

RESEARCH-TO-PRACTICE AND PRACTICE-TO-RESEARCH: CLOSING THE GAP
THROUGH IMPROVING IMPLEMENTATION OF TOKEN ECONOMIES AND
REDUCING CLASSROOM TRANSITIONS

by

JEFFREY FRANK HINE

(Under the Direction of Scott P. Ardoin)

ABSTRACT

The increasing emphasis on implementation of evidence-based practices has been accompanied by a growing awareness of the gap between research and practice, and has become a matter of concern. This dissertation presents two manuscripts linked by the interaction between research and practice. The first manuscript progressed from the need to connect *research to practice* and highlights problematic scenarios encountered during implementation of token economies and how revisiting basic experimental research can guide practitioners as to how to integrate this seemingly esoteric research into their everyday practices. The second manuscript exemplifies the progression from *practice to research*. Conceived through practical observations (and subsequent frustrations), it presents a research study that is directly applicable to everyday problems confronted in practice: that of inefficient and lengthy transitions.

INDEX WORDS: Token economies, Experimental analysis of behavior, Transitions, Classroom behavior management, Computer-assisted instruction

RESEARCH-TO-PRACTICE AND PRACTICE-TO-RESEARCH: CLOSING THE GAP
THROUGH IMPROVING IMPLEMENTATION OF TOKEN ECONOMIES AND
REDUCING CLASSROOM TRANSITIONS

by

JEFFREY FRANK HINE

B.A., University of California at Santa Barbara, 2002

M.Ed., Peabody College of Vanderbilt University, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

© 2014

Jeffrey Frank Hine

All Rights Reserved

RESEARCH-TO-PRACTICE AND PRACTICE-TO-RESEARCH: CLOSING THE GAP
THROUGH IMPROVING IMPLEMENTATION OF TOKEN ECONOMIES AND
REDUCING CLASSROOM TRANSITIONS

by

JEFFREY FRANK HINE

Major Professor: Scott P. Ardoin

Committee: Kevin M. Ayres
Amy L. Reschly
Nathan A. Call

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
Research-to-Practice	2
Practice-to-Research	3
References	6
2 TOKEN ECONOMIES: USING EXPERIMENTAL RESEARCH TO GUIDE PRACTICAL APPLICATIONS	9
Abstract	10
Introduction	11
Part I: Conditioning Tokens as Effective Reinforcers	12
Part II: Acknowledging and Investigating First- and Second-order Schedules of Reinforcement	24
Conclusion	41
References	42
3 DECREASING TRANSITION TIMES AND INCREASING ON-TASK BEHAVIOR IN ELEMENTARY SCHOOL CLASSROOMS: AN EVALUATION OF THE <i>KEEP BUSY AND CARRY ON</i> SYSTEM	54

Abstract	55
Introduction.....	56
Method	61
Results.....	74
Discussion.....	79
References.....	85
4 CONCLUSION.....	89

LIST OF TABLES

	Page
Table 3.1: Phase Means and Ranges for Baseline, Intervention, Withdrawal, and Reinstatement Conditions	91
Table 3.2: Phase Means and Ranges for All Baseline, Intervention, Withdrawal, and Reinstatement Conditions for Teachers' Delivery of Prompts and Praise Statements	92

LIST OF FIGURES

	Page
Figure 3.1: Effects of class-wide intervention on transition duration and on-task behavior for Leo and Samantha	93
Figure 3.2: Effects of class-wide intervention on transition duration and on-task behavior for Jonathan and Jamie	94
Figure 3.3: Effects of class-wide intervention on frequency of teacher prompts and praise statements.....	95

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Extensive resources and energy are directed toward activities designed to promote the use of evidence-based practices. Evidence-based practices are generally defined as those that have been recommended by respected clinical experts and professionals as well as those for which there is some level of empirical evidence regarding its effectiveness (APA Presidential Task Force on Evidence-Based Practice, 2006). The increasing emphasis on implementation of evidence-based practices has been accompanied by a growing awareness of the gap between research and practice, and has become a matter of concern. Research is intended to generate socially useful knowledge and it is assumed that research is supposed to make a difference and improve practice. The reality is, despite these intentions, the gap between research and practice remains, and there are many views on why this gap exists and what to do about it (Stoiber & Kratochwill, 2000). Kennedy (1997) suggested that the gap is due in part to inaccessibility of research to classroom teachers while others (Boardman, Arguelles, Vaughn, Hughes, & Klingner, 2005) have suggested a lack of trust among teachers of claims made within research. Cochran-Smith (2005) noted that regardless of the presumed causes, the gap between research and practice in education is of critical importance because research should be the foundation from which teaching and learning practices are developed and improved. This dissertation presents two manuscripts linked by the interaction between research and practice.

The manuscript presented in Chapter 2 progressed from the need to connect *research* to *practice* and highlights problematic scenarios encountered during implementation of token

economies and how revisiting basic experimental research can guide practitioners as to how to integrate this seemingly esoteric research into their everyday practices. The second manuscript is the result of observations recognized during my time as an educational-behavioral consultant to classrooms. This study exemplifies the progression from *practice* to *research*. Conceived through practical observations (and subsequent frustrations), Chapter 3 presents a research study that is directly applicable to everyday problems confronted in practice: that of inefficient and lengthy transitions.

Research-to-Practice

Researchers most often publish findings in professional journals. Teachers often may not have continuing contact with these journals and subsequently might find reading research a low priority. Carnine (1997) points out that if it is difficult and time-consuming for teachers to locate and interpret research, they are not likely to make the effort. Abbott, Walton, Tapia, and Greenwood (1999) suggested that because of differing reward systems, teachers and researchers ask different questions about instruction and speak different languages. Teachers speak the understandable language of the general population, whereas researchers speak the technical language of research design and experimental control. To be useable and accessible, research findings must be clearly written and deal with topics of importance to practitioners. Addressing questions grounded in practice, focusing on interventions that are efficient and manageable to implement, and broadening the context for successful research demonstrations provide an excellent basis for efforts to improve the perceived and actual usability of research (Carnine, 1997). Consequently, before practitioners can successfully implement research-validated instructional strategies in the classroom, researchers must “translate” this knowledge into teacher-friendly instructional forms.

Hackenberg (2009) suggested token reinforcement systems “stand as among the most successful behaviorally-based applications in the history of psychology” (p 257). Multiple reviews (e.g., Kazdin, 1982; Maggin, Chafouleas, Goddard, & Johnson, 2011) designate token economies as evidence-based and universally applicable to many contexts. The multitude of reports demonstrating token reinforcement’s effectiveness is too long of a list for the purposes of this dissertation but it is well-known that token economies are a popular choice for parents, teachers, and behavior analysts alike. Since these programs are widely used across various contexts, individuals, and behaviors, learning how to appropriately implement and troubleshoot these programs is relevant to all groups. Despite the fact that the flexibility of token economies may have led to its widespread use, all too often practitioners are *too* flexible in implementation, often leading to treatment failure. Many reports have suggested that evidence-based interventions often fail due to fidelity of implementation (Noell et al., 2005; Witt, Noell, LaFleur, & Mortenson, 1997); furthermore, token economies require well-informed and systematic decision-making in order to effectively establish and monitor the program when unexpected results surface. Although most practitioners have a general understanding of token reinforcement and related behavioral principles, a more comprehensive knowledge base of the underlying mechanisms of token reinforcement may be warranted. The purpose of Chapter 3 is to describe the intricacies of conditioned reinforcement and reinforcement schedules and provide practitioners a brief guide for trouble shooting token economies.

Practice-to-Research

Many have observed that transitions tend to compromise a substantial proportion of a child’s school day. Some studies estimate that anywhere from 9-30% of a student’s school day is spent engaging in tasks unrelated to academics (Berk, 1976; Vitiello, Booren, Downer, &

Williford, 2012; Ysseldyke, Christenson, Thurlow, & Bakewell, 1999). Given the problems created for teachers by excessive transitions and the degree to which longer transitions decrease actual instructional time, reducing transition times within schools is of utmost importance. Although there are some evidence-based interventions that teachers can use to decrease transitions (e.g., Sainato, 1990; Campbell & Skinner, 2004), these interventions are unlikely to be as effective if teachers do not implement as prescribed within supporting research. Moreover, research dictates that even if teachers find an intervention acceptable, are trained on the procedures, and use the treatment correctly while being observed, most teachers do not maintain adequate treatment implementation across during long-term implementation (Noell et al., 2005; Witt et al., 1997). Thus, teachers may be less likely to use evidence-based practices without consistent cues and feedback on their own behavior. Fortunately, advances in technology and subsequent increases in the use of computer-assisted instruction (CAI) may allow teachers to use pre-packaged interventions that automate some of the components that they typically would implement themselves (e.g., Stromer, Kimball, Kinney, & Taylor, 2006). Automated components also might provide prompts and cues for appropriate behavior for both students *and* teachers. CAI is more likely to be free of human error and eliminates some problems associated with procedural fidelity. Implementing the intervention in this way can help close the research-to-practice gap by allowing teachers to use a pre-packaged intervention backed by research without the need for lengthy training and intense follow-up.

Overall, the outcome of the research-to-practice gap is often frustration and dissatisfaction. Completing practice-based research is time-consuming and labor intensive. Solutions to these barriers and problems will bring enormous benefits to teachers and students. The following studies will hopefully serve as models for ways in which practitioners can

translate and apply past research to guide practical applications of common interventions, as well as how to conduct research to solve common problems observed in practice.

References

- Abbott, M., Walton, C., Tapia, Y., & Greenwood, C. R. (1999). Research to practice: A blueprint for closing the gap in local schools. *Exceptional Children, 65*(3), 339-352.
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *The American Psychologist, 61*(4), 271.
- Berk, L. E. (1976). How well do classroom practices reflect teacher goals? *Young Children, 32*, 64-81.
- Boardman, A. G., Argüelles, M. E., Vaughn, S., Hughes, M. T., & Klingner, J. (2005). Special education teachers' views of research-based practices. *The Journal of Special Education, 39*(3), 168-180.
- Campbell, S., & Skinner, C. H. (2004). Combining explicit timing with an interdependent group contingency program to decrease transition times: An investigation of the timely transitions game. *Journal of Applied School Psychology, 20*(2), 11-27.
- Carnine, D. (1997). Bridging the research-to-practice gap. *Issues in Educating Students with Disabilities, 363-373*.
- Cochran-Smith, M. (2005). Teacher educators as researchers: Multiple perspectives. *Teaching and Teacher Education, 21*(2), 219-225.
- Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior, 91*(2), 257-286.
- Kazdin, A. E. (1982). The token economy: A decade later. *Journal of Applied Behavior Analysis, 15*(3), 431-445.
- Kennedy, M. M. (1997). The connection between research and practice. *Educational Researcher, 26*(7), 4-12.

- Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology, 49*(5), 529-554.
- Noell G. H., Witt J. C., Slider N. J., Connell J. E., Gatti S. L., Williams K. L.,...& Duhon, G. J. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*, 87-106.
- Sainato, D. M. (1990). Classroom transitions: Organizing environments to promote independent performance in preschool children with disabilities. *Education & Treatment Special Issue: Organizing Caregiving Environments for Young Children with Handicaps, 13*, 288-297.
- Stoiber, K. C., & Kratochwill, T. R. (2000). Empirically supported interventions and school psychology: Rationale and methodological issues—Part I. *School Psychology Quarterly, 15*(1), 75.
- Stromer, R., Kimball, J. Kinney, E. M., & Taylor, B. A. (2006). Activity schedules, computer technology, and teaching children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 21*(1), 14-24.
- Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly, 27*, 210-220.
- Witt J. C, Noell G. H, LaFleur L. H, & Mortenson B. P. (1997). Teacher use of interventions in general education settings: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30*, 693–696.

Ysseldyke, J. E., Christenson, S. L., Thurlow, M. L., & Bakewell, D. (1989). Are different kinds of instructional tasks used by different categories of students in different settings? *School Psychology Review, 18*, 98–111.

CHAPTER 2

TOKEN ECONOMIES: USING BASIC EXPERIMENTAL RESEARCH TO GUIDE PRACTICAL APPLICATIONS¹

¹Hine, J. F., Ardoin, S. P., & Call, N. A. To be submitted to *Behavior Analysis in Practice*.

Abstract

This paper highlights the applicability of patterns seen within basic-experimental token-economy research. Through this review, multiple barriers that encumber practical application of token economies are presented. Following the discussion of each barrier is a review of the basic research founded in the distinct and often overlooked features of conditioned reinforcement (Part I) and reinforcement schedules (Part II). Revisiting this research may assist practitioners in identifying potential barriers to client success and implementing a more fundamentally sound and thus effective intervention. To further guide practitioners in using this knowledge in everyday settings, each area concludes with (a) recommendations specific to each barrier conceived through basic research and (b) relevant applied research and practical examples.

INDEX WORDS: Token economy, Conditioned reinforcement, Experimental analysis of behavior, Schedules of reinforcement, Motivating operations

Introduction

Since first proposed by Ayllon and Azrin (1968) and subsequently refined by Kazdin (Kazdin, 1977; Kazdin & Bootzin, 1972), the use of token economies has become one of the most venerable and widespread applied interventions for producing behavior change (Kazdin, 1982; Matson & Boisjoli, 2009). The use of token economies has an adequately strong research base to be applied across contexts (Christophersen, Arnold, Hill, & Quilitch, 1972), ages (Christensen, Young, & Marchant, 2004; Conyers et al., 2004), and behaviors (e.g., Higgins, Williams, McLaughlin, 2001; LeBlanc, Hagopian, Maglieri, 2000; McLaughlin & Malaby, 1972). Given this widespread use and effectiveness across settings, many practitioners (e.g., teachers and applied behavior analysts) may have a general understanding of the procedures behind establishing a token economy. However, during implementation of a token economy, practitioners may encounter complex barriers to behavior change and may struggle to enact solutions targeting those barriers (Bailey, Gross, & Cotton, 2011; Drabman & Tucker, 1974; Kazdin, 1982). Revisiting foundational basic research on the underlying mechanisms of token economies can assist practitioners in overcoming such situations encountered in practice (McIlvane, 2009). Widespread failure to implement token economies without an understanding of these mechanisms likely has an unseen impact on client progress *and* the reputation of the field of behavior analysis.

Much basic research exists demonstrating the effects of the underlying mechanisms of token economies (for review see Foster, Hackenberg, & Vaidya, 2001; Hackenberg, 2009). Ideally, practitioners would consult this literature when faced with a practical dilemma. A variable that often interferes with this venture, however, includes the considerable effort required for practitioners to assimilate and apply information gained from reading basic experimental

research (Carr & Briggs, 2010). Practitioners might view basic research as inapplicable to everyday practice; yet, general patterns of performance found with animals in the laboratory consistently emerge in applied research (for discussion see Mace & Critchfield, 2010). This paper will highlight the applicability of these patterns that may assist practitioners in identifying potential barriers to client success and implementing a more fundamentally sound and thus effective intervention. This paper is divided into two parts. Within each part, multiple barriers that encumber practice are presented. Following the discussion of each barrier, is a review of the basic research founded in the distinct and often overlooked features of conditioned reinforcement (Part I) and reinforcement schedules (Part II). To further guide practitioners in using this knowledge in everyday settings, each area concludes with recommendations specific to each barrier conceived through this basic research and relevant applied research and examples.

Part I: Conditioning Tokens as Effective Reinforcers

Although there may be some differences in the specifics of establishing a token economy (e.g., Drabman & Tucker, 1974; Miltenberger, 2008; O’Leary & Drabman, 1971), there seems to be general consensus that establishing an effective token economy should at least include the following:

- 1) Identifying and operationally defining appropriate target behaviors.
- 2) Selecting appropriate tokens (e.g., durable, engaging, individualized).
- 3) Identifying backup reinforcers (e.g., primary reinforcers, other conditioned reinforcers).
- 4) Determining values of tokens and exchange rates for backup reinforcers.
- 5) Determining methods of exchange.
- 6) Determining how clients can earn or lose tokens (e.g., rules).

- 7) Accurately monitoring the program's effects on the target behaviors.
- 8) Adjusting the program to meet the long-term goals and addressing barriers to success.

If practitioners implement these steps in a consistent and systematic manner, positive behavior changes are likely to occur. In general, practitioners can use the above framework as a “base” behavior management system with which practitioners can enact multiple options.

Before tackling barriers encountered during implementation of token economies, a brief introduction to tokens acting as conditioned reinforcers is also warranted. A conditioned reinforcer has been defined as an initially neutral event or stimulus acquiring value through its relation to primary reinforcers and subsequently can serve as an effective independent reinforcer (Skinner, 1974; Williams, 1994a, 1994b). Comprehensive reviews such as Fantino (1977), Kelleher (1966), and Williams (1994a) collectively demonstrate that several species' response rates will increase if responding produces conditioned reinforcers. Perhaps the most widely cited investigations of the effects of tokens as conditioned reinforcers are the classic token reinforcement studies of Wolfe (1936) and Cowles (1937). Through these programs of research, chimpanzees learned various tasks with poker chips as rewards. The chimpanzees were taught to insert tokens into a vending machine for primary reinforcers (e.g., food). Contingent presentation of tokens maintained responding across multiple experiments even if the chimpanzees were not allowed to exchange the tokens until the end of an experimental session. Malagodi (1967a, 1967b, 1967c) added to this research by demonstrating that rats acquired new responses through use of token reinforcement alone and that token-specific response rates were similar to those seen under primary reinforcement. Conditioned reinforcers such as tokens can also maintain behavior over a sustained period of time without that behavior ever being contingently followed by the primary reinforcer (e.g., Zimmerman, Hanford, & Brown, 1967).

Undoubtedly, there are various ways in which token economies are not successful. As stated before, practitioners, and those that they train, may suspect a decrease in client progress is due to an inherent flaw in the intervention, rather than to the casual and untailed design and implementation of intervention components. When token economies are not as effective as projected, Kelleher and Gollub (1962) suggested the need for practitioners to investigate a number of general factors relating to the effectiveness of the token. These general factors include (a) the quality of the backup stimuli, (b) the pairing schedule, and (c) motivating operations in relation to backup stimuli.

Barrier: Insufficient Quality of Backup Reinforcers

One barrier with multiple possible solutions includes the case in which the token has not been established as an effective conditioned reinforcer. One way this problem may become evident is when the client ceases to readily exchange tokens for previously used backup reinforcers. If this is the case, decreased responding will likely ensue and the client may discard tokens instead of exchanging them. An initial area that practitioners need to investigate is the quality of the backup stimuli.

Basic experimental research. Some manipulable dimensions of reinforcement found to increase the likelihood of one response over another within token economies include reinforcer rate (Baum, 1974; Herrnstein, 1970), magnitude, and quality of reinforcement (Mace & Roberts, 1993). Quality of reinforcement is often described as involving reinforcer potency or efficacy (i.e., prior establishment as an effective reinforcer) and one way to measure preference is to consider stimuli that are reliably selected as highly preferred in stimulus preference assessments (e.g., Neef, Mace, Shea, & Shade, 1992; Neef, Shade, Miller, 1994). That is, a client's preference for the to-be-paired stimuli will undoubtedly influence a token's effectiveness as a

conditioned reinforcer. For instance Wolfe (1936) demonstrated chimpanzee's proclivity to select tokens that had been paired with food (as opposed to nothing) and tokens that had been paired with two pieces of food (rather than one). Additional animal studies demonstrate that with all other dimensions of reinforcement held constant (e.g., amount and rate) subjects will bias responding toward reinforcers of higher quality. This phenomenon has been shown in rats (wheat vs. brain stimulation; Hollard & Davidson, 1971), cows (hay vs. dairy meal; Matthews & Temple, 1979), and pigeons (buckwheat vs. hemp; Miller, 1976). Thus, if manipulation of quality of primary reinforcement is an effective method of biasing responding, it will likely impact the reinforcing effectiveness of tokens paired with the primary (or backup) reinforcers.

In applied settings. Tokens will not become effective reinforcers if they are paired with stimuli that are not of sufficient quality or that have not been established as effective reinforcers themselves. A large body of applied literature has identified numerous strategies for selecting stimuli most likely to function as effective backup reinforcers (DeLeon & Iwata, 1996; Roane, Vollmer, Ringdahl, & Marcus, 1998). These include single stimulus, free operant, paired stimulus, and multiple stimulus without replacement formats. Given clients' preferences may fluctuate over time (Hanley, Iwata, & Roscoe, 2006), practitioners might consider having an assortment of backup stimuli and institute periodic assessment of preferences. Systematically rotating preferred items can maintain the reinforcing properties of the backup stimuli. For example, DeLeon, Anders, Rodriguez-Catter, and Neidert (2000) demonstrated that providing access to only a single set of toys limited the effectiveness of their intervention due to satiation effects. Instead, when providing access to a rotating set of toys, the percentage of time the participant contacted the toys (competing response) increased; thus, automatically maintained self-injurious behavior was reduced.

Overall, fluctuating preference for backup stimuli is likely to influence (both positively and negatively) the reinforcing properties of the tokens with which they are paired. Furthermore, tokens can maintain their reinforcing properties despite fluctuating preferences if tokens can be exchanged for a variety of high quality reinforcers. Although, tokens as “generalized reinforcers” will be discussed in later sections, it is important to note that one of the real advantages of token reinforcers is that they can be paired with many reinforcers, and as such, become much more flexible as a reinforcer in treatment programs. That is, if tokens can effectively be exchanged for many different backup reinforcers, the convenience and social validity of the program increases by not having to require practitioners to keep a wide range of reinforcers constantly and immediately available.

Barrier: Insufficient or Inconsistent Pairing

Another factor that might impede consistent responding (after ensuring quality backup reinforcers) is the association of the token with the backup reinforcer. These associations arise through the original token-backup pairing and how often this pairing is presented to the client.

Basic experimental research. Early basic research by Wolfe (1936) and Cowles (1937) demonstrated the importance of pairing tokens with primary reinforcers by training chimpanzees to respond differentially to tokens with exchange value (exchangeable for primary reinforcers) as opposed to those without exchange value. A token will not likely have a reinforcing influence over an organism’s behavior if the token was not paired with the backup stimuli a sufficient number of times or temporally close enough in time. Williams and Dunn (1991) provided some evidence for the necessity of token-backup pairing through a series of three experiments examining conditioned reinforcement in pigeons. Overall, the effectiveness of conditioned reinforcers depended on the frequency with which the stimulus was paired with the primary

reinforcer as well as how often the stimulus was followed by reinforcement. With rats, Kelleher and Gollub (1962) also noted the significance of the number of pairings between the eventual conditioned stimuli (i.e., tokens) and primary reinforcers and suggested an optimal interval between token and backup would be 0.5-1 s.

Research investigating respondent conditioning further supports that with repeated pairing, the token should retain the reinforcing properties of the backup reinforcer without the client needing to engage in any behaviors outside of accepting and consuming the reinforcer (Kelleher & Gollub, 1962; Shahan, 2010; Williams, 1994a). For instance, pigeons will begin pecking a lighted key when repeatedly exposed to the paired presentation of an activated grain hopper and the lighted key (Hearst, 1975). Shahan (2010) equates this circumstance to the principles of respondent conditioning that result in stimuli acquiring the capacity to act as conditioned stimuli when paired with unconditioned stimuli (Mackintosh, 1974; Williams 1994a). Except in this case, neutral stimuli (tokens) acquire the capacity to function as reinforcers when paired with primary reinforcers. Research demonstrating application of conditioned reinforcers to shape new responses (e.g., Malagodi, 1967a, 1967b, 1967c; Kelleher & Gollub, 1962) provides further supporting evidence for the importance of the foundational relationship between tokens and backups.

In applied settings. If the token does not seem to function as a conditioned reinforcer, practitioners will most likely need to pair the two stimuli more frequently, more consistently, or temporally closer. Before ever requiring the client to engage in a behavior in order to gain access to the token (e.g., exchange behavior), practitioners would benefit from repeatedly and contiguously pairing the token with a backup reinforcer. When establishing tokens as conditioned reinforcers, practitioners can model their protocol after procedures described by

Foster et al. (2001) who trained non-human subjects to reliably exchange tokens for access to edible stimuli. For instance, the practitioner can fill all spaces on a 15-space token board and on a variable-time schedule (e.g., 30 s) remove a token from the board and immediately give the client a small piece of his or her favorite food (or other primary/backup reinforcer). This process would be repeated until the client accepts both the token and primary reinforcer a majority of the time. In an applied study, Moher, Gould, Hegg, and Mahoney (2008) successfully established tokens as conditioned reinforcers by pairing tokens with backup reinforcers in two stages. The first stage involved the experimenter delivering a backup reinforcer within 0.5 s of delivering a token noncontingently. In the second stage, the participant was required to physically exchange the token for the backup reinforcer. When evaluated in a preference assessment, tokens contingently paired with highly preferred edibles became preferred stimuli themselves. This study provided validation that tokens can be established as conditioned reinforcers that match the reinforcement value of the backups with which they are paired.

Barrier: Overcoming Problematic Effects of Motivating Operations

Even after ensuring a strong pairing between tokens and backup reinforcers, clients may still engage in inconsistent responding. One potential cause is the variable effectiveness of backup stimuli from moment to moment. In an effort to prevent inconsistent effectiveness of backup reinforcers (and thus responding), practitioners can investigate the effects of motivating operations. By definition, motivating operations can alter the reinforcing effectiveness of reinforcers either by increasing (establishing) or decreasing (abolishing) the effectiveness of a given consequence (Michael, 1993; Laraway, Snyckerski, Michael, & Poling, 2003; Vollmer & Iwata, 1991). It might be the case that motivating operations are affecting responding in unforeseen ways.

Basic experimental research. Inconsistent responding due to the effects of motivating operations may be neutralized (or used to one's advantage) by creation of generalized conditioned reinforcers. A token becomes a generalized conditioned reinforcer when it can be exchanged for a variety of backup reinforcers (Skinner, 1953) and are less sensitive to motivating operations (Ferster & Culbertson, 1982). An early basic experimental study by Wolfe (1936) demonstrated this fact by exposing chimpanzees to various states of food deprivation. Specifically, three chimpanzees were given choices between tokens (some exchangeable for peanuts and some exchangeable for water) while under alternating deprivation conditions (16 hrs) from food or water. All chimpanzees preferred the tokens corresponding to the current deprivation conditions; however, the preference was not exclusive due to the relatively modest deprivation states. Therefore, two additional chimpanzees were given the same choices between tokens under 24-hr deprivation conditions and the researchers allowed them access to the alternate reinforcer prior to each session (1 hr). In this phase of the study, the two chimpanzees under the more stringent conditions preferred the deprivation-specific reinforcer to a higher degree. Thus, this study demonstrated chimpanzees' rate of token-exchange was consistent with the state of deprivation specific to each backup reinforcer and that lessening the motivation for one reinforcer strengthened the motivation for the other.

Another motivating-operation factor studied within basic research (and applicable to practical application) involves the degree to which the reinforcer is available outside of the experimental session. Hursh (1980, 1984) described *closed* economies as those where primary reinforcers are only available through an organism's interaction with the experimental environment, and *open* economies where consumption of the reinforcer is not completely dependent on within-session performance. For example, two studies (Felton & Lyon, 1966;

Catania & Reynolds, 1968) performed experiments in which pigeons were given supplemental (noncontingent) feedings outside of the experimental session that were not the result of the rats' interaction with the experimental environment (open economy). Relative rates of responding were markedly less under these conditions than in two comparable studies in which organisms (rats and monkeys) were only given primary reinforcers in response to their interaction with in-session schedules of reinforcement (Collier, Hirsh, & Hamlin, 1972; Hursh, 1978). Overall, these results suggested that supplemental (extraexperimental) access to reinforcers can have unforeseen effects on motivating operations and subsequent within-session performance.

In applied settings. Motivating operations can be seen as an advantage to practitioners who can regulate the amount of access to a single backup reinforcer. This can be achieved by ensuring that the client does not have access to the preferred backup item outside of the ongoing token economy. For instance, Roane, Call, and Falcomata (2005) demonstrated more responding during closed economies in which participants were only able to obtain reinforcement through interaction with progressive-ratio schedules of reinforcement during session. This is in contrast to open-economy sessions during which participants demonstrated decreased responding while obtaining both within-session reinforcers and supplemental access to reinforcers outside of session.

Given the nature of the primary backup reinforcer, motivating operations may constitute a disadvantage if the client satiates too quickly (e.g., Vollmer & Iwata, 1991). Moher et al. (2008) reported that the effectiveness of conditioned reinforcers (tokens) decreased during periods in which participants were satiated on the backup reinforcer. Rotation and choice across multiple backup reinforcers may guard against motivating operations' unforeseen effects on the reinforcing value of tokens (Sran & Borrero, 2010; Tiger, Hanley, & Hernandez, 2006).

Furthermore, increasing the number of backup reinforcers with which the token is paired (e.g., a choice between two or three backup reinforcers) should result in the maintenance of responding even when clients are satiated on the most preferred backup reinforcer (Moher et al., 2008). Thus, efforts can be made to decrease the potential negative effects of motivating operations (i.e., abolishing operations) by having a menu of options from which a client can “buy” backup reinforcers.

There is, however, a potential trade-off between using a menu of backups that will account for these fluctuations versus the challenge of pairing tokens with many different stimuli. Adding backup stimuli may interact with the original token-backup pairing. Depending on the functioning level of the client and the original number of pairings required, it may present a challenge to add additional backups that require the same intensity of pairing. In response to adding more backup stimuli, if the client begins to respond inconsistently to a previously effective token, it may be the case that the token no longer acts as a discriminative stimulus for access to the original backup (Kelleher & Gollub, 1962). Adding more backup stimuli may have essentially “watered down” the original pairing. Rotating backup reinforcers can be effective in maintaining client progress within a token economy but practitioners should be wary of adding too many backup stimuli without considering the original token-backup pairing.

Barrier: Difficulty Shaping the Exchange Response

Depending on the functioning level of the client, his or her apparent reluctance to exchange the token will likely be solved by ensuring high quality backup reinforcers, sufficient and consistent pairing, and investigating the effects of motivating operations. However, given that most practitioners themselves have had a long history of operating within a token economy (e.g., receiving and cashing paychecks), there may be some inclination to assume the client can

exchange tokens for backups spontaneously. For ease and efficiency of token-backup exchange, there may be some benefit in removing the exchange response altogether. That is, by exchanging the tokens *for* a client who is struggling to learn the exchange response, the reinforcing properties of the token might stay intact with practitioner-mediated token-backup pairing. However, explicit teaching of the token exchange may be necessary and beneficial if the practitioner intends on allowing the client to determine components such as the magnitude and rate of reinforcement (i.e., having control over when to exchange tokens and for which backup).

Basic experimental research. Animal studies often rely on multiple stages to shape the exchange response. After magazine training (where animals are taught to run to the food receptacle and consume primary reinforcers), a shaping procedure is used to reinforce approximations of lever pressing. Experimenters then focus on teaching the animal the token deposit response. For instance, Malagodi (1967a) distributed 80 marbles on the floor of an operant chamber and reinforced successive approximations of rats depositing the marbles into a receptacle. Stimulus control over the response was gained by reinforcing deposit responses on a continuous reinforcement schedule in the presence of a discriminative stimulus (receptacle light and clicker). In the early Wolfe (1936) and Cowles (1937) studies, exchange opportunities were freely available for chimpanzees and depositing of the token was initially modeled by the experimenter. Each token deposited by the chimpanzees was reinforced immediately with food. In another example with pigeons, Foster et al. (2001) began token exchange training with 30 LED lights illuminated on a board. On a VT30-s schedule one light was turned off and the hopper was raised until the pigeon completed 2 s of access to food. Once the pigeons were reliably eating from the food hopper, researchers shaped successive approximations of the pigeon pressing a specified (illuminated) key. After this response was established, the key was

made available (illuminated) at irregular intervals during which if the pigeon pecked the key, it would turn off and the pigeon received access to food. Subsequent to the key turning off, an LED light was turned off (i.e., token removed) until all 30 tokens had been exchanged.

In applied settings. Even when tokens have acquired the properties of effective conditioned reinforcers, not all clients will immediately have mastery over the response chain necessary to physically exchange the token and consume the backup reinforcer (e.g., taking the token, handing it over, waiting, and consuming the backup reinforcer). Numerous empirically validated methods for teaching response chains are possible including graduated guidance (Sisson, Kilwein, & Van Hasselt, 1988), errorless learning (Terrace, 1963), constant-time delay (Wolery et al., 1992), and video modeling (Bellini & Akullian, 2007). Initially, the act of exchanging the token should be the primary task in which the client must engage in order to gain access to the reinforcer. In addition to the client not being fluent with token exchange, it could also be the case that the “demand” of the exchange response is too great. That is, if the demand of the act of exchanging is too great (e.g., walking over to a new area, locating the practitioner, making choices, engaging in a communicative response), clients may save tokens so they can increase the amount of reinforcers they receive per exchange (Cole, 1990; Sousa & Matsuzawa, 2001; Yankelevitz, Bullock, Hackenberg, 2009). For instance, an establishing operation for tokens might be in effect during low-effort tasks if the exchange response is too demanding; at least until a sufficient number of tokens have been accumulated to override the abolishing operation of the exchange response. Thus, practitioners should consider the overall demand of the exchange itself, as it may influence the reinforcing effectiveness of tokens and the backup stimuli.

Part II: Acknowledging and Investigating First- and Second-order Schedules of Reinforcement

Reinforcement schedules do not operate in isolation; instead, one schedule (a first-order schedule) can be a unit of behavior upon which another schedule operates (higher- or second-order schedules). In other words, completion of the first-order schedule (e.g., FR5) is a behavioral unit that is reinforced according to second schedule (e.g., VI25). An oversimplified view of client responding within token economies would include practitioners viewing responding as vulnerable to only the “local” contingencies available through first-order reinforcement schedules. If this were the case, client behavior patterns under token economies would mimic those seen under programs using primary reinforcement as it is well documented that an organism’s responding varies as a function of the primary reinforcement schedule in effect (Ferster & Skinner, 1957; Malagodi, 1967b). Fixed ratio schedules, for instance, produce post-reinforcement pauses (also referred to as “pre-run pauses;” see Schlinger, Derenne, & Baron, 2008), in which responding briefly ceases following reinforcement delivery. This momentary lag in responding is often followed by an increase in response rate until the organism meets the requirement for reinforcement. Conversely, a variable-ratio schedule produces relatively higher and steadier rates of responding (Ferster & Skinner, 1957).

Patterns of behavior within an extended token economy, however, should be considered as unitary responses influenced and reinforced according to two other higher-order schedules. For instance, Kelleher (1958, 1966) described token economies as involving three interconnected schedules of reinforcement and client behavior that is responsive to a token economy will be jointly determined by both the first- and second-order reinforcement schedules (Foster et al., 2001; Webbe & Malagodi, 1978). The three schedules include the following:

- Token-production schedule: the first-order schedule of reinforcement under which the behavior targeted for change will result in tokens (e.g., FR5: organism must emit 5 responses to receive one token).
- Exchange-production schedule: the schedule that determines when the opportunity to exchange tokens for backup reinforcers is available (e.g., FT5: the organism can exchange tokens for backups every 5 min).
- Token-exchange schedule: the rate of exchange or “cost” for backup reinforcers (e.g., FR5: the organism must exchange 5 tokens for a backup reinforcer).

Practitioners might view ongoing patterns of behavior as being reinforced by the local contingencies available through the immediate token-production schedule; however, responding is not under the sole control of any of the three schedules at any given point in time. Given research supporting the separate and combined effects of these schedules (Kelleher, 1966; Hackenberg, 2009), practitioners are advised to take notice of the three specific schedules of reinforcement.

Whereas the token-production schedule refers to the number of target responses that must be emitted by clients to receive a token (Kelleher, 1956), the exchange-production schedule refers to how often a client is given the opportunity to exchange tokens for the backup reinforcers. For example, a client can purchase 5 pieces of candy after accumulating 5 tokens (FR5-exchange-production schedule) or has to wait until the end of every hour when the school store opens (FT60-min). Basic experiments (e.g., Kelleher, 1958; Waddell, Leander, Webbe, & Malagodi, 1972; Webbe & Malagodi, 1978) provide evidence of responding that is sensitive to the exchange-production schedule by manipulating exchange-schedule variables and subsequently producing rates and patterns of behavior typical to higher-order schedules of

primary reinforcement. Webbe and Malagodi (1978) demonstrated the effects of the interaction of the token- and exchange-production schedules in a laboratory setting. Specifically, rats were trained to press a lever that produced tokens on a FR20 token-production schedule. Patterns of responding were examined under two different exchange-production schedules. In one phase, earned tokens could be exchanged for food pellets on an FR6 token-production schedule; in the other phase, a VR6. Under the fixed exchange-production schedule, the pattern of responding contained relatively long and frequent pauses at the initiation of each new token-production sequence. When the exchange-production schedule was variable, patterns of responses contained sharp reductions in the frequency and duration of these initial pauses resulting in an increased overall rate of lever pressing.

The other second-order schedule, the token-exchange schedule, refers to how many tokens are required for a specific backup reinforcer, or the token-specific “unit-price” of the backup reinforcer. For example, access to the computer may require the exchange of 10 tokens, whereas a candy bar might require three. It is not always the case that once an exchange opportunity is earned, the client can exchange only one token for one unit of a backup reinforcer. What is more likely to occur in clinical and school settings is the option for clients to “purchase” from a variety of backup reinforcers, which likely vary in price (or token-exchange schedules). While there is undoubtedly an individual effect of the token-exchange schedule on client responding, later sections in this paper will describe a more refined explanation of the “true” price of backup reinforcers (i.e., “unit price”).

In summary, there is some value in investigating the more immediate effects of the token-production schedule, such that modifications specific to local contingencies can have significant effects on behavior patterns (similar to those under primary reinforcement). Practitioners should

be wary of solely relying on this information; however, because the influence of the other schedules can have effects that are often overlooked. Thus, client responding is influenced by all three schedules and practitioners need to investigate both the “local” contingencies of the token-production schedule as well as the often superseding contingencies of the second-order schedules (the exchange-production and token-exchange schedules).

Barrier: Appropriately Adjusting Local Contingencies of the Token-Production Schedule

Specific to the token-production schedule, practitioners may struggle in deciding whether they want to provide tokens after a fixed or variable number of responses (FR or VR); after a fixed or variable duration the client engages in the behavior (FRD or VRD; Dixon et al., 1998); or after the first response following a fixed or variable amount of time (FI or VI). Most token economies implemented in applied settings run on fixed schedules. There are obvious logistical benefits of using fixed token-production schedules (e.g., ease of implementation, predictability) and FR schedules can often result in high, steady responding and are especially beneficially when teaching new behaviors. One important property of fixed schedules, however, is that they can introduce discriminable periods during which reinforcements would not occur, and responding could appear erratic (e.g., scalloped responding under an FI schedule where client responses would be slow in the beginning of an interval and more rapid just before token delivery). Furthermore, FI schedules can become problematic when practitioners attempt to reinforce longer durations of continuous appropriate behavior such as increasing clients’ time on task.

Basic experimental research. As stated before, the token-production schedule does exert some control over response patterns resembling those obtained under schedules of primary reinforcement (Kelleher, 1956; Malagodi, 1967b). For example, Kelleher (1958), trained

chimpanzees to press a lever that produced poker chips according to an FR30 schedule (i.e., every 30 responses produced one poker chip). Exchange periods were scheduled following an FR50 exchange-production schedule. Corresponding to basic experimental research with simple FR schedules using primary reinforcement, responding occurred at a high-steady rate, with short pauses prior to each ratio run (e.g., break-run patterning). Researchers then increased the token-production schedule to an FR125 while keeping the exchange-production schedule constant. Again, emulating effects seen with FR schedules of primary reinforcement (Felton & Lyon, 1966; Mazur, 1983), overall response rates decreased and pre-ratio pausing increased. Furthermore, by switching the token-reinforcement schedule to variable schedules, under local contingencies of the token-production schedule, one can expect high and steady responding under VR token-production schedules, and slow and steady responding under a VI schedule (Ferster & Skinner, 1957).

In applied settings. During acquisition phases, the practitioner has the option to begin with an FR1 schedule (i.e., continuous reinforcement) so that every occurrence of the new behavior receives reinforcement. After the client's behavior progresses, the practitioner can systematically adjust the token economy to include intermittent schedules (e.g., thinning the schedule to an FR5). Alternatively, moving to a VR token-production schedule will result in maintenance of the replacement behavior over longer periods of time and may guard against post-reinforcement pauses and satiation of backup reinforcers (Skinner, 1953). Under FI schedules, basic research has suggested that practitioners would observe long periods of inactivity with slight increases in responding surrounding the interval's end (scalped responding; Kelleher, 1956). To modify this pattern, practitioners could change to either a VI token-production schedule or a response duration schedule. Under a VI schedule, responding

would be steadier and more moderate due to the unpredictability of the interval's end (Malagodi, 1967c). With an FRD or a VRD schedule, practitioners have the option to only reinforce behaviors that are of a specific duration. An example of effectively using an FRD token-production schedule would involve a client receiving a token after 5 min of continuous appropriate conversation, and not receiving a token if the duration of the conversation were less than 5 min.

Barrier: Adjusting the Token-Production Schedule to Reduce Inappropriate Behavior

Common barriers not only to token economies, but to applied practice as a whole, can include the client engaging in problem behavior. Practitioners most likely are well versed in the use of differential reinforcement procedures to reinforce an appropriate behavior in place of an inappropriate one (DRA or DRI), but they may struggle to enact appropriate measures to respond to inappropriate behavior using the token economy. Modifying the token-production schedule to include response cost procedures (rather than relying on punitive measures) could be an effective addition to target problem behavior.

Basic experimental research. Azrin and Holz (1966) defined response cost as a punishment procedure in which conditioned token reinforcers are removed response contingently. For example, Pietras and Hackenberg (2005) used LED lights as tokens to reinforce pecking in pigeons. Each LED token, when illuminated was worth 2.5 s of food. Use of these lights allowed researchers to easily remove tokens by turning the light off. Key pecking was maintained on two separate schedules and when FR schedules of response-cost were introduced in one of the schedules, response rates under only that schedule decreased. It is interesting to note that in this experiment, response rates for the response-cost schedule were not completely suppressed; rather, only during extinction did response rates decrease to near-zero

levels. In a basic experiment using humans (Weiner, 1962), responses produced brief stimuli (lights) signaling availability of reinforcers (points) according to either VI or FI schedules. By subtracting one point from a counter during response cost conditions, response rates were suppressed and did not recover with continued exposure to the response-cost contingency. Thus, basic experimental research has consistently demonstrated decreased response rates of target behaviors using contingent removal of conditioned reinforcers.

In applied settings. A response cost procedure can be effective within any type of token-production schedule as long as the tokens are acting as conditioned reinforcers. Much of the applied token economy research implementing response cost involves the client being given a number of tokens at the beginning of an interval and losing tokens for each inappropriate response (e.g., Conyers et al., 2004; Iwata & Bailey 1974; McGoey & DuPaul, 2000). If the client has enough tokens at the end of the set interval, he or she can exchange them for a backup reinforcer. Conyers et al. (2004) found that both a response-cost and DRO procedure were effective in reducing problem behavior when implemented in isolation; however, they recommended implementing them together to increase treatment acceptability. Thus, a response cost procedure might be even more valuable of a procedure if implemented in conjunction with a DRO token-production reinforcement schedule. For instance, Repp and Deitz (1974) implemented a 15-min DRO token production schedule in which the participant received tokens for *not* engaging in the target behavior. If the participant did engage in the target behavior, however, he lost all accrued tokens.

A practical limitation of using response cost can involve an instance in which a client loses all tokens before an exchange period and has a long wait before an opportunity to earn them back. This instance might produce a segment of time in which contingencies for

appropriate behavior are vague, perhaps creating an establishing operation for problem behavior. Some other practical limitations are encompassed by the potential negative side effects of punishment procedures in general. Specifically, punishment procedures such as response cost can produce negative side effects that include collateral increases in punishment-elicited aggression, escape behaviors, and emotional reactions (Lerman & Vorndran, 2002). Lastly, using response cost in isolation might be disadvantageous considering exchange opportunities depend on responding (FR or VR exchange-production schedules). That is, if response-cost conditions result in low response rates, infrequent pairings of tokens with backup reinforcers may reduce the reinforcing value of the tokens.

Barrier: Unforeseen Effects of the Exchange-Production Schedule

Although we can modify the token-production schedule to effectively influence client responding, unconventional patterns in client behavior under token economies may in fact be attributed to a practitioner overlooking the higher-order effects of the exchange-production schedule. This may be due to focusing too narrowly on the more local effects of the token-production schedule or other features described in Part I. Practitioners might come across decreased responding even with a saturated token-production schedule. One area that needs to be emphasized in this instance is the influence of the exchange-production schedule.

Basic experimental research. Basic research has suggested that the exchange-production schedule may in fact have greater control over responding patterns than the local contingencies operating within the token-production schedule (Webbe & Malagodi, 1978). Foster et al. (2001) highlighted the relatively greater influence of the exchange-production schedule over token-production schedules by comparing one condition with a VR-token-production schedule and a FR-exchange-production schedule with another condition involving

an FR-token-production schedule and a VR-exchange-production schedule (i.e., FR-token/VR-exchange vs. VR-token/FR-exchange). Because pause durations were longer in the VR token-production schedule (a schedule known to produce relatively pause-free, constant rates of responding) relative to the FR token-production schedule (a schedule known to produce break-run patterns), overall rates of behavior were primarily organized by the exchange-production schedule requirements. Bullock and Hackenberg (2006) extended these results by showing more pronounced effects of the exchange-production schedule when the token-production ratios were higher (i.e., when more responses per token were required). Trials that included lower token-production schedules (e.g., FR2), response rates varied much less with the exchange-production schedule. That is, the schedules that allowed more frequent access to tokens mimicked those of primary reinforcement schedules; whereas, when more responses were required per token (i.e., more effort) the frequency of exchange periods had more influence over responding.

Furthermore, response rates increased and pausing decreased with proximity to exchange periods and food. These findings extend prior research (e.g., Foster et al., 2001; Kelleher, 1966) by showing that the effects of the exchange-production ratio depend on the token-production ratio.

In applied settings. A practical example of the token-production schedule acting as a unitary response, and thereby producing reinforcement according to the exchange-production schedule, would include a client earning tokens for completing an activity of daily living such as folding clothes. For every 2 items the client folds, he will receive a token (an FR2-token-production schedule). Given a relatively frequent exchange period (e.g., FR4-exchange-production schedule), one could expect responding to adhere to patterning seen under simple-schedules with primary reinforcement (i.e., the client would fold clothes rapidly to accumulate more tokens within a given time period). The frequency of exchange periods becomes even

more influential once practitioners thin the token-production schedule. For instance, if the practitioner now requires the client to fold 15 items for one token (FR15-token-production schedule), these clusters of behaviors are now more vulnerable to the effects of the higher-order exchange-production schedule. The practitioner could now allow the client to exchange all the tokens for a backup reinforcer either (a) after folding a fixed number of items (e.g., FR45-exchange-production schedule), or (b) after an average number of items (e.g., VR45). Under the FR45 exchange-production schedule, the client is likely to engage in post-reinforcement pausing with a quick transition to rapid responding until reinforcement is received (i.e., break-run patterning). Under the VR45 exchange-production schedule, however, the client should fold clothing quicker while pausing for shorter time periods (Kelleher, 1966). Thus, even though the token-production schedule is the same in either scenario (FR15), it is the exchange-production schedule that disproportionately controls the overall rate of responding.

Barrier: Overcoming Ratio Strain

It may be fairly obvious that client responding will decrease in response to low reinforcement frequency. What might not be obvious is that ratio strain can unexpectedly occur through the interaction of the token- and exchange-production schedule. Sifting through why ratio strain is occurring and which schedule is influencing responding the most can be a complex task. Long pauses in ratio performance could occur with too high of requirements within the token-production schedule, or too much time between exchange opportunities (Bullock & Hackenberg, 2006).

Basic experimental research. Within token economies, ratio strain occurs with abrupt increases in ratio requirements, resulting in decreases in behavior similar to that seen during extinction (Ferster & Skinner, 1957). Specific to token-production, organisms whose schedules

of reinforcement are thinned too quickly are required to engage in more responses (or a greater wait time) before reinforcement is earned. Hodos and Kalman (1963) demonstrated decreased responding in rats was a function of increased ratio requirements through use of progressive-ratio schedules in which the organism is required to emit a systematically increasing number of responses for each successive reinforcement. Ratio strain can also occur through the interaction of the token- and exchange-production schedules whereby organisms might earn tokens at an appropriate rate, yet would show decreased responding if the requirement for exchange-production was too stringent (i.e., exchange opportunities are not often enough). For instance, Bullock and Hackenberg (2006) demonstrated ratio strain in pigeons by showing an inverse relation between responding and the token-production ratio (i.e., high response requirements decreased responding due mainly to long pauses and low response rates in early segments). Researchers also demonstrated, however, an inverse relation between response rate and exchange-production ratios when token-production ratios were kept constant.

In applied settings. This instance could occur within applied environments when practitioners delay opportunities to exchange tokens until the end of the day (FT- or VT-exchange-production schedules) or overlook the prospect of restoring the exchange-production schedules back to the ratio in which responding was adequate. For example, if a client receives tokens on an FR5-token-production schedule, and can exchange tokens after accumulating an average of 5 tokens (VR5-exchange-production schedule), all other things equal, the practitioner can expect relatively rapid responding with short post-reinforcement pauses. If, however, the practitioner abruptly requires the client to either (a) engage in 50 responses for 1 token (FR-50-token-production schedule), (b) exchange tokens only after accumulating an average of 100 tokens (VR100-exchange-production schedule), or (c) both, one would expect decreased

responding due to ratio strain, and potentially complete extinction of responding before reinforcement at the new schedule can occur. Interestingly, Hodos and Kalman (1963) demonstrated when large ratio requirements were used the number of responses increased as a function of the volume of the reinforcer (i.e., reinforcer magnitude) and decreased during small ratio requirements due to satiation. Thus, ratio strain can be avoided by increasing response requirements gradually, temporarily reducing ratio requirements, *or* by increasing backup magnitude or quality (e.g., Roane, Falcomata, & Fisher, 2007). In practice, designing token economies that avoid *both* satiation and ratio strain can become quite complex.

Barrier: Adjusting Prices of Backup Reinforcers

Token economies most often employ use of generalized reinforcers, where clients can exchange tokens for a variety of backup reinforcers listed at different “prices.” Prices (token-exchange schedules) for backup reinforcers often may be chosen arbitrarily or might even be based on the actual retail price of the item. Price, however, can represent more than simply the token-exchange schedule, or the amount of tokens the practitioner would require the client to exchange for specific backup reinforcers. Therefore, practitioners likely need to consider the effects of the overall cost-benefit tradeoff using all three reinforcement schedules.

Basic experimental research. Unfortunately, basic research does not provide extensive information about the specific influence of token-exchange schedules (Hackenberg, 2009). The only published study on this matter suggested that the token-exchange schedule’s influence on responding is similar to the exchange-production schedule in that client responding within token economies is under the control of both first- and second-order schedules (Malagodi, Webbe, & Waddell, 1975). Generally, Malagodi et al. (1975) demonstrated decreased responding in rats’ lever pressing given increased token-exchange schedules: the more demanding the token-

exchange schedule, the longer post-reinforcement pausing. Basic research on “unit price,” or the ratio of responses to every unit of reinforcer (Hursh, 1978), includes a more sophisticated description of price that details the characteristics of cost-benefit tradeoffs (e.g., Delmendo, Borrero, Beauchamp, & Francisco, 2009; Foster & Hackenberg, 2004). That is, “price” does not necessarily refer just to the actual price of the backup reinforcer (token-exchange schedule); rather, the interaction of the token-exchange schedule with the “costs” of producing tokens (the token-production schedule) and the “costs” of producing exchange opportunities (the exchange-production schedule; Bullock & Hackenberg, 2006). Basic research on unit price (Bullock & Hackenberg, 2006; Delmendo et al. 2009) demonstrates that decreases in response rates are associated with increases in unit price (more stringent token-production, exchange-production, and token-exchange requirements).

In applied settings. Malagodi et al. (1975) suggested that practitioners can expect decreased overall responding if the labeled “price” of the backup reinforcer is relatively too high (i.e., too demanding of a token-exchange schedule). In applied settings; however, the actual “price” of the backup reinforcer for the client is better determined by the combined effects of how often responses result in tokens, how often the client can exchange tokens, *and* the number of tokens required to purchase specific reinforcers (Bullock & Hackenberg, 2006; Malagodi et al., 1975). In this sense, Hackenberg (2009) likened the token-production schedule to a worker’s wage, the exchange-production schedule to the effort required to purchase the item (e.g., driving to the store, getting cash out of the bank), and the token-exchange schedule to the number listed on the price tag. The basic premise of “unit price” can be applied to practical contexts by using reinforcer assessments to identify a hierarchy of backup reinforcers. Delmendo et al. (2009) suggested this information could be used to differentially program contingencies for clients’ task

completion based on the subjective effort associated with the task. Within token economies, those tasks that require more client effort can be associated with larger payoffs (i.e., access to more preferred rewards). Conversely, reinforcers that are less preferred may be more suitable for maintaining less effortful responses. An example of this situation could involve a client being able to use tokens to purchase high quality rewards that are not available at other times after a period of effortful tasks (e.g., difficult homework). Less preferred reinforcers, therefore, would be available for the periods that involve less effortful tasks (e.g., sitting at the table while eating).

Barrier: Lack of “Self-Control” and Delay of Reinforcement

Practitioners might encounter clients that are dependent on the immediacy of reinforcers or require frequent access to high-quality reinforcers. Although this is acceptable in the early stages of the token economy, or when a client is learning a new behavior, extended token economies can become quite cumbersome if this level of reinforcement is required. Practitioners most likely would prefer to teach clients “self-control” in which they are able to wait for reinforcers and eventually learn to choose larger (higher quality) delayed reinforcers rather than demand immediate reinforcers (e.g., “impulsiveness”). Fortunately, token economies are especially equipped to teach self-control by employing delayed reinforcement procedures and systematically fading reinforcement.

Basic experimental research. Overall, basic research has suggested that introduction of delay increases resistance to extinction and tokens can maintain behavior despite significant delays to exchange periods (Lattal, 2010; Skinner, 1953; Wolfe, 1936). Tolerance of delays (i.e., self-control) can be developed by increasing delays gradually. For instance, Jackson and Hackenberg (1996) gradually increased the period of time in between token-production and exchange periods by making exchange periods more intermittent across phases until a single

exchange period occurred at the end of an experimental session. Thus, tokens (i.e., LED lights) were initially established as conditioned reinforcers by frequently pairing token presentation with food delivery. Subsequently, by gradually increasing the response requirement to access token exchange, researchers established a rich history of correspondence between tokens and food and minimized the response-weakening effects of potential ratio strain. Interestingly, this study also demonstrated that delays to token reinforcers were less critical than delays to exchange opportunities, suggesting a stronger influence of the token-exchange schedule than the token-production schedule.

Another approach to minimize ratio-strain in this instance is to increase reinforcer magnitude as a reinforcement schedule is thinned. A systematic approach used to study the effects of thinning a token reinforcement schedule includes progressive-ratio schedules. This type of approach involves increasing the response requirement systematically during the course of a session until an organism reaches its “breaking point.” For example, two studies involving rats (Baron, Mikorski, & Schlund, 1992; Hodos & Kalman, 1963) demonstrated that with increasing schedule requirements, post-reinforcement pausing was inversely related to reinforcer magnitude (or the concentration level of sweetened condensed milk). Specific to magnitude (or the number, intensity, and duration of reinforcement), basic research findings have suggested that treatment will be more successful if the magnitude of the reinforcer is increased as the schedule is thinned (Baron & Derenne, 2000). In addition to the findings mentioned in the paragraph above, Jackson and Hackenberg (1996) also demonstrated an ability to maintain pigeons’ preferences for large reinforcers as the delay to a token exchange period was increased. This suggests that in order to establish self-control (i.e., “waiting” for reinforcers of larger magnitude) one would gradually increase delays equally for both small and larger magnitude

reinforcers and then gradually decrease the delay for the small reinforcer while holding the delay for the larger reinforcer constant (e.g., Stromer, McComas, & Rehfeldt, 2000).

In applied settings. Practitioners should consider a number of factors when teaching the client to “wait” for reinforcers (tokens or backups) within a token economy. One primary factor is the client’s ability to tolerate delays or an assessment of “temporal discounting” (Madden, Begotka, Raiff, & Kastern, 2003). Specific to token economies, temporal discounting would refer to both the tokens and the backup reinforcers losing their subjective value as the delay to their receipt increases (Reed & Martens, 2011). Thus, a client with impulsive tendencies would respond less readily to longer delays to reinforcement than a client who exhibits more self-control (Rachlin & Green, 1972). Practitioners can gradually adjust the delay period within both the token- and the exchange-production schedules in three ways: (a) increasing the amount of time between the response and token delivery, (b) increasing the amount of responses required for tokens (e.g., moving from an FR1 to a FR5 token-production schedule), or (c) increasing the amount of time between token accumulation and token exchange (e.g., moving from an FT30- to FT60-min-exchange-production schedule).

As stated above, one approach to introducing delay, and self-control, into token economies is to manipulate the magnitude of backups to offset reinforcer immediacy (Baron & Derronne, 2000). Multiple studies demonstrate the importance of gradually increasing the delay to a reinforcer of larger magnitude (Neef, Bicard, & Endo, 2001; Ragotzy, Blakely, & Poling, 1988; Schweitzer & Sulzer-Azaroff, 1988). In a study by Dixon et al. (1998), for example, researchers initially made both smaller and larger reinforcers immediately available to participants. They established self-control by progressively increasing the schedule associated with only the larger reinforcer. This study used differing *amounts* of reinforcers (e.g., minutes of

attention, small/large cup of soda); however, practitioners may want to foster development of self-control using other reinforcer dimensions such as having clients relinquish immediate access to smaller reinforcers in favor of delayed reinforcers of higher quality.

During increasing schedule requirements in applied settings (i.e., thinning), Roane, Lerman, & Vorndran (2001) demonstrated how participants progressively approached their most preferred stimuli (or those of higher quality) under increasing schedule requirements. This suggests that practitioners might have success with backup reinforcers with uncertain preference strength, but only consistently with low-effort tasks (or low schedule requirements). For high-effort tasks (or high schedule requirements), however, opportunity to exchange tokens for highly preferred backups is likely more influential. Thus, while thinning a token economy (i.e., increasing schedule requirements) practitioners should consider overall response requirements (and task difficulty) when adjusting either the token- or exchange-production schedule. (e.g., giving tokens for more difficult tasks or longer waits for exchange opportunities; or providing higher magnitudes of backup reinforcers for fewer tokens). For example, in a token economy where the client can exchange tokens for “breaks,” and the practitioner is thinning reinforcement by increasing the amount of time between token accumulation and token exchange, he might increase the magnitude of the backup reinforcers by allowing the client to exchange tokens for longer breaks. Alternatively, if the practitioner thins reinforcement by increasing the amount of time between responses and token delivery, he might present two tokens after the delay. The relationship between magnitude of reinforcement and thinning procedures is often complex; yet, practitioners can use knowledge of this relationship to prevent ratio strain and maintain responding within a long-term token economy (Fisher & Mazur, 1997; Roane et al., 2001).

Conclusion

As robust as the literature surrounding token economies is, practitioners may not make regular contact with the basic research that has defined many of the behavior analytic practices that we use today (Carr & Briggs, 2010). However, when faced with practical problems or barriers to client success, it is likely beneficial for practitioners to revisit this literature and examine these underlying principles. Token economies are one of the most widely used interventions to promote behavior change, and this procedure has evolved to be effective across many settings, behaviors, and individuals. Due to this widespread use, casual implementation of the token economy might result in inconsistencies in client responding as well as an overall skepticism in the procedure itself. This is especially true if the practitioner responsible for training others in the implementation of the token economy has a crude understanding of how and why certain procedures function the way that they do. It is true that typical and even substandard implementation of token economies can have positive effects on behavior (e.g., Plavnick, Ferreri, & Maupin, 2010); however, many practitioners are called upon to consult for atypical cases. Atypical cases likely require a deeper understanding of the underlying mechanisms actuating the seemingly everyday practices that we use. The purpose of this paper was to highlight these underlying mechanisms and translate the often complex implications of basic research. Hopefully, practitioners can use this information to improve their own practice as well as their confidence in disseminating the evidence-based reasons as to why token economies have been so effective across so many contexts for so many years.

References

- Ayllon, T., & Azrin, N. H. (1968). *The token economy: A motivational system for therapy and rehabilitation*. New York: Appleton-Century-Crofts.
- Azrin, N. H., & Holz, W. C. (1966). Punishment. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp.380-447). New York: Appleton-Century-Crofts.
- Bailey, J. R., Gross, A. M., & Cotton, C. R. (2011). Challenges associated with establishing a token economy in a residential care facility. *Clinical Case Studies, 10*(4), 278-290.
- Baron, A., & Derenne, A. (2000). Progressive-ratio schedules: Effects of later schedule requirements on earlier performances. *Journal of the Experimental Analysis of Behavior, 73*, 291-304.
- Baron, A., Mikorski, J., & Schlund, M. (1992). Reinforcement magnitude and pausing on progressive-ratio schedules. *Journal of the Experimental Analysis of Behavior, 58*(2), 377-388.
- Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*, 231-242.
- Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73*(3), 264-287.
- Bullock, C. E., & Hackenberg, T. D. (2006). Second-order schedules of token reinforcement with pigeons: Implications for unit price. *Journal of the Experimental Analysis of Behavior, 85*, 95-106.
- Carr, J. E., & Briggs, A. M. (2010). Strategies for making regular contact with the scholarly literature. *Behavior Analysis in Practice, 3* (2), 13-18.

- Catania, A. C., & Reynolds, G. A. (1968). Quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *11*, 327-383.
- Collier, G. H., Hirsch, E., & Hamlin, P. H. (1972). The ecological determinants of reinforcement in the rat. *Physiology and Behavior*, *9*, 705-716.
- Christensen, L., Young, K. R., & Marchant, M. (2004). The effects of a peer-mediated positive behavior support program on socially appropriate classroom behavior. *Education and Treatment of Children*, *27*, 199-234.
- Christophersen, E. R., Arnold, C. M., Hill, D. W., & Quilitch, H. R. (1972). The home point system: Token reinforcement procedures for application by parents of children with behavior problems. *Journal of Applied Behavior Analysis*, *5*, 485-497.
- Cole, M. R. (1990). Operant hoarding: A new paradigm for the study of self-control. *Journal of the Experimental Analysis of Behavior*, *53*, 247-261.
- Conyers, C., Miltenberger, R., Maki, A., Barenz, R., Jurgens, M., Sailer, A.,... & Kopp, B. (2004). A comparison of response cost and differential reinforcement of other behavior to reduce disruptive behavior in a preschool classroom. *Journal of Applied Behavior Analysis*, *37*, 411-415.
- Cowles, J. T. (1937). Food-tokens as incentives for learning by chimpanzees. *Comparative Psychological Monographs*, *12*, 1-96.
- DeLeon, I. G., Anders, B. M., Rodriguez-Catter, V., & Neidert, P. L. (2000). The effects of noncontingent access to single- versus multiple-stimulus sets on self-injurious behavior. *Journal of Applied Behavior Analysis*, *33*(4), 623-626.

- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of multiple-stimulus presentation format for assessing reinforcer preference. *Journal of Applied Behavior Analysis, 29*, 519-532.
- Delmendo, X., Borrero, J. C., Beauchamp, K. L., & Francisco, M. T. (2009). Consumption and response output as a function of unit price: Manipulation of cost and benefit components. *Journal of Applied Behavior Analysis, 42*, 609-625.
- Dixon, M. R., Hayes, L. J., Binder, L. M., Manthey, S., Sigman, C., & Zdanowski, D. M. (1998). Using a self-control training procedure to increase appropriate behavior. *Journal of Applied Behavior Analysis, 31*, 203-210.
- Drabman, R. S., & Tucker, R. D. (1974). Why token economies fail. *Journal of School Psychology, 12* (3), 178-188.
- Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp.313-339). Englewood Cliffs, NJ: Prentice-Hall.
- Felton, M., & Lyon, D. (1966). The post-reinforcement pause. *Journal of the Experimental Analysis of Behavior, 9*, 131-134.
- Ferster, C. B., & Culbertson, S. A. (1982). *Behavior Principles* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton Century-Crofts.
- Fisher, W. W., & Mazur, J. E. (1997). Basic and applied research on choice responding. *Journal of Applied Behavior Analysis, 30*(3), 387-410.
- Foster, T. A., & Hackenberg, T. D. (2004). Unit price and choice in a token-reinforcement context. *Journal of the Experimental Analysis of Behavior, 81*(1), 5-25.

- Foster, T. A., Hackenberg, T. D., & Vaidya, M. (2001). Second-order schedules of token reinforcement with pigeons: Effects of fixed- and variable-rate exchange schedules. *Journal of Experimental Analysis of Behavior*, 76, 159-178.
- Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of Experimental Analysis of Behavior*, 91, 257-286.
- Hanley, G. P., Iwata, B. A., & Roscoe, E. M. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis*, 39(2), 189-202.
- Hearst, E. (1975). Pavlovian conditioning and directed movements. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9; pp. 215-262). New York, NY: Academic Press.
- Hernstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243-266.
- Higgins, J. W., Williams, R. L., & McLaughlin, T. F. (2001). The effects of a token economy employing instructional consequences for a third-grade student with learning disabilities: A data-based case study. *Education and Treatment*, 24, 99-106.
- Hodos, W., & Kalman, G. (1963). Effects of increment size and reinforcer volume on progressive ratio performance. *Journal of the Experimental Analysis of Behavior*, 6(3), 387-392.
- Holland, V., & Davison, M. C. (1971). Preference for qualitatively different reinforcers. *Journal of the Experimental Analysis of Behavior*, 16, 375-380.
- Hursh S. R. (1978). The economics of daily consumption controlling food- and water-reinforced responding. *Journal of the Experimental Analysis of Behavior*, 29, 475-491.

- Hursh, S. R. (1980). Economic concepts of the analysis of behavior. *Journal of the Experimental Analysis of Behavior*, 34, 219-238.
- Hursh, S. R. (1984). Behavioral economics. *Journal of the Experimental Analysis of Behavior*, 42, 435-452.
- Iwata, B. A., & Bailey, J. S. (1974). Reward versus cost token systems: An analysis of the effects on students and teacher. *Journal of Applied Behavior Analysis*, 7(4), 567-576.
- Jackson, K., & Hackenberg, T. D. (1996). Token reinforcement, choice, and self-control in pigeons. *Journal of the Experimental Analysis of Behavior*, 66(1), 29-49.
- Kazdin, A. E. (1977). *The token economy: A review and evaluation*. New York: Plenum.
- 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. (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.
- Kelleher, R. T. (1966). Conditioned reinforcement in second-order schedules. *Journal of the Experimental Analysis of Behavior*, 9, 475-485.
- Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of behavior*, 5(4), 543-597.

- Laraway, S., Snyckerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*(3), 407-414.
- Lattal, K. A. (2010). Delayed reinforcement of operant behavior. *Journal of the Experimental Analysis of Behavior, 93*(1), 129-139.
- LeBlanc, L. A., Hagopian, L. P., & Maglieri, K. A. (2000). Use of a token economy to eliminate excessive inappropriate social behavior in an adult with developmental disabilities. *Behavioral Interventions, 15*(2), 135-143.
- Lerman, D. C., & Vorndran, C. M. (2002). On the status of knowledge for using punishment: Implications for treating behavior disorders. *Journal of Applied Behavior Analysis, 35*(4), 431-464.
- Mackintosh, N. J. (1974). *The psychology of animal learning*. London: Academic Press.
- Mace, F. C., & Critchfield, T. S. (2010). Translational research in behavior analysis: Historical traditions and imperative for the future. *Journal of the Experimental Analysis of Behavior, 93*(3), 293-312.
- Mace, F. C., & Roberts, M. L. (1993). Factors affecting selection of behavioral interventions. In J. Reichle and D. P. Wacker (Eds.). *Communicative alternatives to challenging behavior* (pp. 113-134). Baltimore, MD: Brookes.
- Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology, 11*(2), 139-145.
- Malagodi, E. F. (1967a). Acquisition of the token-reward habit in the rat. *Psychological Reports, 20*, 1335-1342.

- Malagodi, E. F. (1967b). Fixed-ratio schedules of token reinforcement. *Psychonomic Science*, 8, 469-470.
- Malagodi, E. F. (1967c). Variable-interval schedules of token reinforcement. *Psychonomic Science*, 8, 471-472.
- Malagodi, E. F., Webbe, F. M., & Waddell, T. R. (1975). Second-order schedules of token reinforcement: Effects of varying the schedule of food presentation. *Journal of the Experimental Analysis of Behavior*, 24, 173-181.
- Matson, J. L., & Boisjoli, J. A. (2009). The token economy for children with intellectual disability and/or autism: A review. *Research in Developmental Disabilities*, 30(2), 240-248.
- Matthews, L. R., & Temple, W. (1979). Concurrent schedule of food preference in cows. *Journal of the Experimental Analysis of Behavior*, 32, 245-254.
- Mazur, J. E. (1983). Steady-state performance on fixed-, mixed-, and random-ratio schedules. *Journal of the Experimental Analysis of Behavior*, 39(2), 293-307.
- McGoey, K. E., & DuPaul, G. J. (2000). Token reinforcement and response cost procedures: Reducing the disruptive behavior of preschool children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly*, 15(3), 330-343.
- McIlvane W. J. (2009). Translational behavior analysis: From laboratory science in stimulus control to intervention with persons with neurodevelopmental disabilities. *The Behavior Analyst*, 32, 269–280.
- McLaughlin, T. F., & Malaby, J. (1972). Intrinsic reinforcers in a classroom token economy. *Journal of Applied Behavior Analysis*, 5(3), 263-270.
- Michael, J. (1993). Establishing operations. *Behavior Analyst*, 16, 191-206.

- Miller, H. L. (1976). Matching-based hedonic scaling in the pigeon. *Journal of the Experimental Analysis of Behavior*, 26, 335-347.
- Miltenberger, R. (2008). *Behavior modification*. Belmont, CA: Wadsworth Publishing.
- Moher, C. A., Gould, D. D., Hegg, E., & Mahoney, A. M. (2008). Non-generalized and generalized conditioned reinforcers: Establishment and validation. *Behavioral Interventions*, 23(1), 13-38.
- Neef, N. A., Bicard, D. F., & Endo, S. (2001). Assessment of impulsivity and the development of self-control in students with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, 34(4), 397-408.
- Neef, N. A., Mace, F. C., Shea, M. C., & Shade, D. (1992). Effects of reinforcer rate and reinforcer quality on time allocation: Extensions of matching theory to educational settings. *Journal of Applied Behavior Analysis*, 25, 691-699.
- 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.
- O'Leary, K., & Drabman, R. (1971). Token reinforcement programs in the classroom: A review. *Psychological Bulletin*, 75(6), 379-398.
- Pietras, C. J., & Hackenberg, T. D. (2005). Response-cost punishment via token loss with pigeons. *Behavioral Processes*, 69(3), 343-356.
- Plavnick, J. B., Ferreri, S. J., & Maupin, A. N. (2010). The effects of self-monitoring on the procedural integrity of a behavioral intervention for young children with developmental disabilities. *Journal of Applied Behavior Analysis*, 43(2), 315-320.

- Rachlin, H., & Green L. (1972). Commitment, choice, and self-control. *Journal of the Experimental Analysis of Behavior*, 17, 15-22.
- Ragotzy, S. P., Blakely, E., & Poling, A. (1988). Self-control in mentally retarded adolescent: Choice as a function of amount and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 49(2), 191-199.
- Reed, D. D., & Martens, B. K. (2011). Temporal discounting predicts student responsiveness to exchange delays in a classroom token system. *Journal of Applied Behavior Analysis*, 44(1), 1-18.
- Repp, A. C., & Deitz, S. M. (1974). Reducing aggressive and self-injurious behavior of institutionalized retarded children through reinforcement of other behaviors. *Journal of Applied Behavior Analysis*, 7(2), 313-325.
- Roane, H. S., Call, N. A., & Falcomata, T. S. (2005). A preliminary analysis of adaptive responding under open and closed economies. *Journal of Applied Behavior Analysis*, 38(3), 335-348.
- Roane, H. S., Falcomata, T. S., & Fisher, W. W. (2007). Applying the behavioral economics principle of unit price to DRO schedule thinning. *Journal of Applied Behavior Analysis*, 40(3), 529-534.
- 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(4), 605-620.
- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, 93, 269-289

- Schlinger, H. D., Derenne, A., & Baron, A. (2008). What 50 years of research tell us about pausing under ratio schedules of reinforcement. *The Behavior Analyst, 31*(1), 39.
- Schweitzer, J. B., & Sulzer-Azaroff, B. (1988). Self-control: Teaching tolerance for delay in impulsive children. *Journal of the Experimental Analysis of Behavior, 50*(2), 173-186.
- Sisson, L. A., Kilwein, M. L., & Van Hasselt, V. B. (1988). A graduated guidance procedure for teaching self-dressing skills to multihandicapped children. *Research in Developmental Disabilities, 9*(4), 419-432.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.
- Sousa, C., & Matsuzawa, T. (2001). The use of tokens as rewards and tools by chimpanzees (Pan troglodytes). *Animal Cognition, 4*, 213-221.
- Sran, S. K., & Borrero, J. C. (2010). Assessing the value of choice in a token system. *Journal of Applied Behavior Analysis, 43*(3), 553-557.
- Stromer, R., McComas, J. J., & Rehfeldt, R. A. (2000). Designing interventions that include delayed reinforcement: Implications of recent laboratory research. *Journal of Applied Behavior Analysis, 33*(3), 359-371.
- 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), 1-16.
- Terrace, H. S. (1963). Discrimination learning with and without "errors". *Journal of the Experimental Analysis of Behavior, 6*, 1-27.
- Vollmer, T. R., & Iwata, B. A. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis, 24*(2), 279-291.

- Waddell, T. R., Leander, J. D., Webbe, F. M., & Malagodi, E. F. (1972). Schedule interactions in second-order fixed-interval (fixed-ratio) schedules of token reinforcement. *Learning and Motivation, 3*(1), 91-100.
- Webbe, F. W., & Malagodi, E. F. (1978). Second-order schedules of token reinforcement: Comparisons on performance under fixed-ratio and variable-ratio exchange schedules. *Journal of Experimental Analysis of Behavior, 30*, 219-224.
- Williams, B. A. (1994a). Conditioned reinforcement: Experimental and theoretical issues. *The Behavior Analyst, 17*, 261-285.
- Williams, B. A. (1994a). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review, 1*, 457-475.
- Williams, B. A., & Dunn, R. (1991). Preference for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 55*, 37-46.
- Weiner, H. (1962). Some effects of response cost upon human operant behavior. *Journal of the Experimental Analysis of Behavior, 5*(2), 201-208.
- Wolery, M., Holcombe, A., Cybriwsky, C., Doyle, P. M., Schuster, J. W., Ault, M. J., & Gast, D. L. (1992). Constant time delay with discrete responses: A review of effectiveness and demographic, procedural, and methodological parameters. *Research in Developmental Disabilities, 13*(3), 239-266.
- Wolfe, J. B. (1936). Effectiveness of token-rewards for chimpanzees. *Comparative Psychology Monographs, 12*, 1-72.
- Yankelevitz, R. L., Bullock, C. E., & Hackenberg, T. D. (2008). Reinforcer accumulation in a token reinforcement context. *Journal of the Experimental Analysis of Behavior, 90*, 283-299.

Zimmerman, J., Hanford, P. V., & Brown, W. (1967). Effects of conditioned reinforcement frequency in an intermittent free-feeding situation. *Journal of the Experimental Analysis of Behavior*, 10(4), 331-340.

CHAPTER 3

DECREASING TRANSITION TIMES AND INCREASING ON-TASK BEHAVIOR IN ELEMENTARY SCHOOL CLASSROOMS: AN EVALUATION OF THE *KEEP BUSY AND CARRY ON* SYSTEM¹

¹Hine, J. F., Foster, T. E., & Ardoin, S. P. Submitted to *Journal of Applied Behavior Analysis*, 05/30/2014.

Abstract

This study evaluated the effects of a computer-assisted intervention developed to incorporate and automate multiple components previously found to be effective in reducing transition times in classrooms. To promote efficient transitions between classroom activities and increase instructional time, this intervention incorporated explicit timing procedures, transition warnings, interdependent group contingencies, and automated cues to stay on task. Further, the current study examined the effects of the intervention on multiple measures including transition times between classroom activities, students' on-task behavior, and teachers' use of prompts and praise statements. Through a withdrawal (ABAB) design across 2 classrooms (2 participants in each) results suggested implementation of the intervention decreased the amount of time participants spent transitioning across activities during centers-based activities, increased students' on-task behavior, decreased the number of prompts teachers used to cue students to transition and stay on task, and conversely increased the number of praise statements made by teachers. Lastly, this study explored how automated intervention components through use of computer-assisted instruction affected the need for extensive teacher training and overall procedural fidelity.

INDEX WORDS: Transitions, On-task behavior, Computer-assisted instruction, Group contingency, Explicit-timing procedures, Teacher prompts

Introduction

Past school-based research defines a transition as the time during which a student is stopping one academic-related activity and attempting to begin another (Arlin, 1979; Brophy, 1988). Throughout a given school day, students transition between classrooms, subjects, teachers, and school areas (e.g., activity centers, playground, classroom, and cafeteria). Research has suggested that use of effective transitions increases student independence, decreases disruptive behaviors, and maximizes instructional time (Cameron, Connor, & Morrison, 2005; Coddling & Smyth, 2008). Gettinger (1995) noted a strong relation between increased academic instructional time and student achievement; thus, mismanagement of transition time not only leads to fewer opportunities for students to engage in (and respond to) academic inquiries, but it also may compromise academic outcomes. With students in elementary school classrooms often spending up to one half of the school day engaged in tasks not related to academic instruction (e.g., Hollowood, Salisbury, Rainforth, & Palombaro, 1995), parents, educators, and students alike have reason to be concerned that too much of the school day is wasted transitioning across or dawdling between activities.

Appropriate transition-related student behaviors include physically moving from one area to another, putting away and gathering materials in a timely manner, and either quietly waiting for instruction from the teacher or initiating an independent assignment. During problematic transitions, students may appear to be unmotivated (i.e., motivational/performance deficit; Noell, Roane, VanDerHeyden, Whitmarsh, & Gatti, 2000) and to prolong transitions by ignoring directions, being unorganized, delaying their responses to prompts, or engaging in problem behavior (Sainato, 1990). Subsequently, teachers may have to spend time either prompting them through correct behaviors or reprimanding inappropriate behaviors. Interestingly, what often

may be perceived as a failure on the student's part may in fact be exacerbated by a failure on the teacher's part. That is, teachers may often engage in behaviors they perceive to be appropriate in addressing problematic transitions (e.g., reprimanding, waiting; Fudge et al., 2008), but these behaviors might actually prolong transitions (Sainato, Strain, Lefebvre, & Rapp, 1987).

Generally speaking, there is a continual need for easily implemented evidence-based interventions targeting transitions. The purpose of this study was to evaluate an innovative and automated system designed to decrease the amount of time that students spend transitioning between classroom activities as well as the need for teachers to repeatedly prompt and reprimand students who are transitioning inefficiently. Ideally, this system was expected to decrease the amount of time students spent transitioning, thus increasing the amount of time that students were academically engaged.

Common Practices Used to Decrease Transition Time

Educators have several options for improving transitions and increasing instructional time (e.g., Christ & Christ, 2006; Coddling & Smyth, 2008; McCord, Thomson, & Iwata, 2001). One antecedent-based procedure often employed by teachers is the use of a transition signal (e.g., verbal warnings, flicking lights, repeating rhythmic statements). Although transition signals are commonly used within classroom settings, there are few empirical studies supporting their use. In a thorough review of the literature, only one study was identified that demonstrated successful use of a transition signal in isolation from other intervention components. Zeece and Crase (1982) examined the effects of a verbal warning prior to transitions on the amount of time required for 40 preschool children to pick up their toys and transition to a new activity area. Results indicated that the signal alone facilitated quicker transitions.

Conversely, Cote, Thompson, and McKerchar (2005) examined the use of transition signals in a preschool classroom and found that signals alone did not increase transition-related compliance for 3 typically developing toddlers. In fact, participants' compliance did not increase until implementation of escape extinction. Cote et al. suggested that antecedent interventions such as transition signals are likely more effective when used to enhance the effects of other interventions. It is possible that the popularity of transition signals is due to ease of implementation; however, transition signals might have the greatest impact in reducing transition times if added as a secondary component to other evidence-based strategies. For example, Ferguson, Ashbaugh, O'Reilly, and McLaughlin (2004) demonstrated reduction in transition times in a self-contained kindergarten classroom using a multi-component intervention package. Intervention components included a transition signal (ringing of a bell), teacher modeling, contingent reinforcement, and using physical assistance to prompt noncompliant students. Overall, studies such as Ferguson et al. have suggested that classroom management procedures would likely be most effective if educators employed multiple transition time-reducing strategies. In addition to transition signals, another option for improving transitions involves explicit timing procedures (Campbell & Skinner, 2004).

Research has suggested that overtly timing student behaviors and providing students with performance feedback as to the duration of those behaviors can improve speed of responding across various tasks (Rhymer et al., 2002). Although most studies on explicit timing have focused on students' completion of academic work (e.g., math worksheets; Evans, Skinner, Henington, Sims, & McDaniel, 2002), researchers have applied similar procedures to transitions. For instance, Wurtele and Drabman (1984) decreased clean-up time in a kindergarten classroom using a "Beat-the-Buzzer" game, which involved students cleaning up their play area within a

pre-specified amount of time. Using a withdrawal design, the study demonstrated effective reduction of clean-up times without the teacher providing programmed consequences for participants' behavior.

Yarbrough, Skinner, Lee, and Lemmons (2004) also implemented an explicit timing procedure in a study targeting transition behavior in a sixth-grade classroom. As a central component of the intervention (the Timely Transitions Game; Campbell & Skinner, 2004), the classroom teacher used a stopwatch to time transitions, and reported the amount of time required to transition to the class. The teacher then wrote the time on the board (publicly posted feedback) and randomly selected a card from a box of cards labeled with specified times. If the class "beat" this time (i.e., transitioned in a shorter interval) collectively, they earned a token (e.g., letters of a word that would eventually spell out a class-wide reward). In addition to demonstrating the effects of explicit timing, this study illustrated the potential benefits of combining multiple intervention components, such as contingency management and publicly posted feedback. Yarbrough et al. recognized the Timely Transitions Game as an effective and efficient strategy for managing group behavior, as well as a medium for incorporating multiple intervention components. Additionally, Yarbrough et al. noted that the teacher reported that the intervention "required little time and effort and that she no longer had to prompt, pester, or yell at the children to get them settled after lunch" (p. 103). Unfortunately, no data were taken on the number of prompts and reprimands that the teacher delivered across conditions. Furthermore, the teacher received continual assistance from graduate assistants, and the amount of training and feedback required was unclear. Unfortunately, multifaceted interventions like the Timely Transitions Game can be quite complex and, as a result of their complexity, may require considerable teacher support and training.

Procedural Fidelity and Computer-Assisted Instruction

Insufficient teacher support may lead to inaccurate treatment implementation or omission of intervention components (Witt, Noell, LaFleur, & Mortenson, 1997; Wolery, 1997).

Moreover, research indicates that even when teachers find an intervention acceptable, are trained on its procedures, and use it correctly while being observed, they rarely implement it with fidelity across extended periods of time (Noell et al., 2005; Witt et al., 1997). Thus, teachers may require consistent cues and feedback regarding their own behavior in order to change the behavior of their students. Fortunately, advances in technology and subsequent increases in the use of computer-assisted instruction (CAI) may allow teachers to use pre-packaged interventions that automate the components that they would typically be required to implement (e.g., Stromer, Kimball, Kinney, & Taylor, 2006). For instance, computer systems can provide automatic prompts that cue both teacher *and* student behavior. Compared with traditional intervention implementation, CAI is more likely to be free of human error and eliminates some problems associated with procedural fidelity. For example, Kodak, Fisher, Clements, and Bouxsein (2011) compared therapists' implementation of both CAI and one-on-one instructional procedures and demonstrated that therapists implemented CAI with remarkably greater treatment fidelity and less training. Thus, in addition to ensuring procedural fidelity, automation of intervention components might lessen the need for in-depth training on each component of an intervention.

In summary, it may be beneficial for educators to have access to multifaceted interventions that apply an array of intervention components simultaneously to reduce transition times. Unfortunately, implementing such interventions with fidelity can require extensive training and teacher support. CAI may allow teachers to implement interventions with greater fidelity by automating some of their components. However, to date, no studies have evaluated

the use of CAI to moderate transition management. Thus, the current study seeks to evaluate the effects of a CAI-based intervention developed by the researcher to incorporate components found to be effective in previous transition-related research (e.g., Campbell & Skinner, 2004; Yarbrough et al., 2004). This system, the Keep Busy and Carry On system (KBCO), incorporates and automates multiple evidence-based strategies (i.e., explicit timing, transition warnings, cues to stay on task, contingency management) in order to promote efficient transitions between classroom activities and to increase instructional time. The current study examined the effects of this intervention on multiple measures including transition times between classroom activities, students' on-task behavior, and teachers' use of prompts and praise statements. It was hypothesized that implementation of KBCO would (a) decrease the duration of transitions between centers-based activities (CBAs); (b) increase students' on-task behavior; (c) decrease the number of teacher prompts to transition and stay on task; and conversely, (d) increase the number of praise statements made by teachers.

Method

Participants and Setting

Classroom and participant selection. The principal of a suburban elementary school in the Southeast was asked to identify classrooms that might benefit from intervention targeting classroom transitions. The researcher then met with each teacher (one first-grade and one second-grade teacher) and conducted preliminary classroom observations. After confirmation that the classrooms were equipped with an interactive white board (IWB) and the teachers consented to participating in the study, each teacher was asked to select a minimum of 5 students who might benefit from intervention. After receiving parental permission for 5 students in each class, the researcher conducted classroom observations during each classroom's CBAs to assess

the extent to which each student was completing appropriate transitions. Through this informal data collection process, 2 students in each class were chosen as participants due to their average length of transitions in comparison to other students observed. The 2 students with the greatest difficulty transitioning were selected as proxies for their classmates because collecting data on *all* students' transition times could result in compromised measurement reliability. Furthermore, considering that teachers often may refrain from beginning instruction until all students have transitioned, it was expected that decreasing the transition times of the least efficient students might also allow teachers to start lessons earlier, increasing overall instructional time for all students.

Classroom 1. Mrs. Crane had 8 years of teaching experience and was the teacher of a first-grade class of 23 students. The desks in her room were situated in small groups, with students in each group facing each other. During CBAs, students were arranged in groups of five or six students. CBAs were conducted Monday through Thursday, with groups participating in two of four centers (spending 25 min at each activity) each day. Centers included a teacher-led work group, a paraprofessional-led reading group, independent seatwork (i.e., worksheets), and free choice (e.g., learning games, silent reading, or listening center).

Student participants from Classroom 1 included Leo (Caucasian male aged 6 years, 9 months) and Samantha (Caucasian female aged 6 years, 7 months), both of whom were identified by their teacher and parents as typically developing. Neither of the students was identified as having an educational disability, and neither received any supplemental behavioral or academic intervention services.

Classroom 2. Mrs. Stoklasa had 15 years of teaching experience and was the teacher of a second-grade class of 25 students. The desks in her room were situated in small groups, with

students in each group facing each other. During CBAs, students were split into four groups of five to six students. CBAs were conducted Monday through Thursday, with groups participating in four centers (spending 15 min at each activity) each day. Activities included a teacher-led reading group, two rotations of independent seatwork, and a rotating activity consisting of silent reading, worksheets, or time set aside for finishing tasks started at other centers.

Student participants from Classroom 2 included Jonathan (Caucasian male aged 7 years, 8 months) and Jamie (Caucasian female aged 7 years, 10 months). Both students were identified by their teacher and parents as typically developing. Neither of the students was identified as having an educational disability, and neither received any supplemental behavioral or academic intervention services.

Materials

Instructional video. After baseline and before starting KBCO, both classrooms watched an 8-min instructional video explaining the rules for KBCO and how to earn points. The instructional video was created by the researcher using Microsoft PowerPoint 2007 with self-recorded narration and a combination of text, clipart, and images downloaded using an Internet search engine. The video was played from each teacher's desktop computer (PC) and projected using each classroom's IWB (SMART Board 600i; SMART Technologies). The instructional video described appropriate transition behavior and steps and defined on-task behavior using student-friendly language. It also described the distribution of group points, how all students would be given points only if everyone in their small group met criteria, and that groups would earn points for extra recess time.

KBCO videos. Intervention videos were created by the researcher using Windows Live Movie Maker (© 2010 Microsoft Corporation), were converted into .wmv files, and were

uploaded onto both teachers' computers. Videos were created using video clips downloaded from the Internet (www.youtube.com) and were comprised of a visual countdown clock and timer, text, and sound clips. These videos were played from each teacher's desktop computer and projected using each classroom's IWB. Each teacher initiated the video by opening the file on her computer, maximizing the screen, and clicking on the play button. Upon hitting the button, a computer-generated voice was broadcast over speakers, cueing students to "Go to your next center." This voice was accompanied by matching text on the IWB screen. Immediately following these cues, a visual timer (i.e., transition timer) started counting upward in seconds and displayed how long the transition was taking. During the count-up, the IWB also displayed rules indicating steps (see below) that the students needed to follow in order to receive a point for that transition. After reaching a specific time, the video beeped, signaling the end of the transition. The time requirement for completion of transitions was randomized each day and remained unknown to the students throughout the duration of the study.

Upon completion of the transition, a visual clock displayed the duration of the current CBA (i.e., how much time was left before the next transition). During the CBA, an audible tone sounded approximately every 5 min (i.e., the "on-task tone"). Accompanying this tone, a computer-generated voice asked students, "Are you on task?" while the same words were displayed under the clock. On-task tones were randomly distributed across the activity time and occurred five times per activity for Mrs. Crane's class and three times per activity for Mrs. Stoklasa's class. Toward the end of the CBA, the video gave students an auditory (voice) and visual (text on the IWB screen) 1-min transition warning.

Dependent Measure

The primary dependent measure was transition duration, which was defined as the time (in seconds) from the instance when a participant was cued to transition (either by the teacher or KBCO) to the instance when he or she began a new activity. Once the teacher (during baseline) or video (during intervention) cued the class to transition to a new activity (e.g., on the word “go”), an observer recorded the time (start time), then subsequently recorded the time (stop time) when the target student met the following criteria for 3 s consecutively: (a) seated at desk with buttocks touching seat; (b) not speaking, or asking an appropriate (i.e., CBA-related) question; and (c) academically engaged (e.g., writing instrument touching paper, eyes situated toward book or reading material, or eyes directed toward teacher during instructions or when no materials were needed). Mrs. Crane’s class completed three transitions across two activities each day (i.e., a transition to the first activity, a transition between activities, and a transition to complete CBAs). Mrs. Stoklasa’s class completed five transitions across four activities (i.e., a transition to the first activity, three transitions between activities, and a transition to complete CBAs). At the end of each session, mean transition duration was calculated by summing the number of seconds spent during transitions and dividing this sum by 3 (Classroom 1) or 5 (Classroom 2). Due to transition warnings being implemented during both baseline and intervention, recorded transition times did *not* include the length of the transition countdown (1 min); however, if a target student began to transition (e.g., put materials away and got out of seat) prior to being cued by the teacher or video, this data point was flagged for further analysis. Fortunately, this situation did not occur. Duration data were recorded by experimenters using a continuously running timer displayed on a laptop.

Supplemental Dependent Measures

On-task behavior. As a supplemental measure, a 15-s momentary time-sampling procedure was used to estimate the percentage of time that each participant was on task. Data on on-task behavior were recorded for the entirety of CBAs (50 min for Classroom 1 and 60 min for Classroom 2). Observers used an earpiece and were prompted to record participant behavior every 15 s by an audible tone. At each 15-s interval, observers recorded whether the target student was on task, off task, or engaging in another behavior (other). *On task* was recorded if the target student (a) had his/her eyes directed toward a worksheet/book/teacher, (b) was asking a peer for help, (c) was discussing a CBA-based topic with a peer or teacher, or (d) was writing on a CBA-related worksheet. *Off task* was recorded if the student was (a) not looking at either the teacher or work assignment, (b) talking to friends about topics unrelated to the activity, (c) wandering around the room (not gathering materials), (d) gathering materials that he or she should have already had (e.g., specific work folders), (e) sharpening a pencil, (f) doodling on his or her worksheet or folder, or (g) out of the room (e.g., visiting the bathroom or water fountain). *Other* was recorded if the target student was (a) sitting quietly and waiting for the teacher to deliver directions or materials or (b) transitioning across centers. These operational definitions were purposefully stringent such that if a student's behavior was coded as on task, he or she was likely engaged in behaviors leading to work completion. At the end of each session, the number of intervals recorded for each behavior was divided by the total number of intervals and converted into a percentage; the percentage of intervals of on-task behavior was computed in this manner.

Teacher prompts and praise. During transitions and activities, the frequency of teacher prompts and praise toward students was recorded. A prompt was recorded any time the teacher

issued a class- or student-directed statement or reprimand that told the student(s) where to go or to return to work (e.g., “Get back to work,” “You are off task,” “You are supposed to be over at [location],” “You are wasting recess time,” “This work will have to go home with you,” “Time to go to your next center”). Teacher prompts did *not* include explanations of whether students had met criteria for a successful transition or explanations of the rules of the game. A praise statement was recorded if the teacher issued a class- or student-directed statement involving positive praise or “catching” the student(s) engaging in an appropriate behavior (e.g., “You are working so well,” “I like how you are on task,” “You did a nice job transitioning,” “You beat the clock, nice job”). A praise statement was *not* recorded if the teacher gave simple feedback to a group in the absence of a specific praise statement (e.g., “Group B gets a point”).

Experimental Design

An ABAB design was employed to determine the effectiveness of the intervention (KBCO) on the dependent measures for each participant. Phase changes for the group-administered intervention were based on the results of the primary measure (transition duration) for the two target students in each classroom. Phase changes did not occur until both participants’ data were reasonably stable across at least three sessions.

Procedures

Baseline. Each teacher was provided with a brief protocol to follow so that baseline procedures remained consistent across sessions and consistent with what was observed during pre-baseline observations. Each protocol was identical and included behaviors that teachers were already implementing in their classroom. The steps teachers were asked to follow included the following:

- 1) When it is time for students to go to their first center, simply tell them how you usually tell them.
- 2) Prompt the students and give reminders as you do normally.
- 3) When it is time to switch between centers, use whatever strategies you normally use.
- 4) When it is time to end CBAs, simply tell students to clean up and get ready for the next activity as you do normally.

Across classrooms, CBAs began after a segment of large-group instruction. At the beginning of baseline sessions, each teacher cued students to transition by saying, “It’s time to go to your first/next center. Ready? Go.” Teachers used the same prompt across all other transitions except for the last transition of CBAs, at which time teachers instructed students to clean up their materials, stand at their desks, and raise their hands. Both teachers were asked to prompt students as they normally would, both for transitions and for staying on task (e.g., “We have 1 min left until we switch”). Recording of transition duration, on-task behavior, and teacher prompts and praise began once the teacher finished her transition cue for the first transition of CBAs (e.g., on the word “go”). During the withdrawal phase, teachers were asked to follow the same protocol they were given for baseline and to tell the class that they were not going to be playing KBCO that day.

Pre-intervention. Both teachers received a written list of instructions regarding how to implement KBCO. The instructions included the following steps:

- 1) Open the movie file labeled with the day’s date, and make sure the volume is at an appropriate level.
- 2) After each transition is over, give brief feedback to each group (e.g., label the groups that received a point). Mark all points on the daily tally sheet for each group.

- 3) When the on-task tone goes off, scan the room and verbally label which group(s) received a point. Mark points on tally sheet.
- 4) Repeat steps 2 and 3 for all CBAs.
- 5) For the last center, give out a point to groups that meet criteria and choose a member of each group to mark all of their points.

To keep track of points, each teacher was asked to post a 8.5×11 -in sheet of paper containing a large thermometer image for each group. Stages on the thermometer (to be colored in by students) coincided with the number of points the group earned each day. Once a group filled the thermometer, they earned 10 min of extra recess and replaced their old thermometer sheet with a new one. Within the context of the current study, teacher-researcher interaction was kept to a minimum, with the teacher only receiving information regarding which video file to play (marked by date) and which phase of the study should be implemented (e.g., baseline or KBCO).

Instructional video. On the day prior to the start of intervention, students watched an 8-min instructional video explaining the rules for KBCO and how to earn points. At the end of the instructional video, students were cued to ask their teacher any questions they had regarding the intervention. This large-group training took less than 10 min (including the instructional video) and occurred once for each class (i.e., after baseline and before intervention).

KBCO implementation. For the first intervention phase, each teacher started the KBCO video at the beginning of CBAs. Just as in baseline, data recording procedures commenced once the video provided the cue for students to go to their first activity. The length of the time criterion for each transition was randomized by the researcher, and the video automatically stopped when the pre-specified randomized time had expired. Students were unaware of the time criterion for each transition and were most likely unable to predict when the timer would

expire. Once the video stopped, the teacher gave brief feedback only to the groups meeting criteria and recorded points on a tally sheet. Student-specific criteria for transitioning included (a) walking safely; (b) being seated; (c) speaking quietly about CBA-related material or not talking; (d) having all materials for CBAs (e.g., writing instrument, worksheet, and/or book if needed); and (e) starting work. For a group to receive a point for a transition, all students in that group had to meet criteria.

The end of a transition triggered the start of the next CBA. During the activity portions of the session, the on-task tone would sound according to the specifications described above. When the on-task tone sounded, the teacher scanned the room and gave brief feedback only to groups meeting criteria for on-task behavior. Student-specific criteria for being on task included (a) being seated at their workplaces and (b) having their eyes on their books/worksheets *or* the teacher (if they did not have books/worksheets). For a group to receive points for being on task, all students in that group had to meet criteria.

This sequence was repeated for all transitions and activities across each session. For the first 3 days of intervention, randomized time criteria for transitions were predetermined each morning and prepared by the researcher. Transition criteria were determined using the three median baseline times across both students within each classroom. After 3 days of intervention, transition criteria were determined by rank ordering the average transition times across both target students (within each classroom) and selecting the 2nd, 3rd, and 4th fastest times. These transition times were then randomly distributed and rotated across three researcher-created videos for each classroom. These three videos (containing different criteria for each transition within each video) were rotated for use in each classroom for the duration of study. At the end

of CBAs, the teacher selected a student from each group to color in the group's thermometer with the number of points they had earned for that session.

Fading. Fading procedures were implemented only in Mrs. Stoklasa's classroom. Once the ABAB phases of the study were complete, both teachers completed a social validity measure (see below) and were asked to describe how they would alter KBCO if given the opportunity. Mrs. Stoklasa remarked that she wanted her students to become more personally responsible for staying on task during activities; thus, it was decided to remove the on-task tones for Mrs. Stoklasa's classroom. Instead, Mrs. Stoklasa downloaded a smartphone-based interval timer that vibrated every 5 min. At the end of each interval, she scanned the room and provided groups with points for meeting on-task criteria. All other KBCO components remained the same. Mrs. Crane stated that she wanted to make it more difficult for groups to earn extra recess time; thus, the number of points required to receive extra recess time was doubled for each group. Mrs. Crane implemented this strategy after formal data collection was completed, thus no related data collection occurred.

Interobserver Agreement

Secondary data collectors were undergraduate students trained by the researcher on operational definitions and data collection. Training was provided using multiple video examples of classroom transitions and classroom scenarios (created by the researcher) as well as preliminary observations in the classrooms. Training was provided until observers attained at least 95% agreement with the primary observer across three consecutive practice sessions. Interobserver agreement was assessed for at least 33% of all baseline and intervention sessions ($M = 48\%$). For duration measures, the smaller number of seconds recorded was divided by the larger number of seconds recorded and multiplied by 100 for each transition observed.

Regarding measures of on-task behavior, agreement was scored if both observers recorded the occurrence or nonoccurrence of the same behavior during an interval. Percentage of agreement was calculated by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100 for each behavior. For frequency measures (teacher prompts and praise), the number of recorded behaviors was summed (i.e., true frequency) by each observer at the end of each session. The smaller number of behaviors recorded was divided by the larger number of behaviors recorded and multiplied by 100 for each frequency variable. Mean agreement for observations in Classroom 1 was 96% for duration measures (range = 80-100%), 95% for measures of on-task behavior (range = 71-100%), and 89% for prompt/praise frequencies (range = 60-100%). Mean agreement for observations in Classroom 2 was 95% for duration measures (range = 66-100%), 93% for measures of on-task behavior (range = 69-100%), and 91% for prompt/praise frequencies (range = 72-100%).

Procedural Fidelity

During baseline, observers recorded the number of baseline protocol steps completed by each teacher. The protocol used by observers was similar to the one given to teachers but included more detail on the strategies being used during baseline. Procedural fidelity for baseline was calculated by dividing the number of steps implemented correctly by the total number of steps. This number was then converted into a percentage. Procedural fidelity data were taken for 100% of all baseline and withdrawal sessions. Interobserver agreement data were collected on procedural fidelity for 50% of all baseline sessions, with agreement of 100%. Both teachers' procedural fidelity for baseline sessions was 100%.

During KBCO, observers recorded the number of steps completed on a protocol that was similar to the one given to teachers. Procedural fidelity for KBCO was calculated by dividing

the number of steps implemented correctly by the total number of steps. This number was then converted into a percentage. Procedural fidelity data were recorded for 100% of all KBCO sessions. Interobserver agreement on procedural fidelity data was calculated for 48% of all KBCO sessions and was 92%. Mrs. Crane implemented KBCO steps with a mean of 75% accuracy (range = 58-90%), and Mrs. Stoklasa implemented steps with a mean of 85% accuracy (range = 74-96%). Given that teacher procedural fidelity of KBCO sessions was somewhat low (due to teachers not giving feedback to each team that met criteria after *every* transition or on-task tone), a follow-up analysis was conducted to determine teachers' procedural fidelity on the implementation of the basic elements of KBCO. While admittedly subjective, basic elements were defined as components that were integral to implementation of KBCO. Generally, a considerable portion of the intervention was automated, and without implementation of these basic elements (e.g., starting the video), it would be difficult to suggest that KBCO had an effect on the dependent measures. Basic elements of KBCO included the steps listed on the teacher protocol (see "Pre-Intervention" above); however, teachers were given credit for Steps 2 and 3 if they gave feedback *at least* once during the session as opposed to after *every* transition and on-task tone. Mrs. Crane implemented the basic elements (six steps) with a mean of 94% fidelity (range = 83-100%), and Mrs. Stoklasa implemented the basic elements with a mean of 96% fidelity (range = 83-100%).

Social Validity

Following the last session of KBCO, both teachers were asked to complete the Behavior Intervention Rating Scale (Elliott & Von Brock Treuting, 1991) as a measure of social validity (i.e., treatment acceptability). Questions were slightly modified to better fit the current study,

and teachers were required to respond to 24 items by indicating their level of agreement on a 6-point Likert scale (1 = strongly disagree and 6 = strongly agree).

Results

Figure 3.1 presents the results for the participants in Mrs. Crane's classroom (Leo and Samantha), and Figure 3.2 presents the results for the participants in Mrs. Stoklasa's classroom (Jonathan and Jamie). The left two panels of each figure depict the mean transition time per session across baseline and intervention sessions for each student, and the right two panels illustrate students' on-task behavior (percentage of intervals). All phase change decisions were made based upon students' transition duration data, as this was the primary measure of interest. Changes in phases were therefore not always ideal for interpreting the impact of intervention on students' on-task behavior. Means and ranges for students' transition durations and on-task behavior across phases are provided in Table 3.1.

Classroom 1

Shown in the upper-left panel of Figure 3.1, Leo's average daily transition duration across the first phase of baseline was 316 s with high variability and an increasing trend. Upon implementation of KBCO, Leo's daily average transition durations decreased immediately to below 150 s, with his mean transition duration for the first phase of KBCO sessions equaling 104 s. Upon withdrawal of the intervention, Leo's transition durations returned to baseline levels ($M = 283$ s) and increased in variability. Although data were not as stable as during the intervention phase, Leo's transition times again decreased immediately after reinstatement of the intervention ($M = 101$ s).

Leo was on task a mean of 31% of intervals during the first phase of baseline. With implementation of KBCO, his on-task behavior immediately increased to a mean of 65% of

intervals overall. His on-task behavior returned to baseline levels during the withdrawal phase ($M = 33\%$) and increased again once KBCO was reinstated ($M = 66\%$). The percentage of non-overlapping data points between baseline and intervention conditions was 100% for both transition duration and on-task behavior.

Shown in the lower-left panel of Figure 3.1, Samantha's average daily transition duration across the first phase of baseline was 267 s and remained at a high level with somewhat high variability. Once intervention began, Samantha's daily average transition durations decreased immediately to below 125 s, with her mean transition duration for the first phase of KBCO equaling 78 s. Upon withdrawal of the intervention, Samantha's transition durations returned to baseline levels and remained stable ($M = 208$ s). After reinstatement of the intervention, the first data point remained elevated (162 s) when compared to other intervention data points, but Samantha's mean transition duration per session subsequently decreased and remained low and consistent ($M = 83$ s).

The lower-right panel of Figure 3.1 depicts that Samantha exhibited on-task behavior during an average of 46% of intervals during the first phase of baseline. With implementation of KBCO, her on-task behavior increased to a mean of 59% of intervals overall. Samantha's on-task behavior returned to baseline levels during the withdrawal phase ($M = 39\%$) and increased again once KBCO was reinstated ($M = 65\%$). The percentage of non-overlapping data points between baseline and intervention conditions for Samantha was 100% for both dependent measures.

Classroom 2

As seen in the upper-left panel of Figure 3.2, Jonathan's mean daily transition duration across the first phase of baseline was 174 s, with a stable data path throughout the phase. Once

intervention began, Jonathan's daily average transition durations decreased immediately to below 80 s, with his mean transition duration for the first phase of KBCO sessions equaling 60 s. There was a slight increasing trend during this phase; however, due to the dramatic decrease in overall transition duration and the overall stability of both participants' data, a phase change was made. Upon withdrawal of the intervention, Jonathan's transition durations returned to baseline levels and remained stable ($M = 187$ s). After reinstatement of the intervention, Jonathan's transition durations again decreased immediately and remained stable ($M = 57$ s).

Illustrated in the upper-right panel of Figure 3.2, Jonathan was on task a mean of 45% of intervals during the first phase of baseline. With implementation of KBCO, on-task behavior increased to a mean of 76% of intervals overall. On-task behavior returned to baseline levels during the withdrawal phase ($M = 47\%$) and increased again once KBCO was reinstated ($M = 81\%$). During fading, Jonathan's on-task behavior remained at high levels when on-task tones were removed from the KBCO video ($M = 77\%$). The percentage of non-overlapping data points between baseline and intervention conditions for Jonathan was 100% for both dependent measures.

Shown in the lower-left panel of Figure 3.2, Jamie's mean daily transition duration across the first phase of baseline was 210 s, with the data path having a very steep increasing trend. Once intervention began, Jamie's daily average transition durations decreased immediately to below 90 s, with her mean transition duration for the first phase of KBCO sessions equaling 80 s. Upon withdrawal of the intervention, the first data point was slightly higher than intervention data points (marked with an asterisk in the graph). A potential confound within this session involved the addition of two adults' presence during CBAs. For every other session in Classroom 2, the number of adults in the classroom (1) remained constant throughout the study;

however, for Session 8 only, the addition of two more adults may have decreased the amount of time Jamie spent transitioning due to increased adult prompts, novelty of activities, and increased opportunities for adult praise. For the rest of the withdrawal sessions, Jamie's transition durations returned to baseline levels and remained stable at a high level ($M = 182$ s). After reinstatement of the intervention, Jamie's average transition durations remained stable below 84 s with a mean transition duration of 53 s.

The lower-right panel of Figure 3.2 shows that Jamie was on task a mean of 45% of intervals during the first phase of baseline. With implementation of KBCO, on-task behavior increased to a mean of 60% of intervals. Jamie's on-task behavior returned to baseline levels during the withdrawal phase ($M = 44\%$) and increased again once KBCO was reinstated ($M = 71\%$). Similar to Jonathan, during fading, Jamie's on-task behavior remained at high levels when the on-task tones were removed ($M = 81\%$). Including Session 8, the percentage of non-overlapping data points between baseline and intervention conditions was 100% for transition duration and 69% for on-task behavior.

Teacher Prompts and Praise

Figure 3.3 depicts data for both teacher participants. The left panels illustrate the frequency of teacher prompts across phases, and the right panels illustrate the frequency of praise statements. As stated above, all phase change decision were made using visual inspection of data collected for student transition duration. Means and ranges for teachers' frequency of prompts and praise are summarized in Table 3.2.

Shown in the upper-left panel of Figure 3.3, Mrs. Crane delivered an average of 33.0 prompts during baseline sessions. Mrs. Crane's frequency of prompting decreased to an average of 12.3 per session during KBCO. During the withdrawal phase, teacher prompts rose to an

overall average of 28.0 per session and once again decreased to an average of 11.0 per session once KBCO was reinstated. The upper-right panel of Figure 3.3 depicts the frequency of praise statements delivered by Mrs. Crane across phases. She delivered an average of 2.0 praise statements per session during baseline, 7.5 per session during KBCO, 3.5 per session during withdrawal, and 9.2 per session once KBCO was reinstated.

Depicted in the lower-left panel of Figure 3.3, Mrs. Stoklasa delivered an average of 45.0 prompts per session during baseline, but, once KBCO was implemented, her average delivery of prompts decreased to an average of 8.0 per session. Withdrawal of KBCO increased the frequency of prompting to an average of 38.0 per session, but once KBCO was reinstated, Mrs. Stoklasa's prompting again decreased ($M = 12.2$). Once the on-task tones were faded, her average frequency of prompting increased slightly to 16.6 per session. Regarding frequency of praise statements, Mrs. Stoklasa delivered an average of 5.0 per baseline session but increased her praise statements to 16.0 per session once KBCO was implemented. Withdrawal of KBCO decreased her average frequency to 9.5 per session, but reinstatement of KBCO resulted in an increase in her average frequency of delivered praise ($M = 29.2$). Frequency of praise increased even more ($M = 37.4$) once on-task tones were removed.

Social Validity

On the Behavior Intervention Rating Scale, both teachers rated all items between 4 and 6, suggesting that they agreed with all items and deemed the intervention appropriate. Mrs. Crane recorded a mean rating of 5.0, while Mrs. Stoklasa recorded a mean rating of 5.7. Items that teachers both rated as *strongly agree* or *agree* included "This was an acceptable intervention for the child's problem behavior," "Most teachers would find this intervention appropriate for this problem behavior," "I would suggest the use of this intervention to other teachers," "The

intervention is appropriate for a variety of children,” “This intervention is a good way to handle the child’s problem behavior,” “Other behaviors related to the problem behavior also are likely to be improved by this intervention,” and “The intervention quickly improved the child’s behavior.” The lowest-rated items (although still rated as 4 for both teachers) were “The child’s behavior will remain at an improved level even after the intervention is discontinued,” and “Using the intervention should not only improve the child’s behavior in the classroom, but also in other settings.”

Discussion

The purpose of this study was to evaluate the effects of a pre-packaged intervention (KBCO) on elementary students’ transitions between classroom activities. Secondary purposes included evaluating the effects of KBCO on the amount of time students remained on task and the frequency of prompts and praise statements issued by teachers. This study also explored the extent to which teachers could implement KBCO with ease and at a high level of procedural fidelity. Extending the results of past research implementing multi-component interventions (Ferguson et al., 2004; Yarbrough et al., 2004), the current results indicated that KBCO was successful in reducing all 4 participants’ transition durations by a mean percentage of 67% across all baseline and intervention sessions and increased all participants’ time on task by an average increase of 70%. Implementation of KBCO also decreased the frequency of teacher prompts by a mean of 58% across both teachers and increased the use of teacher praise by 304%. Lastly, although both teachers had some difficulty implementing all procedures with a high rate of fidelity ($M = 80\%$ procedural fidelity across teachers), they both implemented the core elements of KBCO at a higher level ($M = 95\%$ procedural fidelity across teachers). These results suggest that implementation of an intervention consisting of multiple evidence-based components can be

successful in reducing the length of transitions across CBAs. Past research has suggested that multi-component interventions that incorporate the components used in KBCO have had some success in reducing transition times (Campbell & Skinner, 2004; Yarbrough et al., 2004); however, the current study was the first to automate these central components.

Upon implementation of KBCO, participants not only transitioned across activities more rapidly, but they also spent more time engaged in completion of schoolwork, as evidenced by increases in the percentage of time that participants spent on task. KBCO's success in increasing the on-task behavior of students may be attributed to two possible mechanisms. First, automated components of KBCO cued students to remain on task during CBAs, and students may have monitored their own behavior and stayed on task in order to receive points for meeting on-task criteria. Second, KBCO was successful in decreasing transition times; thus, students spent more time seated at their workstations and had more time available to receive instruction or to be on task. For example, when comparing the amount of available instructional time across baseline and intervention, implementation of KBCO increased available instructional time by 28% for Mrs. Crane and 24% for Mrs. Stoklasa. Across classrooms, this increase resulted in an average of 39 extra minutes of instruction per week. These results extend previous findings suggesting that simple reductions in transition time can increase instructional time and student engagement (Coddling & Smyth, 2008). It is also expected that, simply by reducing selected participants' transition times, implementation of KBCO may have increased *all* students' instructional time. That is, since teachers often did not start small-group lessons until all students were finished transitioning, increasing the efficiency of participants' (i.e., the least efficient students') transition behavior likely allowed teachers to begin their lessons earlier.

Results also suggest that KBCO decreased the frequency of both teachers' prompting during CBAs. KBCO cued students to transition rapidly and to stay on task; thus, these automated cues may have supplanted the prompts that are traditionally reserved for the teacher. While KBCO was designed to provide these cues to students, the extent to which KBCO also provided similar cues to the *teacher* (e.g., to stay on task) is unknown. Supplemental measures for the frequency of teacher-delivered praise statements suggest that KBCO may have, in fact, reminded the teachers to provide praise to their students. Moreover, due to teachers' decreased need to continually prompt students to continue working, in conjunction with increases in students' appropriate behavior, it is possible that teachers also may have had more opportunities to praise student behavior.

Overall, the most influential element of KBCO may be its automation of intervention components through CAI. Treatment packages such as the Timely Transitions Game (Campbell & Skinner, 2004) are comprised of many components that require extensive teacher training and, unfortunately, are complex enough to compromise procedural fidelity and treatment acceptability. The KBCO system automated many intervention components that, if not automated, would likely be susceptible to lapses in procedural fidelity. Interestingly, procedural fidelity data suggest that both teachers in this study had little difficulty implementing the core elements of KBCO, mostly due to automation. However, both teachers did not implement more *optional* components of KBCO (i.e., giving students feedback and praise after *every* transition and on-task tone) with such a high level of fidelity. Specifically, while both teachers initially issued feedback for a majority of opportunities, both had difficulty maintaining this high degree of implementation. Thus, the structure and predictability of the KBCO system may have been sufficient to override suboptimal procedural fidelity and to significantly impact transition

behavior. Follow-up studies should investigate the importance of teacher-delivered feedback and whether a higher level of procedural fidelity (i.e., more feedback) would further improve student outcomes.

Limitations

This study had a number of notable limitations. First, this study employed a treatment package; thus, intervention components could not be tested in isolation. Therefore, the necessity of each component and the extent to which individual components affected the results are unknown. There is some reason to believe that transition signals alone were not responsible for behavior change due to the fact that both teachers used transition signals before implementation of KCBO, albeit inconsistently and with no known positive effects. It is possible that an increase in the consistent use of transition signals affected the data to some degree, but past research would suggest otherwise (e.g., Cote et al., 2005). It is more likely that explicit timing procedures and contingency management had a more substantial effect on student behavior. Although the current study did not allow for evaluation of component effects, this study extended the results of past research by demonstrating decreased transition durations in response to a multi-component intervention (Coddington and Smyth, 2008; Ferguson et al., 2004).

Second, given that phase changes were made based on transition duration data and were less than ideal in relation to supplemental measures, caution should be exercised when interpreting findings related to such measures. Future researchers should isolate components of KBCO and make phase change decisions based solely on measures affected directly by each particular component (i.e., the effects of on-task tones on on-task behavior).

Lastly, data were only collected for two students per classroom. Intuitively, the effects of KBCO on the *most* efficient students were likely minimal, given that it would have been difficult

for these students to decrease their transition times as dramatically as did the study participants. Other collateral effects, however, most likely included exposure to an increase in overall instructional time and an increase in teacher praise for appropriate behavior. Future research incorporating the use of videotaping within the classroom would allow for a more comprehensive assessment of the class as a whole.

Implications for Practice

Despite aforementioned limitations, the current study presented an innovative system for increasing the efficiency and productivity of CBAs in the classroom setting. Using KBCO to automate multiple strategies eliminated some of the taxing features associated with complex interventions and increased teachers' potential for success even despite suboptimal procedural fidelity. Automated interventions such as KBCO may provide structure to a class with poorly regulated transitions, which are likely influenced by both student and teacher behavior. KBCO provided automated cues to both students and teachers to transition quickly and to stay on task. These cues likely prompted both parties to monitor their own behavior and, in turn, significantly impacted the efficiency and productivity of the classroom. Automated cues available through systems like KBCO also may lessen the degree to which teachers must interrupt or stop their teaching. By relying on automated cues, teachers are less likely to have to provide prompts to manage student behavior. Thus, along with decreased transition durations, fewer interruptions to teaching may increase the amount of time that teachers can devote to providing instruction. Furthermore, with less transition time, more instructional time, and higher student engagement (on-task behavior), teachers likely are more available to provide praise for appropriate behavior.

Future Research

Although the addition of on-task tones was meant to bolster the effects of KBCO on on-task behavior, future research could investigate how individual components of KBCO (i.e., decreased transition times and consistent cues for students to stay on task) comparatively regulate on-task behavior. Future studies should also investigate the effects of the frequency and schedule of on-task reminders and the extent to which these reminders can be managed by automation, the teacher, or the students themselves (i.e., self-monitoring).

Second, an advantage of interdependent group contingencies is that they often encourage classmates to prompt one another to engage in expected behavior (Coogan, Kehle, Bray, & Chafouleas, 2007; Davies & Witte, 2000). Although the current study demonstrated a decrease in teacher prompting, future research should investigate KBCO's effect on peer prompting and whether working toward a common goal increases the frequency of peer prompts. Similarly, the current study only reported on the amount of teachers' class-directed praise statements. Future research is needed to investigate the importance of teacher praise toward specific students during KBCO and whether it has a differential effect on behavior.

Lastly, future studies should focus on making KBCO more user-friendly and adaptable. KBCO is a system that can be easily packaged and downloaded for many types of contexts and classrooms and, with its inclusion of an instructional video, can be used by teachers without any professional or expert involvement. Although most schools provide training to their teachers in the use of CAI for classroom instruction, teachers are very infrequently taught to use CAI for classroom management. There is also a lack of peer-reviewed research investigating the use of CAI for behavior management within the classroom setting. Future research in these areas has the potential to uncover additional CAI strategies that will benefit students and teachers alike.

References

- Arlin, M. (1979). Teacher transitions can disrupt time flow in classrooms. *American Educational Research Journal*, 16 (1), 42-56.
- Brophy, J. (1988). Educating teachers about managing classrooms and students. *Teaching & Teacher Education*, 4, 1-18.
- Cameron, C. E., Connor, C. M., & Morrison, F. J. (2005). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology*, 43, 61-85.
- Campbell, S., & Skinner, C. H. (2004). Combining explicit timing with an interdependent group contingency program to decrease transition times: An investigation of the Timely Transitions Game. *Journal of Applied School Psychology*, 20(2), 11-27.
- Christ, T. J., & Christ, J. A. (2006). Application of an interdependent group contingency mediated by an automated feedback device: An intervention across three high school classrooms. *School Psychology Review*, 35, 78-90.
- Codding, R. S., & Smyth, C. A. (2008). Using performance feedback to decrease classroom transition time and examine collateral effects on academic engagement. *Journal of Educational and Psychological Consultation*, 18, 325-345.
- Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *School Psychology Quarterly*, 22, 540-556.
- Cote, C. A., Thompson, R. H., & McKerchar, P. M. (2005). The effects of antecedent interventions and extinction on toddler compliance with transitions. *Journal of Applied Behavior Analysis*, 38, 235-238.

- Davies, S., & Witte, R. (2000). Self-management and peer-monitoring within a group contingency to decrease uncontrolled verbalizations of children with attention deficit-hyperactivity-disorder. *Psychology in the Schools, 37*, 135-147.
- Elliott, S. N., & Von Brock Treuting, M. (1991). The behavior intervention rating scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology, 29*, 43-51.
- Evans, T., Skinner, C. H., Henington, C., Sims, S., & McDaniel, E. C. (2002). Conspicuous and covert timing during CBM of mathematics across African-American and Caucasian students: An investigation of situational bias. *School Psychology Review, 31*, 529-539.
- Ferguson, A., Ashbaugh, R., O'Reilly, S., & McLaughlin, T. F. (2004). Using prompt training and reinforcement to reduce transition times in a transitional kindergarten program for students with severe behavior disorders. *Child & Family Behavior Therapy, 26* (1), 17-24.
- Fudge, D. L., Skinner, C. H., Williams, J. L., Cowden, D., Clark, J., & Bliss, S. L. (2008). Increasing on-task behavior in every student in a second-grade classroom during transitions: Validating the color wheel system. *Journal of School Psychology, 46*, 575-592.
- Gettinger, M. (1995). Increasing academic learning time. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (3rd ed.; pp. 943-964). Washington, DC: National Association of School Psychologists.
- Hollowood, T. M., Salisbury, C. L., Rainforth, B., & Palombaro, M. M. (1995). Use of instructional time in classrooms serving students with and without severe disabilities. *Exceptional Children, 61*, 242-252.

- Kodak, T., Fisher, W. W., Clements, A., & Bouxsein, K. J. (2011). Effects of computer-assisted instruction on correct responding and procedural integrity during early intensive behavioral intervention. *Research in Autism Spectrum Disorders, 5*, 640-647.
- McCord, B. E., Thomson, R. J., & Iwata, B. A. (2001). Functional analysis and treatment of self-injury associated with transitions. *Journal of Applied Behavior Analysis, 34*, 195-210.
- Noell, G. H., Roane, H. S., VanDerHeyden, A. M., Whitmarsh, E. L., & Gatti, S. L. (2000). Programming for communication in the classroom following an assessment of skill and performance deficits. *School Psychology Review, 29*, 429-442.
- Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L.,... & Duhon, G. J. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*, 87-106.
- Rhymer, K. N., Skinner, C. H., Jackson, S., McNeill, S., Smith, T., & Jackson, B. (2002). The 1-minute explicit timing intervention: The influence of mathematics problem difficulty. *Journal of Instructional Psychology, 29*, 305-311.
- Sainato, D. M. (1990). Classroom transitions: Organizing environments to promote independent performance in preschool children with disabilities. *Education & Treatment Special Issue: Organizing Caregiving Environments for Young Children with Handicaps, 13*, 288-297.
- Sainato, D. M., Strain, P. S., Lefebvre D., & Rapp N. (1987). Facilitating transition times with handicapped preschool children: A comparison between peer-mediated and antecedent prompt procedures. *Journal of Applied Behavior Analysis, 20*, 285-291.

- Stromer, R., Kimball, J. Kinney, E. M., & Taylor, B. A. (2006). Activity schedules, computer technology, and teaching children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 21*(1), 14-24.
- Witt, J. C., Noell, G. H., LaFleur, L. H., & Mortenson, B. P. (1997). Teacher usage of interventions in general education: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30*, 693-696.
- Wolery, M. (1997). Encounters with general early education: Lessons being learned. *Journal of Behavioral Education, 7*, 91-98.
- Wurtele, S. K., & Drabman, R. S. (1984). "Beat the Buzzer" for classroom dawdling: A one-year trial. *Behavior Therapy, 15*, 403-409.
- Yarbrough, J. L., Skinner, C. H., Lee, Y. J., & Lemmons, C. (2004). Decreasing transition times in a second grade classroom: Scientific support for the Timely Transitions Game. *Journal of Applied School Psychology, 20*, 85-107.
- Zeece, P. D., & Crase, S. J. (1982). Effects of verbal warning on compliant and transition behavior of preschool children. *Journal of Psychology, 112*, 269-274.

CHAPTER 4

CONCLUSION

The increasing emphasis on implementation of evidence-based practices has been accompanied by a growing awareness of the gap between research and practice, and has become a matter of concern. This dissertation presented two manuscripts linked by the interaction between research and practice. The first manuscript highlighted problematic scenarios encountered during implementation of token economies and how revisiting basic experimental research can guide practitioners as to how to integrate this research into practice. As robust as the literature surrounding token economies is, practitioners may not make regular contact with the basic research that has defined many of the behavior analytic practices that we use today. The purpose of this manuscript was to highlight the underlying mechanisms of token economies and translate the often complex implications of basic research. Hopefully, practitioners can use this information to improve their own practice as well as their confidence in disseminating the evidence-based reasons as to why token economies have been so effective across so many contexts for so many years.

The second manuscript presented a research study directly applicable to everyday problems confronted in practice: that of inefficient and lengthy transitions. Results suggested implementation of the intervention decreased the amount of time participants spent transitioning across activities during centers-based activities, increased students' on-task behavior, decreased the number of prompts teachers used to cue students to transition and stay on task, and

conversely increased the number of praise statements made by teachers. This study presented an innovative system for increasing efficiency and productivity in the classroom setting.

Overall, the outcome of the research-to-practice gap is often frustration and dissatisfaction. Completing practice-based research is time-consuming and labor intensive. Solutions to these barriers and problems will bring enormous benefits to teachers and students. The manuscripts presented in this dissertation will hopefully serve as models for ways in which practitioners can translate and apply past research to guide practical applications of common interventions, as well as how to conduct research to solve common problems observed in practice.

Table 3.1

Phase Means and Ranges for Baseline, Intervention, Withdrawal, and Reinstatement Conditions

	Transition Duration					On-Task				
	Baseline	KBCO	Baseline	KBCO	Fading	Baseline	KBCO	Baseline	KBCO	Fading
	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)
Leo	316 (218-432)	104 (62-149)	283 (228-364)	101 (54-157)	N/A	31% (26-39)	65% (58-76)	33% (26-40)	66% (53-81)	N/A
Samantha	267 (210-333)	78 (28-122)	208 (183-240)	83 (43-162)	N/A	46% (44-47)	59% (51-70)	39% (34-45)	65% (55-86)	N/A
Jonathan	174 (161-187)	60 (45-80)	187 (175-203)	52 (37-66)	62 (49-89)	45% (28-56)	76% (67-84)	47% (34-63)	81% (77-90)	77% (71-91)
Jamie	210 (186-307)	80 (62-90)	182 (115-221)	55 (29-84)	52 (45-60)	45% (38-57)	60% (58-61)	44% (31-69)	71% (64-76)	81% (73-85)

Note. Fading data were only collected for Jonathan and Jamie (Classroom 2).

Table 3.2

Phase Means and Ranges for all Baseline, Intervention, Withdrawal, and Reinstatement Conditions for Teachers' Delivery of Prompts and Praise Statements

	Frequency of Prompts					Frequency of Praise				
	Baseline	KBCO	Baseline	KBCO	Fading	Baseline	KBCO	Baseline	KBCO	Fading
	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)	<i>M</i> (range)
Mrs. Crane	33.0 (30-37)	12.3 (4-22)	28.0 (21-34)	11.0 (5-18)	N/A	2.0 (1-3)	7.5 (3-13)	3.5 (1-5)	9.2 (6-12)	N/A
Mrs. Stoklasa	45.0 (31-52)	8.0 (2-12)	38.0 (21-54)	12.2 (7-18)	16.0 (10-24)	5.0 (3-8)	16.0 (14-19)	9.5 (5-15)	29.2 (18-36)	37.4 (23-56)

Note. Fading data were only collected for Mrs. Stoklasa's classroom.

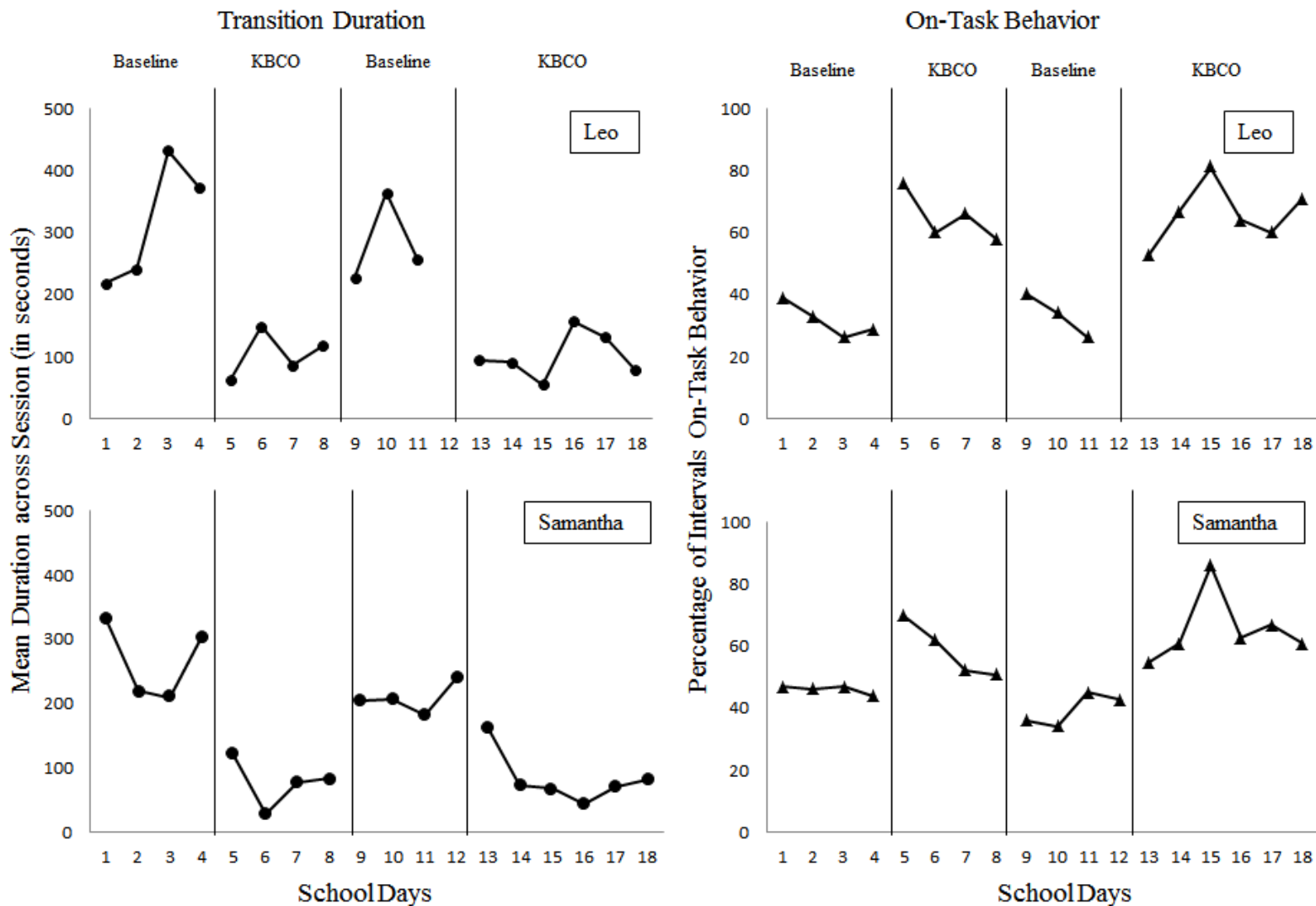


Figure 3.1. Effects of class-wide intervention on transition duration and on-task behavior for Leo and Samantha. Filled circles in the left panels represent the average transition durations across each session for both participants in Classroom 1. Filled triangles in the right panels represent the percentage of intervals during which each participant was on task during each session.

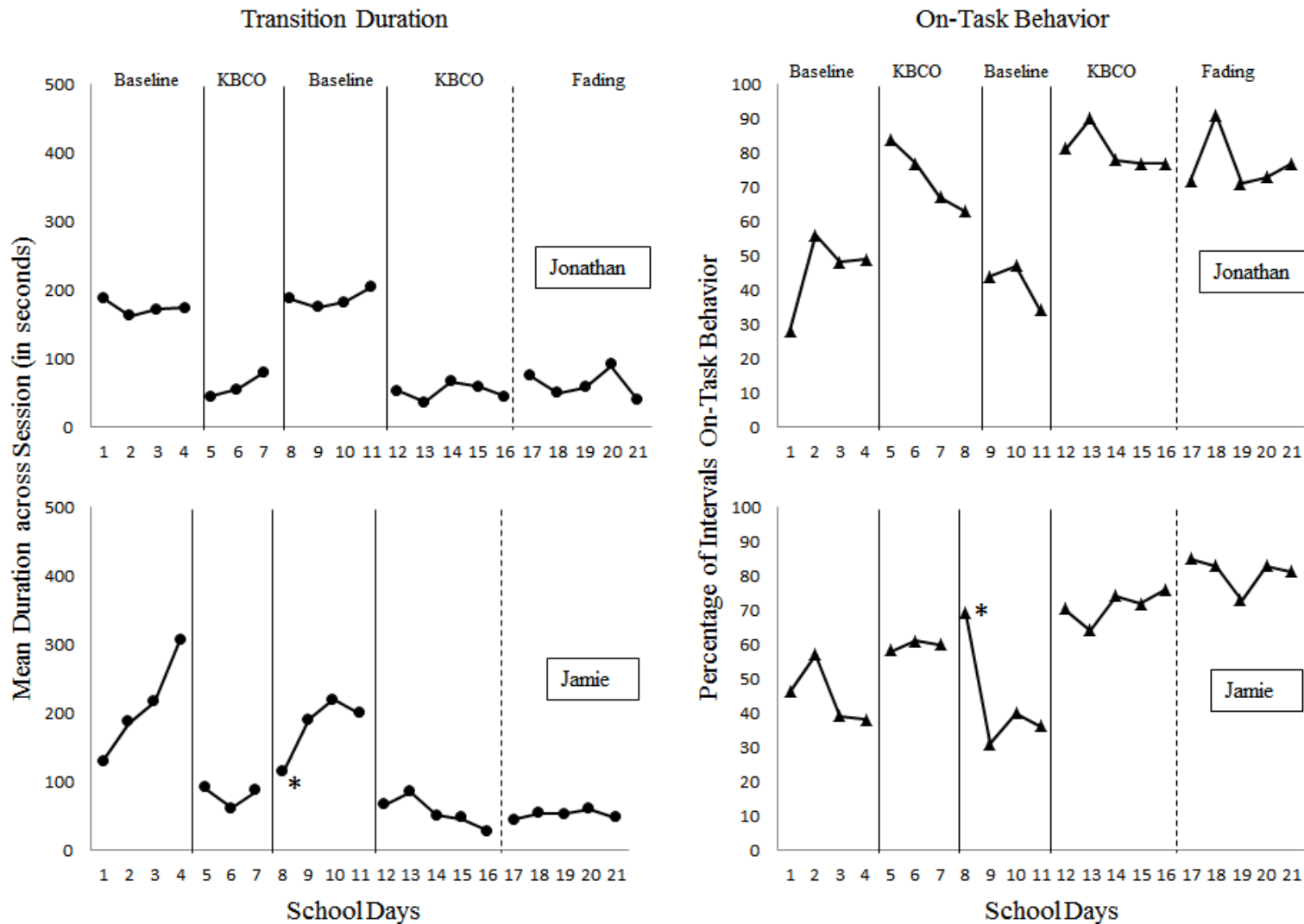


Figure 3.2. Effects of class-wide intervention on transition duration and on-task behavior for Jonathan and Jamie. Filled circles in the left panels represent the average transition durations across each session for both participants in Classroom 2. Filled triangles in the right panels represent the percentage of intervals during which each participant was on task during each session.

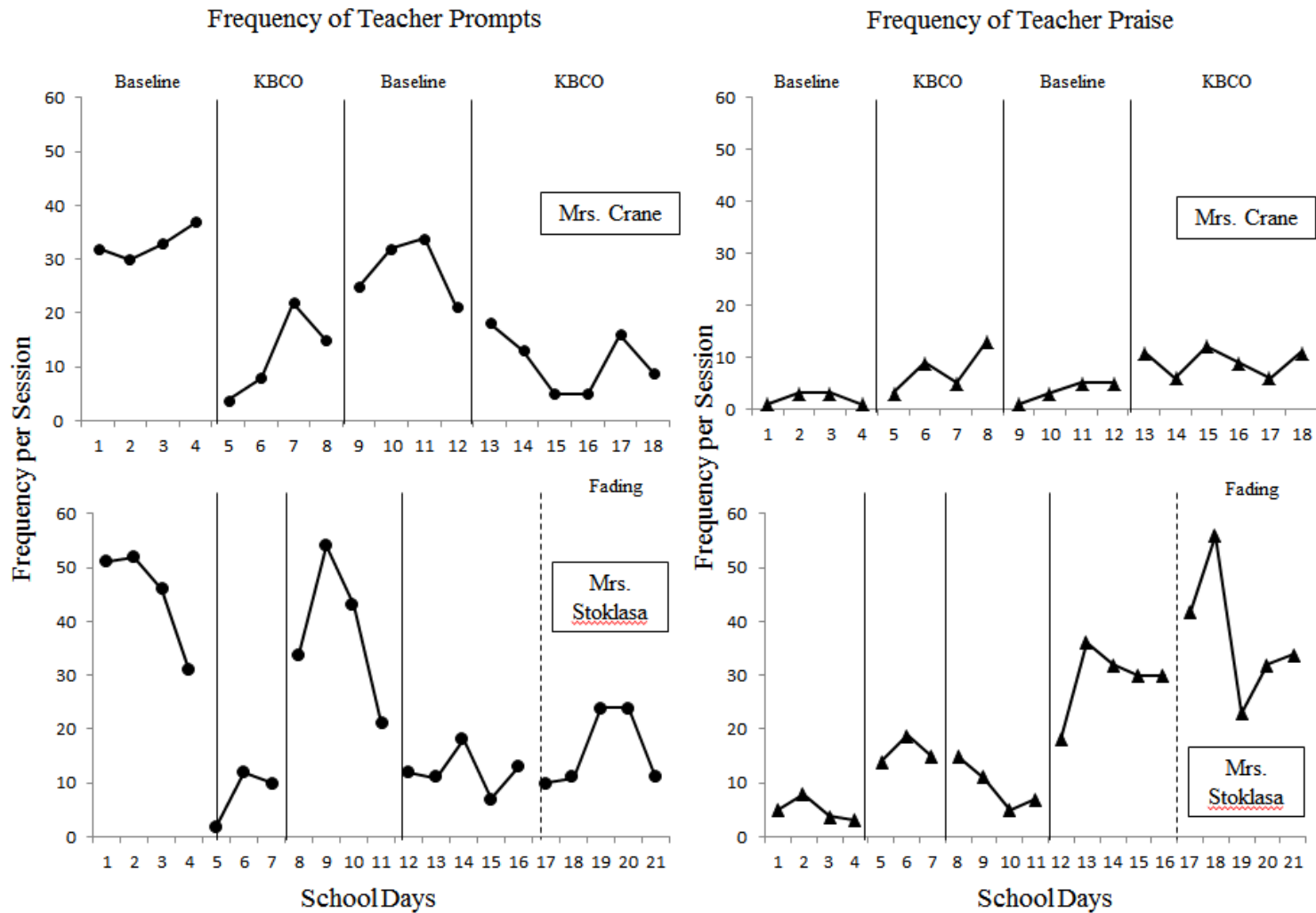


Figure 3.3. Effects of class-wide intervention on frequency of teacher prompts and praise statements. Filled circles in the left panels represent the frequency of teacher prompts across each session for Mrs. Crane (1st Grade) and Mrs. Stoklasa (2nd Grade). Filled triangles in the right panels represent the frequency of praise statements across sessions for both teachers.