. Displays the results from the average component duration obtained during the comparison of second-order token and tandem schedules using an applied task. Components per schedule are displayed on the x-axis with average duration to component completion, measured in seconds, across the y-axis. Token sessions are represented by the black, filled diamonds and tandem sessions are represented by the white squares.

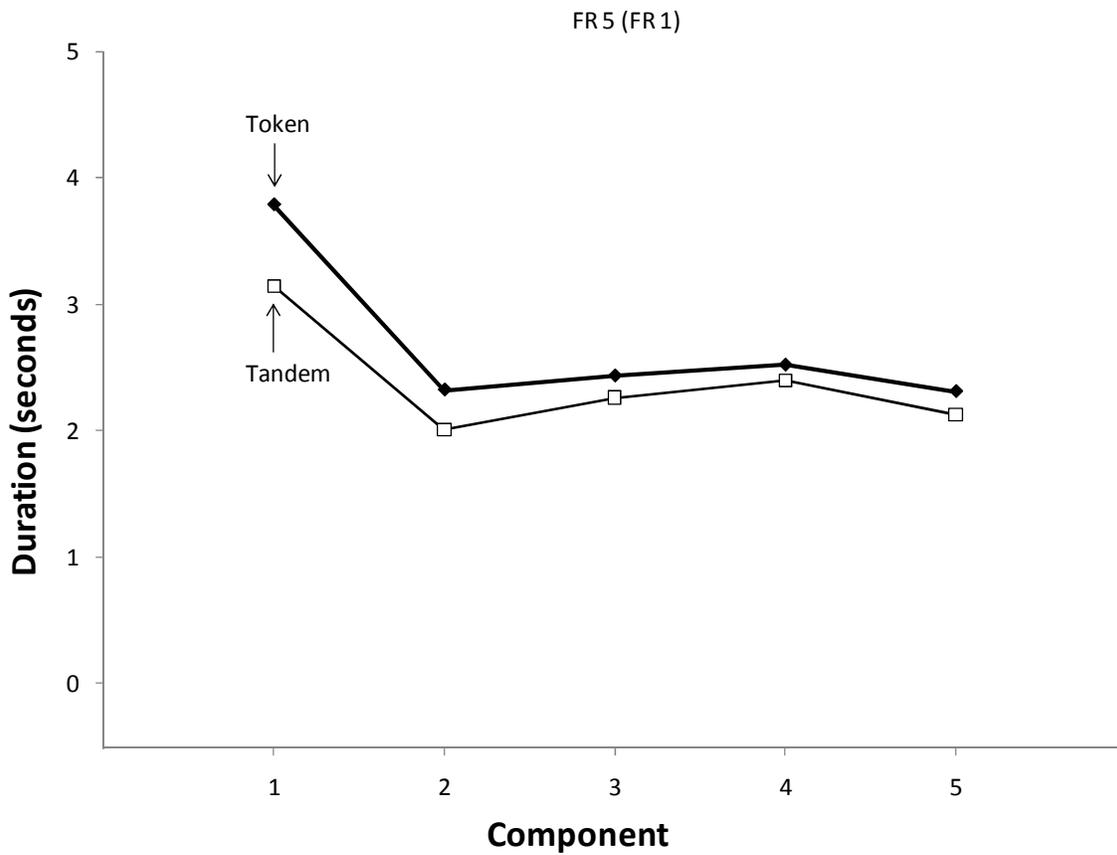


Figure 8. Displays the results from the component duration of the last six schedule completions obtained during the comparison of second-order token and tandem schedules using an applied task. Components per schedule are displayed on the x-axis with the average duration to component completion for the last six schedule completions, measured in seconds, across the y-axis. Token sessions are represented by the black, filled diamonds and tandem sessions are represented by the white squares.