Evaluation of Performance-Based and Pre-set Conventional Criterion for Reinforcement in Check In-Check Out

Lauren Lestremau Harpole
University of Southern Mississippi

Follow this and additional works at: https://aquila.usm.edu/dissertations

Part of the Child Psychology Commons, Educational Psychology Commons, and the School Psychology Commons

Recommended Citation
https://aquila.usm.edu/dissertations/861

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.
EVALUATION OF PERFORMANCE-BASED AND PRE-SET CONVENTIONAL CRITERION FOR REINFORCEMENT IN CHECK IN-CHECK OUT

by

Lauren Lestremau Harpole

Abstract of a Dissertation
Submitted to the Graduate School
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

August 2012
ABSTRACT

EVALUATION OF PERFORMANCE-BASED AND PRE-SET CONVENTIONAL CRITERION FOR REINFORCEMENT IN CHECK IN-CHECK OUT

by Lauren Lestremau Harpole

August 2012

The purpose of the present study was to compare the effectiveness of two methods of criterion-setting, performance-based or pre-set conventional, as evidenced by improvements in children’s behavior. Participant behavior was evaluated through teacher reports of appropriate behavior and observed academically engaged behavior as well as decreases in problem behavior and disruptive behavior. Eight elementary school students in a Southeastern town referred for exhibiting behavior problems served as participants in addition to their teachers. The effects of the different methods of criterion setting on the dependent variables were evaluated. Teacher ratings of appropriate behavior were assessed through evaluation of Daily Behavior Report Card (DBRC) point data. Direct observations were conducted to determine target students’ and control peers’ levels of observed appropriate behavior and problem behavior. Disruptive behavior was evaluated as the frequency of office discipline referrals (ODRs). Treatment integrity was assessed through direct observations as well as a review of permanent products. Acceptability was assessed for adult and child participants. The current study serves as one of the few studies in the Check In-Check Out (CICO) literature to (a) examine various methods of criterion setting, specifically Performance-Based methodology; (b) present teacher ratings of appropriate behavior through use of DBRCs; (c) conduct direct observations of target students and control peers’ appropriate behavior; and (d) evaluate treatment
integrity for all days of CICO implementation with supplementary direct observations of treatment integrity. Results suggest that CICO was effective in improving participant behavior without providing evidence of the superiority of either method of criterion setting.
The University of Southern Mississippi

EVALUATION OF PERFORMANCE-BASED VERSUS PRE-SET CONVENTIONAL CRITERION FOR REINFORCEMENT IN CHECK IN-CHECK OUT

by

Lauren Lestremau Harpole

A Dissertation Submitted to the Graduate School of The University of Southern Mississippi in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Approved:

D. Joe Olmi
Director

Brad A. Dufrene

Michael Mong

Sterett H. Mercer

Susan A. Siltanen
Dean of the Graduate School

August 2012
ACKNOWLEDGMENTS

The writer would like to thank the dissertation chair and mentor, Dr. Joe Olmi, as well as the other committee members, Drs. Brad Dufrene, Michael Mong, and Sterett Mercer. You have all provided great insight and support throughout this project and the whole of my graduate training and I appreciate you all more than you know. Drs. Heather Sterling-Turner and Daniel Tingstrom, thank you both also for your encouragement and guidance throughout my graduate training. You have all provided me with a graduate training experience for which I will remain grateful and remember fondly throughout my career and life.

Thanks especially to my colleagues who have helped with this project, Chandler McLemore, Julie Sherman, and Rachel Mitchell. Thank you also to all of my colleagues and now lifetime friends who have been invaluable through my graduate training, Katie Bellone, Shelley Benshoof, Abby Lambert, Matt LeGray, Katy Menousek, Laura Needelman, and Kim Zoder-Martell.

Special thanks go to my family, my husband, and my parents for their never-ending faith, support, and love.
LIST OF TABLES

Table

1. Grade, School, Race, Referral Concerns, and Criterion for Reinforcement for Each Participant .......................................................... 44

2. Means, Ranges, and Magnitude of Change for Percentage of Teacher Ratings of Appropriate Behavior across Phases for Participants in Pre-Set Conventional Participants .......................................................... 63

3. Means, Ranges, and Magnitude of Change for Percentage of Teacher Ratings of Appropriate Behavior across Phases for Participants in Performance-Based Participants .................................................. 64
LIST OF ILLUSTRATIONS

Figure

1. Percentage of Teacher Ratings of Appropriate Behavior Daily for Each Participant with Pre-Set Conventional Criteria for Reinforcement ..........................60

2. Percentage of Teacher Ratings of Appropriate Behavior Daily for Each Participant with Performance-Based Criteria for Reinforcement ..........................61

3. Academically Engaged Behavior Across Phases for Participants in the Pre-Set Conventional Participants ..................................................................................69

4. Academically Engaged Behavior Across Phases for Participants in the Pre-Set Conventional Participants ..................................................................................72

5. Problem Behavior Across Phases for Participants in the Pre-Set Conventional Participants ........................................................................................................76

6. Problem Behavior Across Phases for Participants in the Performance-Based Group ..................................................................................................................79
CHAPTER I
INTRODUCTION

Approximately 6% of students in a classroom exhibit problem behaviors that require intervention, not including the additional students who exhibit behaviors that hinder their learning and the learning of others (Clunies-Ross, Little, & Kienhuis, 2008). Not only does disruptive behavior interfere with academic and vocational success, but it may also result in chronic maladjustment and unhappiness (Kazdin, 1987). Although the broad goals of schools are academic, addressing increasing levels of disruptive behavior is a necessary precursor for academic achievement (Nelson, 1996). Approximately 76% of teachers surveyed reported they would be better able to provide instruction and educate their students effectively if not faced with such high levels of discipline problems (Public Agenda, 2004). The American public recognizes this growing problem and ranks the behavior problems of children in schools as one of its top concerns (Cheney, Flower, & Templeton, 2008).

Educators and laypersons alike agree on the significance of discipline issues in our schools; however, handling problem behaviors is challenging for teachers, administrators, and consultants (LaRue, Weiss, & Ferraioli, 2008). Disproportionate amounts of time are spent by administrators and teachers dealing with a few problematic students (Cheney et al., 2008), causing these problem behaviors to drain allocated instructional time and monetary resources. Yet, most educators do not possess the expertise needed to implement proactive behavioral strategies in an effective and efficient manner (Fairbanks, Sugai, Guardino, & Lathrop, 2007). For this reason, problem behavior continues to be dealt with through reactive and punitive approaches.
Although punishment-based approaches (e.g., detention, suspension, expulsion) are used commonly, their effectiveness has not been proven. In fact, Covell (2009) indicates that punishment-based approaches have had little positive effect on changing the behavior, attitudes, or achievement of the students exhibiting the behaviors. One approach to misbehavior, zero-tolerance, has been shown to be ineffective for managing the behavior of some students and at times even counterproductive, having been linked to declines in academic achievement and increases in misbehavior, dropouts, and impaired relationships with adults (Armistead, 2008). Therefore, it is an imperative that schools modify their methods of addressing student misbehavior.

Sprick (2009) suggests that schools focus on proactive approaches that better serve all students. School-wide Positive Behavior Supports (SWPBS) represents one of these approaches and includes the use of empirically-supported interventions targeted at creating safer and more positive school climates. SWPBS may and should exist as a component of a larger more comprehensive Response to Intervention (RTI) model. SWPBS is centered on achieving desired student outcomes, such as increased attendance, academic engagement, and appropriate behavior (Sprick, 2009). A well designed RTI model ensures that students who are not responsive to universally available programming are provided increasingly intense individualized interventions tailored to their level of need (Mellard & Johnson, 2008). Sugai et al. (2000) noted the ease with which RTI can be applied within the proactive and prevention-focused SWPBS model. Additionally, Sprick highlights how RTI has made positive changes in the manner that at-risk students are provided services.
For students who have previously slipped through the cracks, this union of SWPBS and RTI increases the likelihood of success for all students. This union represents the recommended continuum of behavior supports: School-wide primary interventions (Tier 1), targeted interventions for students at-risk (Tier 2), and individualized intensive interventions for students exhibiting severe behaviors (Tier 3) (Crone, Horner, & Hawken, 2004; Hawken, 2006). This continuum allows for time and resources to be preserved through implementation of Tier 2 targeted interventions to address the needs of students who do not respond to school-wide primary supports, but are less resource-intensive than the function-based individualized interventions typically implemented in Tier 3 (Crone et al., 2004).

In line with these goals, interventions that help educators overcome obstacles such as lack of training, time, and funding in order to better meet the needs of at-risk students merit continued research. Therefore, this study extends the research on Check In-Check Out (CICO), also referred to as the Behavior Education Program (BEP) (Crone et al., 2004), which serves as an intervention to address the behavioral needs of these at-risk students. Although CICO serves as a promising intervention, the CICO literature is plagued with studies involving poor evaluations of treatment integrity, behavior change data consisting only of office discipline referrals (ODRs), a lack of direct observation data, and a failure of previous studies to discuss challenges incurred during implementation. Therefore, the current study addressed a limitation within the CICO literature by examining methods of setting the criterion for reinforcement within CICO.
Check In-Check Out

Through CICO, behavior support, performance feedback, and positive adult attention are provided daily to the student determined to be in need. It also sets the stage for an increased level of communication between the home and school (Crone et al., 2004). Through the use of CICO, an increase in appropriate behavior is anticipated, thereby likely resulting in greater levels of success in school. Behavior expectations are derived and represent positively stated (i.e., tell the child what to do) alternative behaviors. Behavior change results from the increased access to positive social attention in the form of precorrections about the student’s behavior expectations, praise for exhibiting behavior expectations, and corrective feedback when the student fails to exhibit the behavior expectation. Rewards are also provided contingent on displaying a criterion level of appropriate behavior. Systematic feedback is provided to the target student by his CICO coordinator and his teachers through the use of the Daily Behavior Report Card (DBRC) (Chafouleas, Riley-Tillman, & McDougal, 2002).

This structured feedback centers on the use of a DBRC, which allows for increased structure and feedback in the student’s environment (Todd, Campbell, Meyer, & Horner, 2008). The DBRC does not represent a new intervention; in fact it has been present in the literature for decades in the following forms: The Daily Report Card (DRC) (Dougherty & Dougherty, 1977), home notes (Blechman, Schrader, & Taylor, 1981), school-home notes (Kelley, 1990), and home-school notes (Long & Edwards, 1994). Varied descriptions and uses of DBRCs exist, highlighting its flexibility (Chafouleas, Riley-Tillman, McDougal, 2002). Chafouleas, Riley-Tillman, and McDougal define a
DBRC as a measure in which specific behaviors are rated daily and those ratings are shared with non-raters (e.g., student, parents).

More recently, another term, direct behavior rating (DBR) (Christ, Riley-Tillman, & Chafouleas, 2009), has been born and it is suggested that DBRC falls under the umbrella term of DBR that represents the brief rating of target behavior(s) following a given observation interval (Fabiano, Vujnovic, Naylor, Pariseau, & Robins, 2009). The current discussion of these instruments will refer to DBRC while presenting research associated with both DBRC and DBR. One aspect that has ranged widely within the DBRC and DBR literature represents the type of rating scale used, which Long and Edwards (1994) suggest can be determined, and therefore varied, upon creation of the DBRC. This flexibility has resulted in these rating scales ranging from Likert scores associated with various qualitative or quantitative demonstrations of appropriate behavior to percentages of behavior exhibited during the rating interval. Despite the variability in rating type, the literature has suggested that the DBRC methodology is robust and can be effectively used as an intervention in and of itself, a component of an intervention package, or as a measurement tool to progress monitor the effectiveness of other interventions (Chafouleas, Riley-Tillman, & McDougal, 2002). The adequacy of the DBRC methodology will be highlighted below.

Initial technical evaluations of the DBRC suggest that it possesses internal consistency, temporal stability, and concurrent validity, and that its sensitivity allows for detection of treatment effects (Pelham, Fabiano, & Massetti, 2005). In another evaluation of temporal stability, Fabiano et al. (2009) correlated average DBRC percentages between odd-numbered and even-numbered days and found significant temporal stability.
Chafouleas, Riley-Tillman, Sassu, LaFrance, and Patwa (2007) agree with Fabiano et al. in their postulation that DBRC measures are reliable. In a study by Chafouleas, McDougal, Riley-Tillman, Panahon, and Hilt (2005), 82% to 87% of DBRC ratings and direct observations of on/off task behavior and disruptive behavior were moderately correlated ($r = .67$), highlighting the agreement between direct observation and DBRCs. In a replication study, a significant correlation ($r = .81$) was again found between mean teacher DBRC and mean observer direct observation ratings of on-task behavior (Riley-Tillman, Chafouleas, Sassu, Chanese, & Glazer, 2008). Together these studies suggest that DBRCs can serve as accurate and reliable tools to obtain behavior information in schools.

Before DBRCs can be systematically utilized, target behaviors must be operationalized and then DBRCs should be used consistently, at a given time and in a specific place with a set frequency, and the data gleaned should be scored, graphed, and visually analyzed often (Riley-Tillman, Chafouleas, & Briesch, 2007). Although direct observation represents the standard in assessment, DBRCs help to address the concerns present in applied settings, where collecting direct observation data and monitoring multiple cases is often unrealistic (Chafouleas et al., 2005). DBRCs’ ease of use, requiring less than one minute to complete, represent an additional strength in application (Chafouleas et al., 2005) and represent a critical component within CICO.

Before providing a description of CICO, issues present within the CICO literature must be clarified. CICO is referred primarily in the literature as the BEP (Crone et al., 2004), as few researchers have referred to the program as CICO (Filter et al., 2007; Todd et al., 2008). However, less formally (e.g., in school districts), this intervention is referred
to primarily as CICO. Hawken, one of the seminal researchers in the BEP literature, suggests that the BEP is a CICO intervention (Hawken, Petterson, Mootz, & Anderson, 2006). In addition to its name, across the BEP and CICO literature, another variation exists in regard to the presence or absence of SWPBS. Todd et al. noted that having SWPBS in place during CICO implementation may enhance the effectiveness of CICO. Similarly, Fairbanks, Simonsen, and Sugai (2008) indicated that although Tier 2 interventions are more effective when implemented in conjunction with existing SWPBS (i.e., tier 1) frameworks, this layering is not required. This connotation suggests that although having SWPBS in place prior to implementation of CICO may be beneficial, SWPBS may not be a precursor to CICO, as it is for the BEP (Crone et al., 2004). However, this implication by Todd et al. is largely inconsistent with the literature base of BEP and CICO, as both are predominately discussed within the SWPBS literature. Therefore, to clarify, the literature suggests that BEP and CICO represent the same intervention, can be used interchangeably, and although they typically occur within a SWPBS system, this is not a requirement.

When it is determined that school-wide, primary interventions are not effective, CICO should be implemented (Hawken, 2006). CICO is most often implemented as a Tier 2 intervention, to decrease the likelihood that severe problem behavior will develop in these at-risk students. When implemented in a Tier 2 context, it has been shown to be effective with 60-75% of these at-risk students (Crone et al., 2004). To begin CICO, the student nominates a staff member with whom they enjoy interacting and who has flexibility in his or her schedule to accommodate check-in and check-out times. March and Horner (2002) recommend establishing a written contract between the student,
parent, and the CICO coordinator which outlines the behavior expectations, rewards contingent on obtaining point goals, and parties involved. CICO involves daily checking in by the student before school with his CICO coordinator. During check-in, the Coordinator prepares the student for the day by providing him or her with a DBRC and by providing precorrections about the student’s behavior expectations that are listed on his DBRC. Also listed on the student’s DBRC are his classes, and space for the teacher ratings, assignments, and a positive comment about the student’s behavior.

As the student enters each class or begins a new class period, he is greeted by his teacher who provides precorrections about his behavior expectations. In this manner, the student is repeatedly reminded of what is expected of him in the school and classroom settings, which decreases the likelihood that he will exhibit problem behaviors. The teacher evaluates whether or not or to what degree, the student met his behavior expectations and rates him at the end of the class/time period on his DBRC. Specific performance feedback is then provided to the student (Hawken, 2006; March & Horner, 2002).

At the end of the day, the student visits his CICO coordinator again to check out (Crone et al., 2004). During check-out, the Coordinator determines the number of points the student earned that day, if the student met his point goal, and provides the student a reward (selected based on student reported preference and CICO coordinator approval), if earned. The coordinator sends the DBRC home to the student’s parents to sign via the student and have it returned to the coordinator the next morning at check-in (Hawken, 2006; March & Horner, 2002). Progress monitoring occurs through evaluation of the DBRC data, and these data can be graphed to further aid in data-based decision making.
and to inform any modifications that may be needed (McCurdy, Kunsch, & Reibstein, 2007). Although these procedures represent the standard protocol, modifications can be made to nearly all aspects of the intervention (e.g., criterion for reinforcement changed, rewards varied, and number of check-in times increased). If the student exhibits success with CICO, feedback may be decreased, and the student can be gradually transitioned from CICO (McCurdy et al., 2007).

Although the components of CICO described above do not represent new interventions, their packaging into what is now referred to as CICO has only begun to be researched recently. Studies of CICO have produced desirable, albeit variable, results across a variety of dependent variables and populations. CICO has been used to decrease disruptive behaviors (Harpole, 2010; Hawken & Horner, 2003; Hawken, MacLeod, & Rawlings, 2007; Todd et al., 2008), to increase prosocial behaviors (Harpole, 2010; McCurdy et al., 2007), and to increase academic engagement (Hawken & Horner) across elementary school students (Hawken et al., 2007; McCurdy et al., 2007; Todd et al., 2008), middle school students (Hawken & Horner, 2003; March & Horner, 2002), and high school students (Harpole, 2010). However, despite the gains evidenced across these studies, limitations remain within the CICO literature. Paramount in these limitations is the method and frequency used to assess treatment integrity, the fact that ODRs are used as the primary dependent variable in the CICO literature, the lack of discussion of challenges in implementation, and the method for determining the criterion for reinforcement. Below, evaluations of CICO and the BEP are presented, and their limitations noted. Collectively, these interventions will be referred to as CICO.
In an early investigation, March and Horner (2002) investigated whether student performance on CICO varied based on the hypothesized function of the students’ problem behavior. Two analyses were conducted. First, a descriptive analysis was conducted that involved the application of CICO with no modification based on the function of the problem behavior. The non-responders then participated in an experimental analysis that included a function-based (based on hypothesized function), individualized intervention that was implemented in an attempt to decrease problem behaviors.

Twenty-four middle school students, 20 males and 4 females, exhibiting problem behaviors participated in CICO. CICO had been in place for four years prior to the study’s initiation. Participants were selected if they were nominated by a teacher or parent or had received at least five ODRs within a semester. CICO served to clarify the student’s behavior expectations, increase their routine throughout the day, structure more adult praise, and increase home-school communication. Prior to implementation, a contract was signed by the students, where they agreed to improve behavior. In addition to March and Horner (2002), Crone et al. (2004) indicated that the use of a contract within CICO is beneficial, although this component has not been consistently used across CICO evaluations. Daily, check-in occurred at the office, where the students received social recognition and a CICO form. Throughout the day, teachers rated the student’s compliance with school-wide expectations. Check-out occurred in the main office when the student returned the CICO form. At this time, the student was provided social recognition and a small edible for having each class rated by his or her teachers. Most often, students are rewarded for meeting their point goal (Filter et al., 2007; Hawken,
2006; Hawken et al., 2007, Hawken & Horner, 2003; McCurdy et al., 2007), so having
the contingency tied to the number of ratings represents a different contingency for
reward than is seen in later CICO studies. Although not explicitly stated, it is likely that
this modification was implemented in order to increase the occurrence of students
obtaining all their teacher’s ratings. Also during check-out, a copy was sent home to be
signed by the parent, and a copy was retained at school.

Multiple dependent variables were examined in the study. The primary dependent
variable was the number of discipline contacts, which represented detentions and/or
ODRs for problem behavior. The functions of the problem behaviors were assessed to
determine whether access to peer attention, adult attention, preferred activities tangibles,
or escape from social engagement or academic tasks maintained the problem behaviors.
The Functional Assessment Checklist for Teachers and Staff (FACTS) (March et al.,
2000) was used to determine the hypothesized function of the children’s problem
behaviors and although typically recommended to be used in conjunction with direct
observations (Horner et al., 1999; O’Neill et al., 1997), March and Horner (2002) did not
include direct observations in their assessment of function. The FACTS has been shown
to be strongly correlated to direct observation data and moderately correlated to
functional analysis data; however, this is based on one study that examined existing data
(McIntosh et al., 2008) and additional evaluation of the FACTS measure is merited.
March and Horner (2002) monitored fidelity of implementation through direct
observation and weekly meetings with the teachers; however, no quantifiable measures of
fidelity were gleaned from these interactions. The researchers noted this as a limitation. A
review of permanent products was also conducted to determine whether check-in
occurred, whether the student presented the form to the teacher and received feedback, and if the student returned the form with parent signature; however, these data were not presented.

Preliminary analyses of the results were conducted to determine if the intervention was effective (change in the rate of discipline contacts post-implementation). However, this definition is problematic because even if a change in rate is evidenced it may not represent a meaningful behavior change. It was determined that 5 of the 24 students’ problem behaviors were hypothesized to be maintained by access to adult attention, eight by access to peer attention, and eleven by escape from academic demands. Eighty percent (4 of 5) of those students with problem behaviors maintained by access to adult attention improved with CICO. For students with problem behaviors hypothesized to be maintained by access to peer attention, 62.5% (5 of 8) improved. Only 27% (3 of 11) of the students with problem behaviors maintained by escape from academic demands showed improvements. It is also possible that this discrepancy in response to intervention is related to the students’ intervention targets not being correlated with those behaviors resulting in discipline contacts; however, this information was not included in the study. Also not explicitly evaluated, this lack of response by students with escape-maintained problem behaviors may also be attributed to the student’s inability to perform the academic task due to an academic skill deficit. For this reason, future research should explore whether students with problem behaviors hypothesized to be maintained by escape should be evaluated to determine whether skill or performance deficits exist. Although performance deficits can be addressed through
standard CICO procedures, academic skill deficits likely will require additional skill building in conjunction with CICO to achieve positive outcomes.

Overall, students whose behavior was hypothesized to be maintained by access to adult and peer attention showed the greatest behavior improvements with CICO. In contrast, levels of problem behaviors increased by 50% or more post-implementation for 10 students. Of these 10 students, 7 of the 10 students exhibited problem behaviors maintained by escape, suggesting that these students may be the least responsive to CICO. As described previously, this could be due to these seven students possessing academic skill deficits, but this was not evaluated. Suggesting that for some students, regardless of the function of their problem behavior, CICO in its standard form may not sufficiently address the behavior concerns. For students who do not respond to CICO, more intense or different interventions (i.e., Tier 3 interventions), or perhaps more likely, additional skill building, may be merited.

The second portion of the study was an experimental analysis implemented with the three non-responders from the previous analysis. This marks the transition from targeted interventions to more intensive, individualized interventions. Students were deemed “non-responders” and included in the experimental analysis if the student (a) evidenced no decrease in ODRs following initial CICO implementation; (b) received at least five ODRs within the first four months of the new academic year; (c) was nominated by the CICO team; and (d) assented, along with the parent’s consent. The participants in the experimental analysis included Andy, a 13-year-old seventh grader exhibiting disruptive classroom behavior and physical aggression who received special education services due to academic delays; Bill, a 13-year-old seventh grader with issues
related to defiance and insubordination; and Cathy, a 12-year-old sixth grader who exhibited disruptive and aggressive behavior and received special education services for academic delays.

Partial interval observations were conducted three to five times per week in two of each student’s classrooms and included the target students as well as control peers. This study represents one of a few studies (Campbell & Anderson, 2008; Fairbanks et al., 2007; Hawken & Horner, 2003; Hawken et al., 2007; Todd et al., 2008) of CICO in which direct observations of participant behavior were conducted. Observed behaviors included each student’s problem behaviors as well as engagement, which were defined as the student attending to or looking at the task materials or teacher and problem behaviors were operationalized for each student. The percentages of intervals containing problem behavior and engagement were evaluated. In addition, the teachers’ perceptions of feasibility and utility were recorded for all five teachers through use of a rating scale.

In baseline, variable levels of problem behavior were observed; Andy exhibited a mean of 46% (range 22%-63%), 37% (range 8%-75%) for Bill, and 30% (range 5%-82%) for Cathy. Following implementation of the function-based intervention, problem behavior fell to 14%, 16%, and 16% of intervals, respectively, as compared to control students (13%, 8%, and 10%, respectively). In baseline, academic engagement levels were determined to be 38%, 34%, and 38% for Andy, Bill, and Cathy, respectively. Similarly, academic engagement rose following implementation of the function-based intervention to 68%, 73%, and 65%, respectively. Control peers in each of the three student’s classes exhibited academic engagement similar to the target students, at 74%, 85%, and 76% respectively.
In another early investigation of CICO, Hawken and Horner (2003), using a multiple baseline across subjects design, implemented CICO to increase academic engagement and decrease problem behaviors of four middle school students. Direct observations were conducted to determine disruptive behavior and academic engagement levels for both target students and control peers. Use of direct observations, as well as the inclusion of control peers, is similar to the second portion of March and Horner (2002), yet, sets this study apart from the majority of other CICO investigations. The point goal was arbitrarily set at 80%, which was unrelated to the participant’s baseline level of performance. Weekly, the student’s progress was evaluated by the CICO team. It was determined that the progress was sufficient if the student was earning 80% or more of the possible daily points.

Across all students, decreases in the mean level and variability of problem behavior, as well as increases in the mean level of academic engagement were evidenced. Therefore, Hawken and Horner (2003) demonstrated that CICO functioned as an effective intervention while preserving resources. Additionally, due to the structured adult feedback that is critical to CICO, there was an increased likelihood of adult praise contingent on the student’s appropriate behavior. The authors recommended further investigation to determine methods to improve parental participation in CICO, given its mediocre level of occurrence. The need for the parent component is unclear, as it typically is implemented with the lowest level of integrity (Hawken 2006, Hawken et al., 2007), yet gains in participant behavior continue to occur.

In another example, Hawken (2006) urges school psychologists to serve as systems change agents, particularly with students at risk for engaging in more severe
behavior. To that end, Hawken recommended that CICO be implemented for these students and examined the effectiveness of the program while also monitoring and evaluating treatment integrity. Ten students, nine males and one female, from a rural middle school in the Pacific Northwest participated in the study. The problem behaviors included talking without permission, making inappropriate comments, poor work completion, unpreparedness for class, and failure to keep hands, feet, and objects to self. The students did not exhibit severe behaviors such as aggression, property destruction, or self-injurious behavior. SWPBS was present in the school, and CICO had been implemented for five years. Participants (a) had entered CICO after a minimum of two months of school (to allow for baseline data collection); (b) had received a minimum of five ODRs; (c) were nominated by staff to receive additional supports; and (d) received CICO for at least six weeks. The authors did not address why students already participating in CICO were chosen for participation; however, it may be due to time constraints. Because the students were participating in CICO when the study began, the results must be interpreted cautiously.

The average number of ODRs per week pre- and post-intervention served as the primary dependent variable. For each participant, baseline data were collected for at least eight weeks (range = 8 - 19 weeks), which Hawken (2006) suggested was to allow sufficient time to establish a rate of ODRs, and post-implementation data were collected for at least eight weeks (range = 8 - 23 weeks). In addition, the participants each received intervention for at least six weeks. This suggests that the authors were most interested in the maintenance effects of CICO on students after they were no longer receiving intervention, as data were not presented to reflect the students’ performance during
CICO. However, an informative addition would have been to include the DBRC data, as this would have provided formative information as to the effectiveness of the study. Instead, only summative data of the intervention’s effectiveness were provided, and this is of limited utility in reaching a decision of program effectiveness. No modifications were made to the CICO that was in place at the school prior to the study. The fidelity of implementation was assessed on three randomly selected days of intervention. Because the intervention was in place for at least six weeks, this may be an insufficient number of integrity checks. The following components were assessed for fidelity: (a) checked-in; (b) brought the DBRC to all teachers and obtained feedback; (c) checked out; (d) returned the signed DBRC from the parent; and (e) whether the CICO coordinator recorded the data for that day. Again, fidelity checks did not begin until the study began, which was at least six weeks following the implementation of CICO for the participants. Therefore, the validity of the results must be questioned given that no evaluation of procedural integrity occurred for at least the first six weeks of implementation, and following that time, was assessed on only three days.

Seven of the students improved on CICO, but not all the students, a finding that is supported by previous research (Filter et al., 2007; Hawken et al., 2007) indicating that approximately 65%-75% of students can be expected to respond to CICO when function is not assessed and addressed specifically (Hawken, O'Neill, & MacLeod, 2011). However, the function-based research related to CICO has relied almost exclusively on hypothesized function as assessed through use of the FACTS and, therefore, this relationship should be tempered. Finally, one of the non-responders exhibited an increase in ODRs post-intervention. For these students who do not respond to less intensive
targeted interventions (e.g., CICO), Hawken (2006) recommends use of more intensive interventions as was implemented for the non-responders in March and Horner (2002). Based on review of the integrity data, it was indicated that CICO was implemented with a high level of fidelity, with the exception of the parental feedback component. However, due to the nature of the integrity assessments, these data must be interpreted cautiously. Hawken indicates there is no set level of treatment integrity necessary to produce behavior change, but suggests that the fidelity level in the current study was sufficient to produce positive effects. Despite the behavior gains, Hawken noted that the quasi-experimental design was a limitation, as it is unclear whether other variables may have been responsible for the effects evidenced. In addition, Hawken suggested that despite the reductions in ODRs, reductions in classroom disruptive behavior or increases in academic performance may not have occurred. Nonetheless, Hawken (2006) echoed the findings of Hawken and Horner (2003), suggesting that CICO serves as an efficient and effective intervention. Hawken also noted that teachers and staff can implement CICO with integrity; however, the data presented do not fully support this claim.

Hawken et al. (2007) conducted another study in which the effect of CICO, which was currently in place in an urban elementary school, was evaluated. The effects of CICO on the frequency of ODRs were assessed, as were treatment fidelity and social validity. The elementary school (Grades K-6) of 655 students had been implementing SWPBS for over three years with 88% fidelity as evidenced by the Schoolwide Evaluation Tool (SET) (Horner et al., 2004). The participants (a) had entered CICO after a minimum of two months of school (to allow for baseline data collection); (b) had received a minimum of two ODRs; (c) were nominated by staff to receive additional supports; (d) had received
CICO for at least six weeks; and (c) had exhibited problem behavior across the entire school day and during various activities. As in Hawken (2006), the authors did not explain why baseline data collection occurred for an extended period of time or why the students were receiving intervention prior to the initiation of the study. However, they suggested that due to the smaller number of ODRs earned, aggregated reports across longer periods of time (i.e., monthly) are often required to provide representative data, and likely serves as their rationale for the extended baseline period (Hawken et al., 2007).

Twelve students, ten boys and two girls, were selected for participation. Eight of the participants were receiving free or reduced lunch, and one student was receiving special education services for a learning disability in reading. The participants’ problem behaviors included talking without permission, poor work completion, making inappropriate comments, and failing to keep hands, feet, and objects to self. None exhibited severe problem behavior.

Hawken et al. (2007) assessed fidelity of implementation was assessed as described by Hawken (2006) across three randomly selected days of CICO. Data were evaluated by the CICO coordinator and used for decision-making. Mean fidelity scores for each student for every component were determined. As noted in previous studies, these assessments were not initiated until the study began, therefore, no fidelity assessments were conducted during early CICO implementation. Further, after initiation of the study, fidelity was assessed on only three days. Despite the researcher’s efforts to ascertain levels of treatment integrity, the methods used may be too limited to provide an accurate representation of the level of fidelity of implementation. As in Hawken (2006), the validity of the results is threatened by this procedural limitation.
The primary dependent variable was problem behavior, defined as the total number of ODRs per group of three students per month. Students were placed in groups based on when they entered CICO. Specifically, students who entered within one month of each other were in the same group, which resulted in four groups of three students. These groupings made data analysis and interpretation more difficult. Nonetheless, minor and major referrals were combined for each group of students and evaluated for each month prior to implementation and following implementation of CICO. The researchers also assessed social validity using the five-item CICO Acceptability Questionnaire (Hawken & Horner, 2003).

The effectiveness of CICO was assessed using a multiple baseline design across groups of students. Typical school-wide behavior supports were provided to all students and were in place during baseline. For the participants, CICO was implemented and involved students checking in with a paraprofessional before school where they were provided with their DBRC and evaluated on their preparedness for class (had the required materials). In a different addition from previous studies, praise and a lottery ticket for the weekly drawing were provided for checking in. Again, this suggests that CICO at this school had been modified in order to increase student’s compliance with checking in. Daily goals were also identified and students were provided feedback on these during check-in. During natural transitions in the school day, the students’ behavior expectations were evaluated on their DBRC by their teachers. The students returned their DBRC to the paraprofessional for check-out at the end of the school day. During check-out, daily percentages were calculated, and praise and rewards using a spinner system were provided for meeting the daily goal. The spinner system also represents an addition to the
standard procedures of CICO for delivering rewards. Daily goals for all students were set at 80% of total points per day. This goal was set arbitrarily and represents a limitation to the current study. The behavior support team met bi-monthly to evaluate student progress, which they defined as the students meeting their 80% daily goal.

CICO was implemented with high fidelity (i.e., mean greater than 90% of components implemented), which was the average of all components present across all students for the sample of days analyzed. As is typically seen, the family review and feedback component was the worst implemented, with an average of 36% implementation across the fidelity checks. Therefore, questions exist as to whether this is a critical component to intervention success and represents an area of future research. Mean total ODRs per month decreased across all groups, suggesting a decrease in problem behavior. Specifically, reductions of 51%, 46%, 36%, and 25% in ODRs per month from baseline to CICO implementation were seen across the groups. For individual students, changes from pre- to post-CICO were presented. Nine of the twelve students received fewer ODRs post intervention, with the remaining three students exhibiting increases in ODRs post intervention. The intervention was rated as acceptable by staff, parents, and students. The lowest rating was provided by parents for the interventions’ improvement of academic performance at 4.5 on the 6-point scale.

Overall, Hawken et al. (2007) suggested that CICO can be implemented in a typical school setting by school personnel, although the level of integrity at which typical school personnel implement CICO is unclear. Despite the gains evidenced, the study is limited in that ODRs were the only measure of problem behavior. As mentioned previously, the inclusion of students who were already receiving CICO makes the
validity of the outcomes questionable. Also, problems with interpreting the results were due to the grouping of students and averaging of ODRs across the month. This grouping added no additional clarity to the study. Instead, it made interpretation of individual student’s outcomes more challenging. Also, only a small number of data points in each phase were provided due to the averaging the ODRs per month, with only two data points provided for the fourth group. Therefore, analysis of the intervention’s effects was made challenging due to numerous errors made in the presentation of data.

In a post-hoc analysis of CICO, Filter et al. (2007) evaluated the effectiveness of CICO in three elementary schools using a quasi-experimental design. In order to evaluate the implementation of CICO in a natural school context by typical school personnel, the active training and on-going support of researchers was removed by examining CICO post-implementation. Across the schools, implementation of CICO occurred within a RTI model that included SWPBS. Nineteen students were selected for participation. Problem behavior as measured through major ODRs, minor ODRs, and a combination of both types of ODRs, served as the dependent variable. The inclusion of minor ODRs is unusual, as minor behaviors typically do not result in ODRs. However, these schools maintained separate databases for major (i.e., defiance, aggression, vandalism) and minor (i.e., rule violation, minor inappropriate language, failure to complete an assignment) ODRs. Results indicated that 67% of students experienced decreases in combined ODRs from pre-CICO to post-CICO. The students averaged one ODR (combination of major and minor ODRs) every 5.59 days prior to implementation and one every 8.47 days following intervention.
Treatment integrity and the perceived effectiveness and efficiency also served as the outcome measures. Using a 5-item integrity checklist, treatment integrity was assessed. The components evaluated were those assessed in Hawken (2006) and Hawken et al. (2007) with the addition that the data were graphed for analysis. However, a permanent product review was not conducted. Instead, a survey was completed to determine in general whether or not the steps were occurring at their school. Information specific to the fidelity of implementation for a given participant or CICO component was not known. The survey integrity data suggested that checking in and checking out occurred across the three schools. Another component that was reported to have consistently occurred was the feedback provided to students, as indicated by ninety-four percent of respondents (i.e., 16 of 17). Therefore, survey responses indicated that the three initial components were being implemented fairly consistently. Eighty-two percent (i.e., 14 of 17) of respondents indicated that data were used for on-going data-based decision making. However, only 41% (i.e., 7 of 17) of respondents indicated that a family member was reviewing the DBRC.

Although the focus of this study was to assess the fidelity of implementation by typical school personnel, the methods used to assess treatment integrity fail to provide an accurate assessment of fidelity. Instead, the survey used provides only the adult participants’ perceptions of which components occurred in general. Therefore, which components were occurring, how often they were occurring, and for which student they were occurring is unknown. Therefore, one is unable to determine the role CICO played in the changes evidenced in ODRs. Perceived effectiveness and efficiency measures were presented and were also completed using a Likert measure and the program was rated as
generally effective and efficient. Although it is more common for acceptability and efficiency measures to be obtained through surveys, these remain only adult participants’ perceptions of acceptability and efficiency of the intervention. Goal setting was not discussed.

In a more recent evaluation of CICO, McCurdy et al. (2007) implemented CICO and evaluated the DBRC point data of eight students at an urban northeastern elementary school. Only limited results were presented given the case study format, so three featured participants included a Caucasian male fourth grade student diagnosed with Attention Deficit-Hyperactivity Disorder (ADHD), a Caucasian female fifth grade student receiving special education services, and an African American male fifth grade student. A female special education teacher was responsible for CICO implementation, due to her willingness and approach to interacting with the students participating in the study. Procedures were implemented in standard CICO format with the inclusion of minor added features. One of these included the check of homework completion, which occurred during check-in. During check-out, the student participated in graphing his or daily percentage of points. Also, in addition to daily reward, additional rewards were provided for students earning 100% of their daily points over a period of time.

The dependent variables included the daily point percentages, which has been a long overlooked source of data. This sets the current study apart from other studies that typically monitor only ODRs as the student outcome variable (Hawken, 2006; Hawken et al., 2007). Across each participant’s data, the authors determined which students achieved successful, moderately successful, and unsuccessful outcomes. The authors stated that half of the participants (4 of 8) achieved successful outcomes, two achieved
moderately successful outcomes, and two achieved unsuccessful outcomes. No method of determining which participants obtained successful, moderately successful, or unsuccessful outcomes was described. However, a graphical representation of the participants’ data was provided which represented each outcome. Through examination of the visual representations of the three students daily point percentages, a successful outcome was represented by an immediate increase in level of a large magnitude from baseline to intervention, which remained higher than baseline levels throughout intervention. According to the authors, a moderately successful outcome was initially variable, and then remained near criterion level for a few weeks before deteriorating to within baseline levels. However, considering intervention data within baseline levels as moderately successful seems questionable. Finally, an immediate decrease in percentage of points earned upon implementation of intervention that remained lower than baseline levels throughout intervention was considered an unsuccessful outcome. Other problems in the study included the arbitrary nature of pre-set conventional criterion for reinforcement (i.e., 80% of points), the lack of treatment integrity data, and failure to note limitations or challenges in implementation.

Todd et al. (2008) implemented CICO in a rural elementary school in the Pacific Northwest. Four elementary-age boys were selected for participation and included Trevor, a Native American boy in third grade, Chad, a Caucasian boy in the first grade, Kendell, an African American boy in second grade, and Eric a Caucasian boy in kindergarten, in addition to seven teachers and three staff members. A multiple baseline across participants design was used to evaluate the effectiveness of CICO for these students.
Prior to CICO implementation, the researchers conducted FBAs for each student that included teacher interviews and direct observations. The assessment information indicated that the hypothesized function of the problem behavior for all students was access to adult attention, which researchers often suggest results in the best CICO outcomes (Crone et al., 2004). CICO was implemented as described previously, with another modification to reward delivery (McCurdy et al., 2007) that allowed students to spend points earned for appropriate behavior to purchase tangible items, activities, or privileges.

During baseline, the students engaged in a variable but unacceptable level of problem behavior. Problem intervals were 30%, 26%, 34%, and 27% for Trevor, Chad, Kendall, and Eric, respectively. A reduction occurred across all participants, falling to 14%, 8%, 13%, and 12%, respectively, when CICO was implemented. Across all participants, mean baseline to mean intervention levels of problem behavior decreased by 17.5% (range = 15% - 20%). Chad began ADHD medication on the first day of intervention without the researchers knowing, and therefore, his outcome cannot be exclusively attributed to CICO.

Todd et al. (2008) extended the CICO literature by including direct observation data; however, data from the DBRC again was not presented. The researchers indicated that there was a record of whether students checked in, checked out, and returned their forms signed, but no formal assessments of treatment integrity were reported. Although the researchers attempted to loosely monitor fidelity of implementation, these data should have been presented and more formal attempts at integrity assessments should have been
made. Despite these limitations, this study supports the idea that CICO may be a viable intervention for students with problem behavior maintained by access to adult attention.

Harpole (2010) sought to extend the CICO literature by addressing a number of limitations evident in the CICO literature. Namely, Harpole sought to address a primary limitation by including the DBRC data as the primary dependent variable. Other limitations addressed included examining the effectiveness of CICO with high school students, conducting a more thorough evaluation of treatment integrity, and using a performance-based point criterion for reinforcement. A multiple baseline design was used to evaluate the effects of CICO on the participants’ ratings of appropriate behavior and disruptive behavior.

Three African American high school students in the ninth, twelfth, and eleventh grades served as the participants. They had no previous diagnoses or special education rulings. The participants attended two moderately sized schools (range = 617 students – 957 students). Demographics across the schools ranged from 70% white with 29% African American to 92% African American and 7% White. Both schools had high levels of students receiving free or reduced lunch (range = 73% - 80%), suggesting the largely low socioeconomic status (SES) of the students, which has been shown to serve as an obstacle to treatment (Lavigne et al., 2010).

Prior to implementation, the three participants had each received numerous ODRs for reasons including continued minor behaviors (e.g., off task, talking without permission), defiance, tardiness, and skipping class. Each student nominated a staff member who they wanted to serve as their CICO coordinator. One participant’s CICO coordinator was a second year junior high science teacher, who held a bachelor’s degree
in biology and a master’s degree in education. The other students shared a CICO coordinator who was the TST Coordinator and held a master’s degree in education. In addition, the selected coordinators each nominated an alternate coordinator who would be responsible for carrying out the coordinator’s duties if he or she were absent.

CICO was implemented as described previously with the addition of a performance-based point criterion. Instead of using a pre-set conventional criterion for reinforcement, Harpole (2010) set the student’s initial point criterion at their baseline median. This point criterion was re-evaluated weekly and either increased to the current week’s median or kept the same if the median was lower than during the previous week. The primary dependent variable, teacher ratings of appropriate behavior as noted by percentage of daily points earned, served as an aggregate of each student’s three replacement behaviors on each student’s DBRC. The second dependent variable was disruptive behavior, which was measured through frequency of ODRs.

During baseline, mean levels of teacher ratings of appropriate behavior were 34% (range = 20%-50%), 72% (range = 40%-89%), and 55% (range = 22%-81%) across the three participants. Increases in appropriate behavior from baseline means to CICO means were 41%, 19%, and 24% for the students. Effect size estimates provided the level of impact of the results using methodology described by Parker and Hagan-Burke (2007). The students’ scores suggested that the odds or likelihood of improvement in the CICO phase was 34, 5, and 2.78 times of that during baseline. Percentage of All Non-Overlapping Data was 90.48%, 75%, 56.25%, and 73.58%. Disruptive behavior also decreased for all students by 0.8, 1.6, and 1.3 ODRs per month (i.e., prorated to account for a different number of days per month. Greater decreases in ODRs may have been
evidenced, but floor effects likely occurred as two participants received zero ODRs during CICO implementation.

Treatment integrity was assessed using permanent products data for all days of treatment, which sets this study apart from others who assess treatment integrity a limited number of days of implementation. Across the students, overall treatment integrity, average daily treatment integrity, median daily treatment integrity, and the integrity of the five key components was assessed. These key components included (a) check-in; (b) DBRC provided; (c) teacher ratings; (d) check-out; and (e) parent signature on returned DBRC. Overall implementation was 81%, 72%, and 81% for the participants. The most poorly implemented component across all students was having all teachers rate the students’ DBRCs. Despite the lower levels of treatment integrity evidenced, CICO implementation at this level with the obstacles of intervention implementation in a high school in addition to the low SES of the students and families highlights the utility of CICO.

Recently published evaluations of CICO are beginning to empirically evaluate previously noted hypotheses found within the CICO literature. For example, McIntosh, Campbell, Carter, and Dickey (2009) conducted a study that followed up on an idea introduced by March and Horner (2002) years earlier—does the function of problem behavior play a role in a child’s improvement or lack of improvement during CICO implementation? This idea is commonly presented in CICO literature; however, few researchers have empirically evaluated whether function of problem behavior relates to success or failure with CICO. The study was conducted with 34 participants across six public elementary schools in the Pacific Northwest, with all schools having SWPBIS in
place at above 80% mean implementation on the School-wide Evaluation Tool (Sugai, Lewis-Palmer, Todd, & Horner, 2001). Five of the six schools had CICO interventions in place at the time of the study.

Prior to CICO implementation, teachers completed the FACTS to determine the hypothesized function of the participants’ problem behavior. As the function of the participants’ problem behavior was paramount in evaluating the current study, FACTS data could have been supplemented with direct observation data and conditional probabilities analysis. Teachers also completed the Behavior Assessment Scale for Children – Second Edition (BASC-2) (Reynolds & Kamphaus, 2004) prior to CICO implementation in order to determine norm-referenced levels of problem behavior and adaptive behavior. The Behavioral Symptoms Index and Adaptive Scale of the BASC-2 in addition to ODRs served as dependent measures and were statistically analyzed using a mixed model multivariate analysis of variance repeated across time (i.e., pre and post intervention). Fidelity of implementation was randomly assessed across nine days for multiple students using a checklist of ten critical items. Fidelity evaluations suggested that CICO was implemented with 94% accuracy (range = 83-100%) across observations.

Results of McIntosh et al. (2009) suggest that CICO implementation resulted in statistically significant improvements in ratings of problem behavior, ratings of prosocial behavior, and ODRs for children hypothesized to have attention-maintained problem behavior. On the other hand, students with hypothesized escape-maintained problem behavior did not achieve statistically significant improvements on any measures. In fact, participants with escape-maintained problem behavior exhibited increases in ratings of problem behavior from pre- to post-implementation of CICO, suggesting that function of
problem behavior may serve as a moderator of the effects of CICO (McIntosh et al., 2009).

Further extending the CICO literature, researchers have now begun to evaluate CICO using more stringent research designs and evaluation of CICO component analyses. For example, Campbell and Anderson (2011) noted various limitations within the CICO literature including the limited number of CICO studies, the quasi-experimental nature of these investigations, the over-reliance of ODR data to determine CICO effectiveness, the absence of direct observations of participant behavior, the use of single subject methodology in all but one study, the absence of withdrawal or reversal of treatment, and the absence of component analysis investigating the elements within CICO that may be responsible for producing behavior change. Therefore, Campbell and Anderson (2011) sought to replicate previous CICO findings within the context of a more rigorous experimental design while also assessing the contribution of teacher feedback within CICO.

Participants included four elementary age males who attended a suburban elementary school in the Pacific Northwest. The school had SWPBS in place for approximately five years, and the researchers noted that it was fully in place according to performance on the SET. Functional assessments were conducted initially for all participants. These assessments entailed completion of the FACTS and three 10 minute 5 second partial interval observations to assess problem behavior, adult attention, task avoidance, and peer attention. CICO procedures involved a morning check-in with the CICO coordinator, various feedback meetings with the student’s teacher, an afternoon meeting with the coordinator, and home feedback provided by a parent/guardian.
Criterion for reinforcement was set at 80% of points. As the teacher feedback portion was the focus of the current study, feedback meetings were systematically removed. Initially, the noon feedback meeting was removed, followed by the morning feedback session, and finally, the afternoon feedback session was removed for all participants but one. For the remaining participants, one feedback meeting was later reinstated. Ratings of participant behavior (i.e., to determine points earned) were not completed prior to feedback meetings that no longer took place, therefore, possible points only included times that were rated and discussed during the respective feedback session. Dependent measures included problem behavior and academic engagement assessed via direct observations, the percentage of points earned, and social validity (i.e., teacher’s perceptions of student problem behavior compared to peers) and contextual fit (i.e., effort required to manage the student’s behavior compared to peers) were assessed twice weekly for teachers using a 2-item questionnaire. Fidelity of implementation was assessed across 27% of CICO days by experimenters who observed check-in, check-out, and classroom feedback sessions to evaluate integrity according to a 12-item checklist of CICO key features. Fidelity was determined to be 97% (range = 83-100%) across observations.

Results of the functional assessment suggest that all participants exhibited problem behavior hypothesized to be maintained primarily by access to adult attention. Decreases in problem behavior demonstrated the effectiveness of CICO as compared to baseline and withdrawal phases. Decreases in problem behavior persisted for all students when two and one feedback meetings occurred daily, with increases in problem behavior occurring when zero feedback meetings occurred. For academic engagement, gains were
observed during CICO implementation and persisted when two and one feedback meetings occurred, but decreased when no feedback meetings occurred and during baseline and withdrawal phases. Social validity measures suggested that teachers perceived the students as having problem behavior similar to that of peers during baseline with improvements in problem behavior made during CICO implementation that persisted as feedback sessions were faded to two and one, but that increased slightly when no feedback sessions occurred. Contextual fit measures suggested that adult participants perceived their level of effort in managing student behavior to decrease upon CICO implementation and decrease further when feedback meetings were faded to twice and once daily, with perceived effort increasing when feedback meetings were removed altogether. These results suggest that teacher feedback sessions and ratings of behavior are a critical component of CICO, but that the frequency of daily meetings required to maintain gains in student behavior may be less than was previously believed. It is unclear whether similar results would have been evidenced in regard to CICO involving one feedback meeting daily had CICO with three feedback meetings daily not preceded the less intensive form of CICO.

As recent studies are beginning to answer more questions regarding CICO, Simonsen, Myers, and Briere (2011) further extended the CICO literature by evaluating the effectiveness of CICO using a group design involving an experimental and control group. The researchers indicated that although previous researchers have demonstrated the effectiveness of CICO using single subject methodology, experimental group designs have not been used. Additionally, no studies have evaluated the effectiveness of CICO as compared to a control intervention. Therefore, the current study employed an
experimental group design to evaluate the effectiveness of CICO as compared to standard protocol (i.e., counseling) on decreasing off-task behavior and other problem behavior evaluated using an indirect measure (i.e., Social Skills Rating System, SSRS) (Gresham & Elliot, 1990).

The study was conducted in an urban middle school in New England where SWPBS was in place (i.e., 98% fidelity observed on the SET). Participants were randomly assigned to either group (i.e., CICO or counseling). Twenty-seven students were assigned to the CICO group and 15 were assigned to the counseling group. It is unclear why there was a large discrepancy in group size. Standardized CICO (Crone et al., 2004) was administered, and counseling was administered as typically provided by the school’s three counselors. The counseling was provided weekly in group format by the school counselor to groups of children categorized by perceived need (e.g., anger management, grief counseling) and the skill areas noted by the school counselors to have been taught included anger management, social skills, and positive choices. Dependent measures evaluated included the FACTS, direct observation of off-task behavior, the SSRS, and social validity measures. The FACTS was administered to each student’s teacher prior to intervention to identify behaviors of concern and hypothesized functions of those behaviors (e.g., access to attention, escape/avoidance of task). Twenty minute 10 second partial interval observations were conducted three to five times per student per condition to determine target students’ levels of off-task behavior. The SSRS was completed by each student’s teacher before and after intervention implementation to obtain teacher’s perceptions of student problem behavior, social skills, and academic competence. Additionally, statistical analyses were conducted. Treatment integrity was
assessed across three days for check-in and check-out critical components only.

Acceptability measures were also completed by adult participants.

Results suggest that most students (i.e., 68.4%) were hypothesized by their teachers to have problem behavior maintained by multiple functions—92.1% were rated as having problem behavior maintained by access to peer attention, 55.3% for escape or avoidance of tasks, 42.1% to gain access to adult attention, and 21.1% to escape attention (e.g., escape adult reprimands, negative peer attention). Although no statistically significant differences were found across groups in regard to hypothesized functions of behavior, it should be noted that more students believed to have problem behavior maintained by escape or avoidance of tasks were assigned to the counseling group. As previous research has indicated, CICO has been shown to be less effective for students with problem behavior maintained by escape or avoidance of tasks (McIntosh et al., 2009), therefore gains evidenced in CICO may have been less had more of the students with problem behaviors maintained by escape or avoidance of tasks been in the CICO group.

Direct observation data suggested that students assigned to the CICO group engaged in higher levels of off-task behavior prior to intervention, and made larger decreases in off-task behavior during CICO although participants assigned to the counseling group also exhibited decreases in off-task behavior during intervention. One-way ANOVA results suggested that gain scores (i.e., decreases in recorded levels of off-task behavior) for CICO and counseling groups were statistically significant, F(1, 40) – 10.41, p < .01. SSRS across all subscales was analyzed using a one-way ANOVA and suggested that no changes were statistically significant. However, effect sizes for
problem behavior, social skills, and academic competence according to SSRS scores were small for both groups but favored the counseling group. Treatment integrity evaluations suggested that school staff and interns completed 92.6% (range = 88.9% - 100.0%) of steps as intended. Additionally, it was noted that the primary researcher continued to monitor fidelity and periodically assisted during check-in or check-out to model accurate implementation or help with staff shortages. Acceptability measures completed by teachers suggest that statistically significant differences were not evidenced for adult participants assigned to either group, but that ratings suggested that adult participants assigned to CICO rated CICO more favorably than those of the counseling group. Therefore, results suggest that both groups evidenced gains during intervention, with direct measures highlighting greater improvements for participants receiving CICO than those receiving counseling.

Despite the gains evidenced across these studies, limitations exist within the CICO literature. Treatment integrity (Gresham, 1989) or fidelity of implementation represents the extent to which components of an intervention are implemented as intended and represents a critical component in evaluating the effects of an intervention. However, in the CICO literature, a number of researchers have reported the effectiveness of CICO programs that were already in place in schools (Filter et al., 2007; Hawken, 2006; Hawken et al., 2007), which prevented systematic treatment integrity data from being collected throughout CICO implementation. Filter et al. evaluated treatment integrity using only a rating scale completed by the adult participants as to whether they perceived each component to have occurred across most of CICO implementation. Other researchers have taken advantage of the permanent product data to evaluate integrity
(Hawken, 2006; Hawken et al., 2007; Hawken & Horner, 2003; March & Horner, 2002). However, the limited number of fidelity checks did not begin until the study began, with products were reviewed for only three days of CICO implementation. Therefore, no procedural integrity evaluations occurred during the early weeks of implementation and for only a small sample of intervention days thereafter. Although other researchers have evaluated treatment integrity within the entire CICO implementation period, few integrity evaluations were conducted (i.e., three days assessed across six week CICO implementation, Simonsen et al., 2011). Finally, Todd et al., 2008 failed to present any fidelity data, although the researchers indicated that they had data as to whether students checked in, checked out, and returned the signed form daily. Across these studies, it is evident that greater care must be taken to evaluate treatment integrity accurately and thoroughly in order to effectively assess the effectiveness of CICO.

Additionally, the majority of CICO researchers have ignored a wealth of student behavioral outcome data by failing to include students’ daily point percentages as a dependent variable. Instead, the literature has focused primarily on ODR data or other indirect measures (i.e., BASC-2), as a means of evaluating CICO effectiveness (Filter et al., 2007; Hawken, 2006; Hawken et al., 2007; McIntosh et al., 2009). ODR data are indirect in that they are reliant on teacher report and thereby teacher perception, which can often be inconsistent and vary substantially across teachers. ODR data are also limited in that they are removed from time and place as ODRs are often completion some time after to the noted behavioral incident took place. Finally, ODRs are then interpreted by another staff member who enters them into the school’s database.
On the other hand, the DBRC data are sensitive to more subtle student behavior changes and can detect performance differences across different classes. Few studies (e.g., Harpole, 2010; McCurdy et al., 2007) have presented student daily point percentages as the primary dependent variable. However, McCurdy et al. provided graphical representations of the CICO point data for only three of the five participants. McCurdy et al. indicated these three graphs were representative of the three possible outcomes (i.e., successful, moderately successful, unsuccessful); however, no criteria were presented to explain objectively how these outcomes were categorized. Despite the benefits of DBRC data as compared to ODR data, it is unclear how closely DBRC data align with direct observation data. DBRC data are also rated by teachers, sharing with ODRs a reliance on teacher perception. Similarly, DBRC data is rated at completion of a set rating interval and so is also delayed, although to a lesser degree than ODRs. Early research suggests overlap; however, the body of evidence is limited and requires continued examination. Based on these concerns, these measures serve as complementary procedures and should be used in conjunction with one another as they were in the current study.

Additionally, challenges surface in terms of intervention implementation in applied settings, and this remains true for CICO. Harpole presented these concerns as well as possible solutions to address these problems. More researchers should describe the challenges incurred in implementation in order to aid practitioners in future CICO implementation and determine areas of CICO requiring further evaluation. The final limitation represents the method of criterion setting typically implemented in evaluations of CICO and the concerns with this methodology will be described below.
Criterion for Reinforcement

Determining an appropriate criterion for reinforcement, or criterion for reward, is critical to students’ success with CICO, yet researchers have continued to arbitrarily set the criterion at 80% of points earned on the DBRC, regardless of the student’s baseline level of performance. Although this may be an effective method, no data have been evaluated to suggest this as the optimal method of goal setting. In fact, researchers more often suggest that performance-based criteria are more effective (Cooper, Heron, & Heward, 2007; Heward, 1980). For this reason, the current study will evaluate whether improvements in student behavior are increased with the use of performance-based versus pre-set conventional criteria for reinforcement.

Before a discussion of criteria for reinforcement can occur, reinforcement and reward must be differentiated. Reinforcement and reward are often referred to synonymously, but represent different terms worth distinguishing (Cameron & Pierce, 1994). Reinforcement represents the process of presenting or removing a stimulus immediately following the occurrence of a behavior that increases the future frequency of that behavior in similar conditions (Cooper et al., 2007). A reinforcer, therefore, is a stimulus that when presented following a behavior, increases the future display of that behavior in similar conditions (Cooper et al., 2007). On the other hand, a reward is a stimulus that is assumed to be positive, but is not defined by its effect on behavior (Cameron & Pierce, 1994). Therefore, by definition, reinforcement, at times through the presentation of a reward following a desired behavior, changes the future occurrence of that desired behavior.
For reinforcement to be effective there must be some criteria in place to determine whether or not the reinforcer has been earned and should be presented—a criterion. A criterion, in and of itself, has been shown to have an effect on behavior (Mizes, 1984). In a study by Mizes, college students performed better on a manual wheel turning task when there was a criterion in place, whether it was a self- or experimenter-imposed criterion, and performance further improved when there was a reward provided contingent on meeting the criterion. Despite the utility of setting a criterion, Cooper et al. (2007) indicated that setting the criterion for reinforcement at too high a level initially represents a common mistake in reinforcement programs. The problem with a pre-set conventional (i.e., 80%) criterion for reinforcement, as is typically seen in CICO studies, is that this may be too high for the student to attain early in intervention. Therefore, the student will be unable to access reinforcement. Without the student accessing reinforcement for exhibiting the desired level of appropriate behavior, the contingency between the desired behavior and the reinforcer cannot be formed (Cooper et al., 2007). This contingency serves to increase the frequency of the student performing the desired behavior.

Cooper et al. (2007) recommends that the initial criterion for reinforcement should be one that will allow the student’s first attempts to access the reinforcer. For example, in CICO, the criterion for reinforcement, or point goal, should be set low enough that the student can access the daily reward the first few days that CICO is in place. This criterion should then be gradually increased as the student’s performance improves (Cooper et al., 2007). Heward (1980) proposed a method for determining the initial criteria for reinforcement that is related to the student’s level of performance during baseline. Specifically, Heward suggested that for behaviors that are to be
increased, the initial criterion should be (a) higher than the child’s mean baseline performance and (b) lower than or equal to the highest level of baseline performance. Conversely, Heward recommended that behaviors that are being targeted to decrease, the initial criterion should be (a) below the mean level of baseline performance, and (b) greater than or equal to the child’s lowest (best) baseline performance.

Summary and Rationale

Collectively, the studies described previously suggest that CICO may be an efficient and effective targeted intervention for students at-risk of developing more severe behaviors. It is suggested based on the review of the study data that most students will benefit at some level from CICO, others will not benefit, and some may worsen over the course of CICO implementation. However, despite the gains evidenced across the CICO studies, additional research is required to address specific limitations in the literature. These limitations include (a) the overreliance on ODR data; while DBRC data have been overlooked; (b) the limited and/or inaccurate evaluations of treatment integrity; (c) the lack of description of challenges incurred during implementation and means to address these; (d) the lack of direct observations of student behavior, and (e) overuse of pre-set conventional criterion for reinforcement. To address these limitations and extend the literature on CICO, the current study will focus on determining the method of criterion-setting (i.e., performance-based or pre-set conventional) that produces greater gains in student behavior as well as which produces a higher rate of change. In addition to this, the current study will implement procedures similar to Harpole (2010) in order to address previous limitations by including DBRC data as the primary measure of effectiveness and by reviewing permanent products for all days of CICO implementation to evaluate
treatment integrity. The work of Harpole will be further extended through inclusion of direct observations of student behavior and for the assessment of treatment integrity.

Purpose

CICO serves as a promising intervention; however, a number of questions remain unanswered. The present study sought to address limitations within the literature with procedures similar to those described by Harpole (2010) by evaluating treatment integrity appropriately and across CICO implementation, and through presentation and discussion DBRC data. In addition, an extension of Harpole was conducted through inclusion of direct observation data of target students and control peers’ appropriate behavior. However, the primary focus of the present study was to address whether a performance-based or a pre-set conventional criterion for reinforcement produces higher levels and a higher rate of change in teacher reported appropriate behavior and other dependent measures. It was hypothesized that a performance-based criterion would result in greater gains in student performance. Combined, the present study serves to enhance and extend the CICO literature.
CHAPTER II
METHOD

Setting and Participants

The study was conducted in three participating elementary schools in southern Mississippi. The three schools were all in a county school district of approximately 2,500 students. Three participants attended School A, which served approximately 640 students in grades K through 8. For School A, 67% of students were White, 30% were African American, 1% were Hispanic, and 1% were Asian. Two participants attended School B, which served approximately 171 students in grades K through 6. For School B, 53% of the students were African American, 42% were White, 3% were Hispanic, and 1% were Asian. Three participants attended School C, which served approximately 667 students in grades K through 8. For School C, 91% of students were White, 7% were African American, and 1% were Hispanic. Free or reduced lunch rates across the schools were 64%, 85%, and 66% for Schools A, B, and C, respectively. All schools had School-wide Positive Behavior Interventions and Support (SWPBIS) in place for approximately four years with School-wide Evaluation Tool scores for the 2010-2011 school year of 87.6%, 91.1% (i.e., PBIS model site), and 94.6% (i.e., PBIS model site) for Schools A, B, and C, respectively.

Eight general education students in first through fifth grade with behavior problems warranting intervention were selected for participation. The participants were required to (a) have been recommended by a teacher or administrator due to presenting social behavior concerns; (b) have received problematic rates of office discipline referrals (ODRs) as indicated by a staff member; (c) exhibit problem behaviors that did not
include serious, dangerous or infrequently occurring behaviors, and (d) obtain parent/guardian consent (see Appendix A). Inclusion criteria also required that participating students have a mean level of teacher ratings of appropriate behavior less than 70% following baseline data collection. Participant demographics (i.e., grade, school, race), referral concerns, and criterion for reinforcement (i.e., Performance-Based or Pre-Set Conventional) are provided below in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Grade</th>
<th>School</th>
<th>Race</th>
<th>Referral Concerns</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chris</td>
<td>1</td>
<td>A</td>
<td>W</td>
<td>OT, IV, OOS</td>
</tr>
<tr>
<td>Cameron</td>
<td>1</td>
<td>B</td>
<td>B</td>
<td>IV, OOS</td>
</tr>
<tr>
<td>Bree</td>
<td>2</td>
<td>A</td>
<td>W</td>
<td>OT, IV, OOS</td>
</tr>
<tr>
<td>Samantha</td>
<td>2</td>
<td>B</td>
<td>B</td>
<td>OT, PWO</td>
</tr>
<tr>
<td>Kasey</td>
<td>3</td>
<td>C</td>
<td>H</td>
<td>IV, HFI, PPI</td>
</tr>
<tr>
<td>Nate</td>
<td>4</td>
<td>A</td>
<td>W</td>
<td>OT, IV, OOS</td>
</tr>
<tr>
<td>Rachel</td>
<td>5</td>
<td>C</td>
<td>W</td>
<td>OT, PPI</td>
</tr>
<tr>
<td>Nolan</td>
<td>5</td>
<td>B</td>
<td>W</td>
<td>OT, IV, OOS</td>
</tr>
</tbody>
</table>

Note. B = Black, H = Hispanic, W = White, OT = off task, IV = inappropriate vocalizations, OOS = out of seat, HFI = handling frustration inappropriately, PPI = poor peer interactions, PWO = playing with objects

The target students’ teachers also served as participants in the current study, and their informed consent was obtained as well (see Appendix B). For all participants, one of their teachers served as their Check In-Check Out (CICO) Coordinator. The CICO
coordinators were all certified general education teachers with a mean of 10 years of teaching experience (range = 2-15 years). Two of the CICO coordinators held at least one master’s degree, with one CICO coordinator currently enrolled in a Master’s program, and the remaining CICO coordinators possessing bachelor’s degrees.

Dependent Measures

Multiple behavioral outcomes were evaluated in the present study. The primary dependent variable, teacher ratings of appropriate behavior, represents an aggregate of each student’s three alternative behaviors as rated by their teachers on each student’s Daily Behavior Report Card (DBRC) and is represented as the percentage of daily points earned. Teacher ratings of appropriate behavior as rated on the DBRC were selected as the primary dependent variable as this measured the students’ behavior across all classes daily. Although direct observations may provide a higher level of accuracy, direct observations require a high level of resources (e.g., observers, time) and so could not be conducted in all classes each day in the present study. Therefore, direct observation data may not provide a comprehensive picture of the students’ behaviors across all settings and times in the school day. Teacher ratings of appropriate behavior were evaluated by determining the points earned on the student’s DBRC divided by the total points possible that day and multiplying by 100. Larger percentages indicate greater levels of appropriate behavior. Alternative behaviors for the participants included the following: On task, talked with permission only, in seat, keeping hands still, positive peer interactions, and handled frustration well. On task is defined as actively (e.g., writing, raising hand, answering a question, talking about lesson) or passively (e.g., eyes directed at task or teacher, without breaking contact for 3 seconds consecutively) participating in the
classroom activity. *Talked with permission* was defined as speaking to peers, teacher, or self when instructed or required to by the academic task, unless given permission by the teacher after appropriate request or following teacher initiation. *In seat* was defined as student’s buttocks in contact with the designated seat/area, without breaking contact for more than 3 seconds consecutively. *Keeping hands still* was defined as the student making contact with only objects that are required for the task at hand and using required materials in the manner intended (e.g., writing with pencils). *Good peer interactions* was defined as the student engaging in positive interactions with peers as evidenced by engaging in positive verbal (e.g., saying thank you and please, using appropriate tone of voice, polite response to peer request/comment) and social exchanges (e.g., sharing items). *Handled frustration well* was defined as the student accepting not being called on by teacher and not getting to do as he or she would like without getting frustrated. Frustration included the child hitting objects (e.g., books, desks), hitting objects (e.g., desk) with other objects (e.g., book), hitting self (e.g., contact with self by hand or object), or crying.

Dependent measures assessed via direct observation included academically engaged behavior (AEB) and problem behavior. Twenty-minute observations were conducted weekly for each target student during the time/class period identified by his or her teacher as the most problematic. Trained graduate students with advanced graduate training in behavioral assessment and previously trained to 90% interobserver agreement on the data collection system served as the observers. AEB and problem behavior were observed simultaneously and AEB was scored using 10 second whole interval coding while problem behavior was scored using 10 second partial interval coding using an
observation form (see Appendix C). The decision to code AEB using whole interval and problem behavior using partial interval served to provide a conservative estimate of both the students AEB and problem behavior. AEB was scored if the child was engaged in any combination of the following behaviors: (a) looking at the teacher during instruction, (b) working with a peer when instructed to do so, (c) reading silently or writing to complete assignments when instructed to do so, (d) participating in a teacher-approved activity following the completion of work, or (e) talking with the teacher about academic work. In order for the interval to be coded as AEB, the student was required to exhibit AEB during the entire 10 second interval. The percentage of intervals of AEB was calculated by dividing the number of intervals scored as AEB by the number of intervals possible multiplied by 100. Higher percentages suggested higher levels of appropriate behavior. Problem behavior represented an aggregate measure consisting of each student’s two to three problem behaviors, as identified by his teacher(s). As problem behavior was coded using 10 second partial interval, the interval was scored as problem behavior having occurred if the child exhibited any one of his problem behaviors (e.g., talking without permission) at any point during the 10 second interval. The number of intervals scored as problem behavior divided by the number of possible intervals multiplied by 100 was calculated to determine a percentage of problem behavior. Higher percentages suggested higher levels of problem behavior. As described previously, problem behaviors for the target students included off-task, inappropriate vocalizations, out of seat, playing with objects, poor peer interactions, and handling frustration inappropriately. Off-task was defined as the student breaking eye contact with the task materials or the teacher for more than 3 seconds. Inappropriate vocalizations were
defined as any audible vocalization (e.g., talking, humming, singing) made by the student that was not preceded by teacher permission to speak or initiated by the teacher. *Out of seat* was defined as the student’s buttocks breaking contact with the seat/designate area for more than 3 seconds. *Playing with objects* was defined as manipulating materials or objects that were not related to the task at hand. *Poor peer interactions* was defined as the child making inappropriate vocalizations to peers (e.g., calling names, cursing, commanding directions to peers), failing/refusing to share items, and making inappropriate contact with peers (e.g., pushing, pulling, grabbing). *Handling frustration inappropriately* was defined as the student exhibiting signs of frustration verbally (e.g., sighing loudly, grunting, crying), through facial expressions (e.g., pushing bottom lip out, frowning), hitting objects (e.g., books, desks), hitting objects (e.g., desk) with other objects (e.g., book), hitting self (e.g., contact with self by hand or object). Three control peers were observed as well to provide an aggregate control peer score of appropriate behavior for comparison purposes. Observation intervals were alternated between the target student and each of the control peers.

Disruptive behavior was assessed by monitoring the frequency of ODRs. ODRs often serve as dependent measures due to their availability and sensitivity in measuring the effects of interventions (Irvin, Tobin, Sprague, Sugai, & Vincent, 2004). Additionally, Putnam, Luiselli, Handler, and Jefferson (2003) indicated that ODRs can be used to evaluate outcomes and aid staff in making data-based decisions about interventions’ short term and long term effectiveness. Despite these strengths, research on ODRs has not been extensive, and additional research in the areas of ODR reliability and validity is merited, but not within the scope of this study. Nonetheless, ODRs serve as a common metric for
evaluating behavior across the SWPBS and CICO literature, and therefore, it was used in this study in conjunction with more specific measures.

Materials

A Daily Behavior Report Card (DBRC) was used (see Appendix D), which included the student’s behavioral expectations written across the top of the form with the different class periods as the rows. Assignments and written teacher praise statement could also be written on the DBRC. Boxes for the teachers to initial after rating the student’s behavior as 2, 1, or 0 points were delineated. A 2 was awarded for meeting the behavioral expectation; a 1 was awarded for partial demonstration of the behavioral expectation, and a 0 was awarded for failure to meet a particular behavioral expectation (Crone et al., 2004). At the bottom of the form, the total points earned for the day was written by the CICO coordinator. A line was provided for parent signature as well.

Design

Two single subject ABAB designs were used to assess the effects of criterion setting method on CICO outcomes. Four participants were randomly assigned to receive the Pre-Set Conventional criterion and four participants were randomly assigned to receive the Performance-Based criterion. Participants with the Pre-Set Conventional criterion received point goals (i.e., criteria for reinforcement) associated with their CICO that were 80% of their daily points possible. In other words, these students had to earn 80% of their possible daily points in order to access reinforcement at the end of the day during Check-Out with their CICO coordinator. The participants with the Performance-Based criterion had point goals (i.e., criteria for reinforcement) that were initially based on their baseline performance and that were increased gradually.
Procedures

Selection of Problem and Replacement Behaviors

To determine each student’s problem and replacement behaviors, the primary investigator reviewed each student’s records (e.g., ODRs, behavior logs) and consulted with respective teachers. Problem behaviors were observed during the direct observations, and the replacement behaviors were the behavior expectations noted on each student’s DBRC. The primary investigator operationally defined these behaviors in consultation with the student’s teacher(s). Operational definitions were provided to adult participants and observers during their respective trainings and to students prior to CICO implementation.

Teacher and Coordinator Training

Prior to the initiation of each phase, the primary investigator trained the teachers and CICO coordinator(s) on the required procedures associated with each phase of the study. The training consisted of the primary investigator describing the procedures, modeling the procedures, and having the staff member practice the procedures with the primary investigator providing feedback. This format was implemented until the individuals were able to implement the procedures independent of the primary investigator’s assistance. Re-training for a specific CICO component was provided to the adult participants responsible for the component after the component was not implemented for three consecutive days. Re-training had to be provided to three of the eight CICO coordinators, and was conducted in the manner of training described previously. Re-training was provided for conducting check-out, calculating percentage of points earned, and providing precorrections at the beginning of class periods.
Baseline

The participating teachers completed blind DBRCs for each participating student rating the students’ behaviors on the DBRCs without the students’ knowledge, such that the students were blind to the fact they were being observed and rated. Although the target students’ parents already provided consent for their child to participate, specific details about the intervention and the timeline of implementation was not provided to the parent(s) in order to decrease the likelihood that the child’s performance during baseline was altered. Additionally, baseline data collection began at least five days after the parent/guardian consent was obtained. Additional details were provided to the parent prior to CICO implementation and are described below. At the end of a class period, the teacher rated the target student’s behavior as a 2, 1, or 0 on the student’s DBRC. Data were collected on teacher ratings of appropriate behavior for at least three days during baseline, but more were collected if stability in performance was not established. In this way, a baseline of their percentage of teacher ratings of appropriate behavior was established. Additionally, at least three direct observations of each target student during the baseline condition were conducted to determine a baseline level of AEB and problem behavior. Additionally, the mean and median number of ODRs during baseline was determined.

CICO

The CICO coordinator and participating teachers were trained by the primary investigator on the CICO procedures using the training format previously described. The CICO procedures were as follows:
Setting the DBRC Point Goal. Participants with the Performance-Based criterion for reinforcement had point goals based on their baseline performance, and these were gradually increased across CICO implementation. In other words, their initial point goal was set by determining a criterion that was higher than the student’s baseline mean of teacher ratings of appropriate behavior and lower than the student’s highest teacher rating of appropriate behavior during baseline. Each week thereafter, the criterion for reinforcement was re-evaluated in that manner. For example, at the end of the first week of CICO, the criterion for reinforcement was set higher than the first week’s mean but lower than the first week’s highest teacher rating of appropriate behavior (Heward, 1980). For participants with the Pre-Set Conventional criterion for reinforcement, their point goal remained at 80% throughout CICO implementation.

Check-in. The CICO coordinator conducted morning check-in, which occurred in the prior to the beginning of the first class period or shortly after class beginning, as most participants’ CICO coordinators were one of their primary teachers. Check-in activities include the coordinator (a) collecting the signed DBRC parent copy from the previous day, when applicable; (b) providing the student with a DBRC to take to classes; and (c) reminding the student of his behavioral expectations. These procedures were outlined on the Check In-Check Out Procedures (see Appendix E) form that was provided to CICO coordinators.

Teacher Feedback in Each Class. At the beginning of a new class period, the teacher prepared the student for the class by reminding him or her of their expectations. At the end of each class, the teacher rated the target student’s behavior as a 2, 1, or 0 on the student’s DBRC. Then, the teacher briefly explained his ratings to the student and
provided praise for meeting behavioral goals and/or corrective feedback for behavior
goals that could be improved. These procedures were outlined on the top of each
student’s DBRC (see Appendix D).

*Check-out.* Check-out occurred approximately five to ten minutes prior to the end
of the day. During check-out, the Coordinator determined the percentage of points earned
by the student on his DBRC, delivered praise, delivered corrective feedback (if
necessary), determined whether the student met his daily goal, provided the student with
an incentive (if goal was met), and sent the DBRC home for parent review. These
procedures were outlined on the Check In-Check Out Procedures form (see Appendix E)
that was provided to CICO coordinators.

*Student and Parent Orientation to CICO.* After the CICO coordinator and
teachers were trained on their roles, the primary investigator met with each target student
separately to train him or her in the intervention process (see Appendix F). The primary
investigator spoke with each student’s parent via telephone prior to CICO
implementation to inform the parent of the CICO features described to their child. At this
time, the parent was provided details about CICO that could not be provided earlier as to
not interfere with his or her child’s baseline performance. Topics discussed with the
target student and parent included: (a) the purpose of CICO; (b) the student’s goal
percentage of points; (c) the CICO coordinator’s role in CICO, (d) the teachers’ role in
CICO; (e) how points are scored on the DBRC; (f) how participating students will know
if they are meeting their point goals; (g) how to check in; (h) how to check out; (i) what
would be provided contingent on meeting point goals; (j) how to give the DBRC to each
teacher; and (k) how to get parent/guardian signature and return DBRC the next day.
(Crone et al., 2004). The primary investigator also completed the CICO Contract (see Appendix G; Crone et al.) with the student, which outlined the target behaviors and preferred rewards.

*CICO Implementation.* After all participants provided permission for participation and were trained on the procedures, all parties were informed that CICO implementation was going to begin. Data were collected from target students’ DBRCs to determine levels of teacher ratings of appropriate behavior. Direct observations were conducted at least once weekly for each target student at the pre-determined time (i.e., most problematic time period in school day) to monitor levels of AEB and problem behavior. The monthly rate of ODRs for each target student was monitored to assess disruptive behavior.

Data Analysis

For each method of data analysis, individual outcomes were assessed as well as group (i.e., Performance-Based, Pre-Set Conventional) outcomes. Each target student’s level of teacher ratings of appropriate behavior, AEB, and problem behavior was graphed and visually inspected (Kazdin, 1982, 1984) for all conditions. Level, trend, and variability were examined to determine the effects of the intervention on each student’s behavior. The rate of ODRs was analyzed for each student.

Additionally, quantitative measures of effect size were analyzed in order to demonstrate the level of impact of the intervention outcomes (Borckardt et al., 2008). Percentage of All Non-Overlapping Data (PAND) and odds ratios were calculated for each participant and for both groups. PAND is a non-parametric statistic that represents a calculation of the total number of data points that do not overlap between baseline and intervention phases (Parker et al., 2005; Parker & Hagan-Burke, 2007). PAND is
calculated by dividing the number of overlapping data points by the number of total data points, multiplying by 100, and subtracting that percentage from 100. Odds ratios are effect size estimates that were included in order to demonstrate the level of impact of the results using methodology described by Parker and Hagan-Burke (2007). Odds ratios are calculated by dividing the odds ratio for intervention (i.e., number of points in intervention that do not overlap with baseline data points divided by the number of data points that do overlap) by the odds ratio for baseline (i.e., number of data points that do overlap with intervention data points divided by the number of data points that do not overlap).

Treatment Integrity

Treatment integrity (Gresham, 1989) was evaluated through direct observations and permanent product review. Treatment integrity checks were conducted through direct observation during a mean of 29% (range = 23% - 31%) of morning check-ins, a mean of 27% (range = 23% - 38%) of afternoon check-outs, and 26% (range = 23% - 31%) of the designated class period’s feedback sessions. Morning check-in procedures involved four components (i.e., check-in occurred, DBRC from previous day collected, new DBRC provided, student precorrected about his behavioral expectations). Afternoon check-out procedures involved seven steps (i.e., check-out occurred, praise provided, corrective feedback provided if necessary, percentage of points earned calculated, determination of whether student met point goal, incentive provided if earned, DBRC sent home for parent review). Class procedures involved four steps (i.e., precorrections provided to student about behavior expectations, ratings explained, praise given, and corrective feedback provided if necessary).
Permanent products (i.e., DBRCs) were reviewed across all days of CICO implementation by the primary investigator. Using the DBRCs, the treatment integrity of check-in procedures was evaluated by determining whether a new DBRC was provided. Permanent products were evaluated to determine the treatment integrity of check-out by examining whether check-out occurred, whether percentage of points earned was calculated, and whether parent signature was obtained. Class procedures evaluated included whether each teacher rated the student. Therefore, if a student had five teachers and was rated by only four teachers, permanent products would suggest 80% treatment integrity for classroom procedures for that student’s day’s implementation of CICO.

Treatment integrity was measured as the number of components that occurred as intended divided by the number of possible components and multiplied by 100. Daily treatment integrity was calculated by dividing the number of components completed in a day by the number of possible components and multiplying the total by 100. Mean daily treatment integrity was calculated by totaling the daily treatment integrity scores for a given participant and dividing that amount by the total days of CICO implementation and multiplying the total by 100. Additionally, mean treatment integrity was evaluated for each portion of the participant’s day (i.e., check-in, check-out, class) for observed treatment integrity evaluations by totaling the treatment integrity scores for that portion of the day for a given participant and dividing that amount by the total days of CICO implementation and multiplying the total by 100.

Interobserver Agreement

Interobserver agreement (IOA) was assessed for direct observations of student behavior as well as for treatment integrity. Reliability was assessed across at least 20% of
observations through use of a second trained observer. IOA was calculated by dividing the number of agreements by the number of agreements plus disagreements multiplied by 100.

Acceptability

Treatment acceptability measures were issued to adult and child participants in order to determine their perceptions of the acceptability, feasibility, and perceived effectiveness of CICO. The Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985) is a commonly used measure to evaluate the acceptability of school-based behavioral interventions. The IRP-15 (see Appendix H) was completed by CICO coordinators and teacher participants. Respondents rated the 15 items according to their agreement or disagreement with the statement using a 6-point Likert scale (i.e., 1 = strongly disagree, 6 = strongly agree). Item responses are totaled to glean a general acceptability score, which ranges from 15 to 90, with larger numbers representing higher levels of treatment acceptability. Martens, Witt, Elliott, and Darveaux (1985) suggests that the IRP-15 has excellent internal consistency (alpha coefficient = .98). The IRP-15 was modified for the current assessment as items were changed to past tense to allow for raters to complete the scale post-CICO implementation. Freer and Watson (1999) reported that this modification does not negatively affect the psychometric properties of the IRP-15.

The Children’s Intervention Rating Profile (CIRP) (Witt & Elliott, 1985) is a treatment acceptability measure for children. Children answer seven items using a 6-point Likert scale. The CIRP items assess the fairness (e.g., The program used was fair.), expected effectiveness (e.g., Being in this program helped me do better in school.), and
potential adverse effects (e.g., Being in this program caused problems with my friends.) as related to the intervention. Ratings of 1 (i.e., I definitely disagree) suggested the student agreed with the item with ratings of 6 (i.e., I definitely disagree.) suggesting that the student disagreed with the statement. Items 2, 3, and 4 were reverse scored (i.e., higher scores suggested greater acceptability). Ratings are totaled producing a total treatment acceptability score, ranging from 7 to 42 with lower scores representing higher levels of acceptability. The current study used an adapted version of the CIRP (see Appendix I) in order to clarify the ratings (e.g., I definitely agree, I somewhat agree, I agree) for younger children as the original CIRP was shown to have a fifth grade readability (Rinderknecht, 2006). Additionally, the primary investigator read the items aloud to younger children. It is unclear how these modifications may have affected the psychometric properties of the CIRP.

Institutional Approval

The University of Southern Mississippi Institutional Review Board reviewed the current project in accordance with the University of Southern Mississippi Human Subjects Protection Review Committee in accordance with the Federal Drug Administration regulations, Department of Health and Human Services, and university guidelines to ensure protection of human participants. All procedures and materials used in the current study were approved by the University of Southern Mississippi Institutional Review Board per the Insititutional Review Board Notice of Committee Action (see Appendix J).
CHAPTER III

RESULTS

Teacher Ratings of Appropriate Behavior

The current study sought to evaluate the effect of two methods of determining criterion of reinforcement on various student outcomes. For teacher ratings of appropriate behavior, mean overall gains from baseline to the second Check In-Check Out (CICO) phase were 37% (range = 27-45%) for the Pre-Set Conventional participants and 19% (range = 12-32%) for the Performance-Based participants. Percentages of teacher ratings of appropriate behavior were graphed for each student with the Pre-Set Conventional criterion for reinforcement across phases and are provided in Figure 1 with graphs provided for the Performance-Based participants in Figure 2.
Figure 1. Percentage of Teacher Ratings of Appropriate Behavior Daily for Each Participant with Pre-Set Conventional Criteria for Reinforcement.
Figure 2. Percentage of Teacher Ratings of Appropriate Behavior Daily for Each Participant with Performance-Based Criteria for Reinforcement.
During baseline, mean levels of teacher ratings of appropriate behavior were 47% (range = 37-62%) for the Pre-Set Conventional participants and 45% (range = 15-69%) for the Performance-Based participants. When CICO was implemented, immediate increases were evidenced across participants receiving both methods of goal setting. Mean levels during CICO were 78% (range = 72-91%) for the Pre-Set Conventional participants and 73% (range = 39-93%) for the Performance-Based participants. Average increases from baseline to CICO were 32% (range = 28-35%) for the Pre-Set Conventional participants and 28% (range = 19-36%) for the Performance-Based participants. When CICO was withdrawn, immediate decreases in teacher ratings of appropriate behavior were seen for all participants with the exception of Rachel. Mean levels of teacher ratings of appropriate behavior were 60% (range = 34-74%) for the Pre-Set Conventional participants and 46 (range = 7-78%) for the Performance-Based participants. When CICO was re-implemented, increases were again evidenced for most participants, but not all participants. Mean levels of teacher ratings of appropriate behavior were 83% (range = 70-91%) for the Pre-Set Conventional participants and 64% (range = 27-89%) for the Performance-Based participants. Mean increases from the Withdrawal phase to the second CICO phase were 23% (range = 1-55%) for the Pre-Set Conventional participants and 18% (range = 7-32%) for the Performance-Based participants. Individual participant means, ranges, and magnitudes of change across phase for participants and by criterion for reinforcement method are provided in Table 2 for the Pre-Set Conventional participants and in Table 3 for the Performance-Based participants.
Table 2

Means, Ranges, and Magnitude of Change for Percentage of Teacher Ratings of Appropriate Behavior across Phases for Participants in Pre-Set Conventional Participants

<table>
<thead>
<tr>
<th></th>
<th>Nolan</th>
<th>Samantha</th>
<th>Kasey</th>
<th>Chris</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL Mean</td>
<td>62</td>
<td>37</td>
<td>50</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td>BL Range</td>
<td>53-71</td>
<td>17-50</td>
<td>25-69</td>
<td>32-43</td>
<td>17-71</td>
</tr>
<tr>
<td>CICO1 Mean</td>
<td>91</td>
<td>72</td>
<td>78</td>
<td>72</td>
<td>78</td>
</tr>
<tr>
<td>CICO1 Range</td>
<td>80-100</td>
<td>50-100</td>
<td>27-100</td>
<td>24-98</td>
<td>24-100</td>
</tr>
<tr>
<td>WD Mean</td>
<td>34</td>
<td>64</td>
<td>74</td>
<td>69</td>
<td>60</td>
</tr>
<tr>
<td>WD Range</td>
<td>20-50</td>
<td>60-66</td>
<td>44-94</td>
<td>38-89</td>
<td>20-94</td>
</tr>
<tr>
<td>CICO2 Mean</td>
<td>89</td>
<td>82</td>
<td>91</td>
<td>70</td>
<td>83</td>
</tr>
<tr>
<td>CICO2 Range</td>
<td>80-100</td>
<td>60-96</td>
<td>78-98</td>
<td>33-86</td>
<td>33-100</td>
</tr>
<tr>
<td>Δ BL to CICO1</td>
<td>29</td>
<td>35</td>
<td>28</td>
<td>35</td>
<td>32</td>
</tr>
<tr>
<td>Δ BL to CICO2</td>
<td>27</td>
<td>45</td>
<td>41</td>
<td>33</td>
<td>37</td>
</tr>
</tbody>
</table>

Note. BL = baseline, CICO = Check In-Check Out, WD = withdrawal, Δ = Magnitude of Change
### Table 3

*Means, Ranges, and Magnitude of Change for Percentage of Teacher Ratings of Appropriate Behavior across Phases for Participants in Performance-Based Participants*

<table>
<thead>
<tr>
<th></th>
<th>Cameron</th>
<th>Bree</th>
<th>Nate</th>
<th>Rachel</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BL Mean</strong></td>
<td>15</td>
<td>38</td>
<td>69</td>
<td>57</td>
<td>45</td>
</tr>
<tr>
<td><strong>BL Range</strong></td>
<td>6-22</td>
<td>30-42</td>
<td>61-79</td>
<td>44-67</td>
<td>6-79</td>
</tr>
<tr>
<td><strong>CICO1 Mean</strong></td>
<td>39</td>
<td>71</td>
<td>88</td>
<td>93</td>
<td>73</td>
</tr>
<tr>
<td><strong>CICO1 Range</strong></td>
<td>19-70</td>
<td>22-100</td>
<td>75-100</td>
<td>85-100</td>
<td>19-100</td>
</tr>
<tr>
<td><strong>WD Mean</strong></td>
<td>7</td>
<td>50</td>
<td>49</td>
<td>78</td>
<td>46</td>
</tr>
<tr>
<td><strong>WD Range</strong></td>
<td>5-8</td>
<td>36-64</td>
<td>21-75</td>
<td>65-85</td>
<td>5-85</td>
</tr>
<tr>
<td><strong>CICO2 Mean</strong></td>
<td>27</td>
<td>57</td>
<td>81</td>
<td>89</td>
<td>64</td>
</tr>
<tr>
<td><strong>CICO2 Range</strong></td>
<td>17-48</td>
<td>44-64</td>
<td>67-92</td>
<td>70-100</td>
<td>17-100</td>
</tr>
<tr>
<td>Δ BL to CICO1</td>
<td>24</td>
<td>33</td>
<td>88</td>
<td>93</td>
<td>28</td>
</tr>
<tr>
<td>Δ BL to CICO2</td>
<td>12</td>
<td>19</td>
<td>12</td>
<td>32</td>
<td>19</td>
</tr>
</tbody>
</table>

Note: BL = baseline, CICO = Check In-Check Out, WD = withdrawal, Δ = Magnitude of Change

**PAND**

PAND (Parker et al., 2007) was analyzed in order to provide additional objective means of data interpretation (Parker et al., 2007). Mean PAND was 78% (range = 68-100%) for the Pre-Set Conventional participants and 81% (range = 80-83%) for the Performance-Based participants. Therefore, overall PAND was higher for the Performance-Based participants; however, Nolan (Pre-Set Conventional criterion) had the highest level of PAND across all participants at 100%.
**Odds Ratios of Improvement**

Effect size measures for each participant were calculated. The Odds Ratio for the Pre-Set Conventional participants is greater than 35.3 (range = 1.6-132), which suggests that the odds or likelihood of improvement in CICO are 35 times that of the baseline phase. An exact odds ratio could not be calculated as one participant of the Pre-Set participants, Nolan, had zero overlapping data points. Therefore, an odds ratio could not be calculated for him. However, if he had one overlapping datum, his odds ratio would have been 132, suggesting that the odds of improvement were 132 times more likely to have occurred in CICO than in baseline. Despite the large odds ratio evidenced by Nolan, which increased the collective odds ratio for all participants receiving the Pre-Set Conventional criterion for reinforcement, other participants with the Pre-Set criterion for reinforcement exhibited much smaller odds ratios than those exhibited by the Performance-Based participants. The odds ratios calculated for the other participants with the Pre-Set Conventional criterion for reinforcement (i.e., Samantha, Kelsey, and Chris), aside from Nolan, would have been 6. For the Pre-Set Conventional participants, Nolan’s odds ratio was >132, Chris’ was 5, Kasey’s was 2.7, and Samantha’s was 1.6. On the other hand, the Performance-Based participants’ odds ratios were 17.5 for Nate, 15 for Bree, 12 for Rachel, and 3.5 for Cameron, with a group odds ratio of 12 (range = 3.5-17.5). Across both methods of criterion for reinforcement, odds ratio data suggests a robust effect for CICO.

**Observed Child Participant Outcomes**

Direct observations were used to assess academically engaged behavior (AEB) and problem behavior of target students and control peers. Direct observations were
conducted during 45% of study days (range = 87% of baseline, 26% for CICO, 40% for withdrawal, 27% for CICO) with interobserver agreement (IOA) assessed for 60% of observations. IOA was 95% (range = 94-99%) across participants.

Academically Engaged Behavior

AEB was assessed via direct observation for all participants as well as control peers. For the Pre-Set Conventional participants, mean AEB was 49% (range = 18-72%) for baseline, 79% (range = 49-98%) for CICO, 58% (range = 3-56%) for withdrawal, and 81% (range 62-100%) when CICO was re-implemented. Therefore, overall gains from baseline to CICO were 32% (range = 21-47%) for the Pre-Set Conventional participants.

For the Performance-Based participants, mean AEB was 47% (range = 16-73%) for baseline, 80% (range = 33-97%) for CICO, 70% (range = 56-91%) for withdrawal, and 78% (range = 65-97%) when CICO was re-implemented. Mean overall gains for the Performance-Based participants were 31% (range = 8-73%) from baseline to CICO. AEB levels for control peers across observations conducted with all participants ranged between 71% and 87%.

Pre-Set Conventional Participants

Across participants with the Pre-Set Conventional criterion for reinforcement, gains in AEB were exhibited in three of four participants. These gains increased target students’ levels of AEB above those exhibited by control peers. The effect of CICO on AEB for participants with the Pre-Set Conventional criterion for reinforcement is further discussed below.

Samantha. Mean levels of AEB for Samantha were 59% for baseline, 72% for CICO, 56% during the withdrawal, and 80% when CICO was re-implemented, for an
overall change from baseline to CICO of 21%. During baseline, Samantha’s AEB was variable and lower than that exhibited by control peers. Samantha’s AEB exhibited a decreasing trend during the first CICO phase, with levels remaining within CICO levels during the withdrawal, and with increasing levels of AEB occurring when CICO was re-implemented to levels higher than those observed by peers.

*Nolan.* Mean levels of AEB for Nolan were 48% for baseline, 88% for CICO, 29% during the withdrawal, and 74% when CICO was re-implemented, for an overall change from baseline to CICO of 26%. When CICO was implemented with Nolan, immediate increases in AEB were evidenced. When CICO was removed, immediate decreases in AEB occurred, with an increase of large magnitude occurring upon re-implementation of CICO. Although Nolan exhibited levels of AEB lower than control peers across all phases, levels of AEB were more similar to those evidenced by control peers when CICO was implemented.

*Kelsey.* Mean levels of AEB for Kelsey were 44% for baseline, 81% for CICO, 81% during the withdrawal, and 78% when CICO was re-implemented, for an overall change from baseline to CICO of 34%. When CICO was implemented for Kelsey, an immediate change was observed with levels of AEB remaining above baseline levels, but exhibiting a decreasing trend. When CICO was removed, AEB levels remained within CICO range and decreased further during the re-implementation of CICO phase to levels higher than those observed in control peers.

*Chris.* Mean levels of AEB for Chris were 45% for baseline, 73% for CICO, 67% during the withdrawal, and 92% when CICO was re-implemented, for an overall change from baseline to CICO of 47%. When CICO was implemented for Chris, an increase in
AEB was evidenced with levels of AEB increasing and remaining above baseline levels throughout the CICO phase and within range of those exhibited by control peers. During withdrawal, levels of AEB remained within CICO levels but were below levels observed in peers, but when CICO was re-implemented, immediate gains were evidenced with levels of AEB increasing across observations to levels higher than those observed in control peers. Graphs for the Pre-Set Conventional participants and control peers are provided in Figure 3 below.
Figure 3. Academically Engaged Behavior Across Phases for Participants in the Pre-Set Conventional Participants.
Performance-Based Participants

Across the Performance-Based participants, gains were made for half of the participants in AEB. These improvements decreased the gap between target students levels of AEB and control peers’ levels. The impact of CICO on AEB is discussed for each participant below.

*Cameron.* Mean levels of AEB for Cameron were 57% for baseline, 62% for CICO, 63% during the withdrawal, and 82% when CICO was re-implemented, for an overall change from baseline to CICO of 25%. During baseline, Cameron exhibited a decreasing trend of AEB, with an immediate increase occurring when CICO was implemented, but decreasing during the final observation of the phase. During the withdrawal, levels of AEB remained within baseline and CICO levels, but when CICO was re-implemented, AEB increased and returned to the higher levels evidenced initially during the CICO phase. With the exception of one datum during CICO, Cameron’s levels of AEB were in the range of peers during CICO implementation as compared to much lower levels exhibited than peers when CICO was not implemented.

*Bree.* Mean levels of AEB for Bree were 46% for baseline, 90% for CICO, 56% during the withdrawal, and 65% when CICO was re-implemented, for an overall change from baseline to CICO of 19%. Bree exhibited stable and low levels of AEB during baseline that were far lower than those evidenced by peers, with immediate and stable increases in AEB occurring when CICO was implemented to levels higher than those of control peers. When CICO was withdrawn, AEB levels decreased to baseline levels, and increased only slightly but above baseline levels when CICO was re-implemented. However, Bree’s AEB levels remained below control peers’ levels of AEB.
Rachel. Mean levels of AEB for Rachel were 24% for baseline, 76% for CICO, 91% during the withdrawal, and 97% when CICO was re-implemented, for an overall change from baseline to CICO of 73%. During baseline, Rachel exhibited low levels of AEB that decreased across the phase and that were below levels observed by peers. When CICO was implemented, immediate increases of large magnitude were evidenced but decreased slightly across the phase; however, even with the slightly decreasing trend, CICO resulted in increased levels of AEB that were in range with peers’ levels of AEB. During the withdrawal, Rachel exhibited the highest levels of AEB observed across observations that were higher than levels observed by peers, but when CICO was re-implemented, AEB levels increased further and remained in line with AEB levels of peers.

Nate. Mean levels of AEB for Nate were 61% for baseline, 90% for CICO, 68% during the withdrawal, and 69% when CICO was re-implemented, for an overall change from baseline to CICO of 8%. When CICO was implemented with Nate, immediate increases in AEB were evidenced and continued to increase across observations to levels higher than those observed in peers. When CICO was removed, AEB decreased to baseline levels, increasing only slightly, but above baseline and withdrawal levels, when CICO was re-implemented and remaining below the AEB levels of control peers. Graphs for the Performance-Based participants and control peers are provided in Figure 4 below.
Figure 4. Academically Engaged Behavior Across Phases for Participants in the Pre-Set Conventional Participants.
Problem Behavior

Levels of problem behavior were assessed via direct observation for all participants and control peers. For the Pre-Set Conventional participants, mean problem behavior occurred in 29% (range = 9-47%) of baseline intervals, 12% (range = 1-37%) for CICO, 19% (range = 7-41%) for withdrawal, and 16% (range = 2-40%) when CICO was re-implemented. Therefore, the mean percentage of intervals of problem behavior decreased by 14% (range = increase of 3% -decrease of 36%) from baseline to CICO for the Pre-Set Conventional participants. For the Performance-Based participants, mean problem behavior occurred in 31% (range = 21-47%) of baseline intervals, 10% (range = 0-27%) for CICO, 12% (range = 4-20%) for withdrawal, and 12% (range = 2-27%) when CICO was re-implemented. Therefore, the mean percentage of intervals of problem behavior decreased by 19% (range = 3-39%) from baseline to CICO for the Performance-Based participants.

Pre-Set Conventional Participants

Two of the four participants assigned to the Pre-Set Conventional participants exhibited decreases in problem behavior across CICO implementation. The remaining two participants exhibited variable outcomes. The problem behavior observed for each participant is provided below.

Samantha. Mean levels of problem behavior for Samantha were 30% for baseline with levels divergent from control peers. When CICO was implemented, a decrease was observed to a mean level of 19%, bringing Samantha’s problem behavior more in range of peers. When CICO was removed, mean levels of problem behavior increased to 21%,
but were approaching levels observed by peers. Mean levels did not decrease further when CICO was re-implemented, remaining at 21%.

_Nolan._ Mean levels of problem behavior for Nolan were 30% for baseline, 15% for CICO, 40% during the withdrawal, and 33% when CICO was re-implemented. Despite the gains evidenced for Nolan across CICO implementation, he exhibited increases in problem behavior during the last observation, which increased means for the phase. Therefore, Nolan’s mean level of problem behavior increased from baseline to CICO by 3%.

_Kelsey._ For Kelsey, mean levels of problem behavior were 18% for baseline, 9% for CICO, 8% during the withdrawal, and 7% when CICO was re-implemented, for an overall decrease from baseline to CICO of 8%. Kelsey’s levels of problem behavior during baseline were largely divergent from control peers. When CICO was implemented, Kelsey’s levels of problem behavior were similar to those exhibited by control peers. When CICO was removed, Kelsey’s percentage of intervals of problem behavior increased above control peers’ levels, but decreased again when CICO was re-implemented.

_Chris._ Chris exhibited 39% intervals of problem behavior during baseline, 6% during CICO, 7% during the withdrawal, and 3% when CICO was re-implemented, for an overall decrease from baseline to CICO of 36%. During baseline, Chris exhibited increasingly high levels of problem behavior that were elevated as compared to control peers. When CICO was implemented, problem behavior immediately decreased and was within range of control peers. When CICO was removed, Chris’ problem behavior remained low and at a similar level to that observed by control peers, with further
decreases being evidenced when CICO was re-implemented. Graphs for the Pre-Set Conventional participants are provided below in Figure 5.
Figure 5. Problem Behavior Across Phases for Participants in the Pre-Set Conventional Participants
**Performance-Based Participants**

Across the Performance-Based participants, gaps between observed levels of problem behavior exhibited by participants and control peers were lessened through CICO implementation. Despite these gains, experimental control was not demonstrated for all participants as increases in problem behavior were not observed for all participants during the withdrawal of CICO phase. The effects of CICO on problem behavior for each participant with the Performance-Based criterion for reinforcement are discussed below.

*Cameron.* Mean levels of problem behavior for Cameron were 23% for baseline, 20% for CICO, 18% during the withdrawal, and 12% when CICO was re-implemented, for an overall decrease from baseline to CICO of 11%. In baseline, Cameron exhibited an increasing level of problem behavior that was at a higher level exhibited than that of his peers. When CICO was implemented, problem behavior immediately decreased to the level of control peers, but increased across CICO implementation. During withdrawal, levels of problem behavior remained within CICO levels and at a higher level than control peers. When CICO was re-implemented, Cameron’s levels of problem behavior exhibited a decreasing trend across observations to a level within range of his peers.

*Bree.* Mean levels of problem behavior for Bree were 35% for baseline, 5% for CICO, 20% during the withdrawal, and 13% when CICO was re-implemented, for an overall decrease from baseline to CICO of 22%. During baseline, Bree’s levels of problem behavior were higher than levels exhibited by control peers, but decreased across CICO implementation to a level lower than that of control peers. When CICO was removed, Bree’s problem behavior remained at a level similar to control peers, and remained low and in range of peers when CICO was re-implemented.
Rachel. Mean levels of problem behavior for Rachel were 41% for baseline, 10% for CICO, 4% during the withdrawal, and 2% when CICO was re-implemented, for an overall decrease from baseline to CICO of 39%. During baseline, Rachel exhibited levels of problem behavior that were higher than levels evidenced by control peers that increased across the phase. When CICO was implemented, Rachel’s problem behavior immediately decreased to levels of peers but increased across the phase, and decreased further during the withdrawal phase. When CICO was re-implemented, Rachel’s problem behaviors decreased to very low levels that were within range of control peers.

Nate. Nate exhibited mean levels of problem behavior of 24% for baseline, 6% for CICO, 6% during the withdrawal, and 21% when CICO was re-implemented, for an overall decrease from baseline to CICO of 3%. During baseline, Nate exhibited levels of problem behavior that increased, diverging from peers’ levels, but that immediately decreased when CICO was implemented and continued to decrease across the phase. When CICO was removed, Nate’s levels of problem behavior increased above levels evidenced by peers, but remained elevated as compared to peers when CICO was re-implemented. Although levels were elevated, Nate’s problem behavior was more closely aligned with levels exhibited by peers during CICO implementation. Figure 6 depicts the levels of problem behavior for participants assigned to the Performance-Based participants.
Figure 6. Problem Behavior Across Phases for Participants in the Performance-Based Group
Disruptive Behavior

Disruptive behavior as indicated by rate of office discipline referrals (ODRs) per month was analyzed for all participants. Monthly rates were prorated to account for variable number of school days per month as well as to account for the fewer number of days during the withdrawal phase. For all students with the exception of Cameron, monthly ODR data were available for the duration of the school year. Cameron began attending the school immediately prior to participating in the current study, so limited ODR data were available prior to initiation of the current study.

**Pre-Set Conventional**

The monthly rate of ODRs for participants in the Pre-Set Conventional group was 0.15, with a small increase occurring during CICO implementation to 0.17 ODRs per month. Samantha and Chris exhibited 0 ODRs across the school year. Kelsey exhibited 0.45 ODRs per month prior to CICO implementation with an increase to 0.67 ODRs per month during CICO implementation. Noah received 0.13 ODRs per month prior to CICO and 0 ODRs per month during CICO. Therefore, one Pre-Set Conventional participant received fewer ODRs during CICO, one received more ODRs, and two received the same number of ODRs.

**Performance-Based**

The monthly rate of ODRs for participants with the Performance-Based criterion for reinforcement was 2.36, with a decrease occurring during CICO implementation to 0.08 ODRs per month. Across all participants with the Performance-Based criterion for reinforcement, decreases in rates of ODRs per month were evidenced. Rates of ODRs per month prior to CICO were 0.29 for Bree, 1 for Nate, 0.95 for Rachel, and 7.26 for
Cameron as he received an ODR during the withdrawal phase. During CICO implementation, ODRs per month decreased to 0 for Bree, Rachel, and Cameron, with Nate’s rate of ODRs per month decreasing to 0.33.

Treatment Integrity

Treatment integrity was assessed via direct observation and through a review of permanent products. Direct observations occurred for check-in, check-out, and during the observed class time. For the Pre-Set Conventional participants, mean treatment integrity was 97% (range = 67-100%), 98% (range = 67-100%), and 85% (range = 0-100%) for check-in, check-out, and class procedures, respectively. Mean levels of treatment integrity for Samantha were 93%, 93%, and 80% for check-in, check-out, and class procedures, respectively. For Nolan, mean levels of treatment integrity were 100%, 97%, and 94% for check-in, check-out, and class procedures, respectively. Treatment integrity for Kelsey was 95%, 100%, and 80% across procedures, respectively. IOA was assessed across 30% (range = 22-42%) of treatment integrity checks and IOA was 100%. Chris’ treatment integrity was 100% for both check-in and check-out and 85% for class procedures. Collectively, levels of treatment integrity were high across the Pre-Set Conventional participants, with the lowest levels occurring for the class procedures.

Across the Performance-Based participants, mean levels of treatment integrity were 99% (range = 75-100%) for check-in, 95% (range = 67-100%) for check-out, and 91% (range = 50-100%) for class procedures. Cameron’s levels of treatment integrity were 100% for check-in and check-out and 90% for class procedures. For Rachel’s, treatment integrity was 100%, 97%, and 94% across check-in, check-out, and class procedures, respectively. Nate’s treatment integrity was 100% for both check-in and
check-out and 92% for class procedures. Bree’s treatment integrity ranged from 99%, 95%, and 91% across check-in, check-out, and class procedures, respectively. These levels of treatment integrity suggest that across the Performance-Based participants, CICO was largely implemented as intended.

Permanent products were reviewed for all participants. Mean overall treatment integrity was calculated for each group and participant as well as for each component across criterion for reinforcement groups. For the Pre-Set Conventional participants, the mean treatment integrity was 89% of the CICO components reviewed occurred as intended. For the Performance-Based participants, the mean treatment integrity was 93%. Within the Pre-Set Conventional participants, mean levels of treatment integrity were 90%, 80%, 98%, and 89% for Samantha, Nolan, Kelsey, and Chris, respectively. Across these participants, treatment integrity was 100% for Daily Behavior Report Card (DBRC) given, 99.5% (range = 99-100%) for the student being rated across all time periods, 93% (range = 86-100%) for check-out occurring, 91% (range = 86-100%) for percentage of points earned being calculated, and 64% (range = 24-92%) for parent signing DBRC. However, the treatment integrity of parent signatures were approximately 64% and above for all participants except for Nolan. Within the Performance-Based participants, mean levels of treatment integrity were 93%, 81%, 92%, and 100% for Cameron, Bree, Rachel, and Nate, respectively. Across these participants, treatment integrity was 100% for DBRC given, 98% (range = 94-100%) for the student being rated across all time periods, 100% for check-out occurring, 88% (range = 56-100%) for percentage of points earned being calculated, and 88% (range = 24-92%) for parent signing DBRC.
Acceptability

The IRP-15 was issued to adult participants and the CIRP was issued to child participants following completion of CICO in order to determine their perceptions of the acceptability, feasibility, and perceived effectiveness of CICO. Adult participants reported a mean level of acceptability of 77 (range = 72-89), suggesting that the adult participants found CICO to be acceptable as scores above 52.5 are typically considered acceptable. The statement they indicated agreeing with the most was *This procedure was effective in changing the child’s problem behavior* and the least with *This was reasonable for the child’s problem behavior*. The Pre-Set Conventional participants exhibited a mean total of 76.6, with the Performance-Based participants exhibiting a mean total of 78.

Child participants reported a mean level of acceptability of 12 (range = 7-16) based on responses to items of the CIRP, suggesting that the child participants perceived CICO as acceptable. Child participants rated the statement of *I liked the program we used* as the statement they agreed with the most and the statement of *There were better ways to help me with my behavior* as the statement they agreed with the least. Participants assigned to the Pre-Set Conventional criterion for reinforcement rated a mean total of 10.2 with the Performance-Based participants rating an average of 14, suggesting that the Performance-Based participants perceived CICO as more acceptable and effective. Overall, it appeared that child and adult participants perceived CICO as acceptable and effective.
CHAPTER IV
DISCUSSION

Although the broad goals of schools are academic, addressing increasing levels of disruptive behavior is a necessary precursor for academic achievement (Nelson, 1996). Approximately 76% of teachers surveyed reported they would be better able to provide instruction and educate their students effectively if not faced with such high levels of discipline problems (Public Agenda, 2004). However, handling problem behaviors is challenging for teachers, administrators, and consultants (LaRue et al., 2008).

Interventions that help educators overcome obstacles such as lack of training, time, and funding in order to better meet the needs of at-risk students merit continued research. Therefore, this study extended the research on CICO. Although CICO serves as a promising intervention, it is currently without sufficient empirical support. For this reason, the current study addressed a limitation within the CICO literature by examining methods of setting the criterion for reinforcement within CICO.

Across the dependent measures, the Performance-Based and Pre-Set Conventional participants each responded positively to CICO, with participants of each group of goal setting demonstrating greater gains across different dependent variables. Overall, results demonstrated the effectiveness of CICO, without providing evidence for the superiority of either method of criterion for reinforcement. For teacher ratings of appropriate behavior, the mean change from baseline to CICO was 18% greater for the Pre-Set Conventional participants. Despite the greater gains in teacher ratings of appropriate behavior evidenced by the Pre-Set Conventional participants, higher levels of variability were evidenced by these participants. It is unclear why these participants had an elevated
level of variability; however, it is suspected to be attributed to these students being required to exhibit a high level of appropriate behavior too soon. In line with this, it is hypothesized that as these students received poor ratings early in the day they may have believed that they could no longer meet their criterion for reinforcement and, therefore, continued to exhibit a high level of problem behavior. Overall there was an intervention effect for three of the four Pre-Set Conventional participants, albeit with limited data points for one participant. Similarly, there was a strong intervention effect for two of the four Performance-Based participants; however, there was not a replication of intervention effects for the remaining two participants. It is hypothesized that the low levels of teacher ratings of appropriate behavior for Bree may have been associated with the extenuating family circumstances that were taking place at this time. For Cameron, it appears that skill deficits may have prevented him from exhibiting a higher level of appropriate behavior consistently; suggesting that he likely required a more intensive level of interventions and supports.

Statistical analyses based on teacher ratings of appropriate behavior suggested that levels of PAND exhibited for participants with the Performance-Based criterion for reinforcement were 3% greater as compared to participants with the Pre-Set Conventional criterion for reinforcement; however, both groups of participants exhibited similar scores that suggested overall high levels of PAND. Odds ratios calculated for teacher ratings of appropriate behavior suggest that odds ratios were greater for the Pre-Set Conventional participants when all four participants’ scores were taken into account. However, when the odds ratio was calculated based on the three participants who did not exhibit 100% non-overlap between phases, the odds ratio was 3.1 (range = 1.6 - 5) as compared to the
odds ratio of 12 (range = 3.5 - 17) for the Performance-Based participants. It is important to note that Nolan’s 100% non-overlapping score was an extreme score that was not representative of the odds ratios of participants in either group. Despite the inclusion of these statistical analyses, visual analysis should be the primary method used to interpret the data as these statistical measures may magnify small and inconsistent differences in effect size across the participants of the different groups.

Disruptive behavior (i.e., ODRs) evaluations suggested that participants with the Performance-Based criterion for reinforcement evidenced larger decreases from pre-CICO to CICO implementation (i.e., 2.28 for Performance-Based, 0.08 for Pre-Set Conventional). As the situation with Nolan’s odds ratio, Cameron’s rate of ODRs without CICO was far greater than the levels evidenced by the other participants across both groups of participants. Cameron received an ODR during the withdrawal phase. As the number of days within this phase was small, the rate of ODRs for non-CICO instances for the month was elevated. As Cameron had limited non-CICO data due to his late enrollment in the year, he did not have the lower levels of ODRs across other months in which to lessen the overall monthly rate of ODRs during non-CICO months. However, if the outlier (i.e., Cameron) for the Performance-Based participants is not included in the group mean, the Performance-Based participants continue to exhibit greater decreases (i.e., change of 0.7 as compared to change of 0.02 for Pre-Set Conventional participants) from non-CICO months to months in which CICO was implemented.

Across observed child dependent variables, participants in both groups exhibited generally large gains in academically engaged behavior (AEB) (i.e., 31% and 32% for Pre-Set and Performance-Based, respectively) from baseline to CICO, highlighting large
improvements made during CICO implementation. Participants with the Performance-Based criterion for reinforcement evidenced larger decreases in levels of problem behavior from baseline to CICO as compared to the Pre-Set Conventional participants. Both of these observed measures support the effects evidenced in teacher ratings of appropriate behavior and highlight the gains made by participants in regard to lessening the gap between their performance and their peers. Use of direct observations in the current study enhances the interpretation of outcomes as control peers’ levels of problem behavior and AEB were included. The goal of Tier 2 interventions are to aid recipients in exhibiting levels of performance (e.g., behavior, academic performance) in line with the majority of their peers who are receiving only Tier 1 interventions and supports. These direct observations allow the reader to assess the magnitude of difference between target students and control peers’ levels of performance, thereby increasing the social validity of outcomes. In this way, participants’ behavior change can be evaluated as compared to non-intervention phases as well as compared to peers’ levels of the same behavior.

Acceptability ratings by teachers and children suggest that both found CICO to be acceptable and effective. Teacher ratings indicated that those within the Performance-Based participants rated CICO as slightly more acceptable; however, the magnitude of difference in mean ratings compared to participants assigned to the Pre-Set Conventional group was small. Child participants assigned to the Pre-Set Conventional group rated CICO as more acceptable overall, with Kelsey and Samantha reporting the highest levels of acceptability. These results are similar to those previously shown within the CICO literature (i.e., Filter et al., 2007; Hawken, 2006; Hawken & Horner, 2003; Hawken et al.,
2007; March & Horner, 2002; Simonsen et al., 2011), suggesting that participants find CICO acceptable and perceive it as effective.

Treatment integrity that was directly observed and assessed via permanent products suggests that the adult participants were able to implement CICO as intended. It is believed that the elementary settings are more conducive to implementing CICO with higher levels of treatment integrity. Challenges that were present in Harpole (2010) such as the various teachers each student had throughout the day, the challenge in having the students leave their classes to attend check-in and check-out and loss of Daily Behavior Report Cards (DBRCs) were avoided in the current study due to the elementary setting. The elementary setting allowed for the students’ teachers to serve as their CICO coordinators, making check-in and check-out easier as most participants were present in their CICO coordinator’s classes at these times. Additionally, as elementary students typically change classes as a group, teachers were able to directly exchange DBRC forms between each other; therefore, students were not in possession of the forms, which prevented them from being lost, as was a problem evidenced with high school students in Harpole (2010).

Overall, results suggest that participants experienced decreases in problem behaviors and gains in teacher ratings of appropriate behavior and AEB during CICO implementation. Gains evidenced were substantial for both groups with greater improvements being evidenced for the Performance-Based participants on the dependent measures of PAND for teacher ratings of appropriate behavior, observed problem behavior, and disruptive behavior. On the other hand, greater gains were made for the
participants of the Pre-Set Conventional group on the measures of teacher ratings of appropriate behavior and odds ratios based on this measure and AEB.

It should be noted that greater levels of improvement and more stable levels of performance may have been evidenced by Bree if possible extenuating home situations were not occurring, which required the school to make two reports to the Department of Human Services. Additionally, Cameron’s baseline levels of teacher ratings of appropriate behavior were 15%, which is 22% lower than the next lowest baseline scores of other participants. Therefore, his scores may have served to overall lessen group means of improvement. Similarly, Nate exhibited the highest levels of teacher ratings of appropriate behavior in baseline; therefore, ceiling effects may have prevented greater gains from occurring.

Although participants with the Pre-Set Conventional criterion for reinforcement evidenced more improvements than were hypothesized by the authors, it should be noted that the participants in this group were able to meet or exceed the 80% criterion for reinforcement within the first three days of implementation (i.e., one day for Nolan and Chris, two days for Kelsey, and three days for Samantha). Therefore, the participants were able to access the reward, strengthening the contingency between a level of appropriate behavior and access to rewards. This response time suggests response to intervention occurred for these participants, as Campbell and Anderson (2008) highlighted the rapid response (i.e., within one week) incurred with Tier 2 interventions. Had these participants been unable to access the reward after some number of days, it is likely that they may have no longer attempted to exhibit the desired level of appropriate behavior. Data collected suggest that participants who exhibited a mean baseline level of
teacher ratings of appropriate behavior of at least 37% were able to meet the 80% criterion for reinforcement within three days. However, it is unlikely that Cameron (i.e., baseline mean of 15% teacher ratings of appropriate behavior) would have been able to access the 80% criterion for reinforcement in a timely manner.

Therefore, it may be reasonable for schools to implement an 80% criterion for reinforcement for students who exhibit at least a 37% mean level of teacher ratings of appropriate behavior in baseline and to use a lower initial criterion for reinforcement for students exhibiting lower levels in baseline. Mean gains from baseline to the first CICO phase for all participants were 30% (range = 19-36%). This suggests that for participants who are exhibiting very low levels of teacher ratings of appropriate behavior in baseline (i.e., less than 37%), an initial criterion for reinforcement could be their mean baseline score plus 20%. In this way, students with low levels of baseline teacher ratings of appropriate behavior would be able to access the reward, increasing the likelihood that they continued to attempt to exhibit a desired level of appropriate behavior. However, additional evaluation of this method is required. Another option for schools may be to set a number of days at which to modify the criterion for reinforcement if the child was unable to meet the 80% criterion for reinforcement. Both of these methods would likely increase the success of all children who received CICO as an intervention for their behavior.

For children who exhibit very low levels of teacher ratings of appropriate behavior in baseline (e.g., Cameron), additional intervention or skill building may be required in place of or in addition to CICO. For example, although Cameron exhibited gains from baseline to CICO of 12%, these gains are insufficient for Cameron to be
successful in the classroom. Therefore, Cameron may have benefitted from a more intensive, individualized, or function-based intervention or skill building as it appears that his inability to meet his criterion for reinforcement consistently was due to a skill deficit. On the other hand, the participants who were able to immediately exhibit a high level of appropriate behavior (i.e., 80%) when reinforcement was introduced likely exhibited problem behavior associated with performance deficits in that they had the skills to exhibit appropriate behavior but did not exhibit these behaviors until reinforcement was introduced. Taking this information into consideration, baseline data collection could be used to identify students who likely exhibited student behavior associated with performance deficits versus those who exhibited problem behavior associated with skill deficits as this would identify students who could likely respond to use of a Pre-Set Conventional criterion for reinforcement versus a Performance-Based criterion for reinforcement potentially supplemented with skill building, respectively. In this way, early assessments of teacher ratings of appropriate behavior and brief experimental analyses or probe sessions may be utilized to inform teachers and schools in efficiently planning services and supports for children in order to increase the likelihood of success for all students.

Limitations

The current study had a number of limitations that should be considered. The first represents the limited direct observation data. Although the study includes direct observation data, observations were only conducted during one class period on a limited number of days for each child. Therefore, direct observation data cannot provide an overall representation of the child’s behavior across the school day. In line with this,
there was an overreliance on DBRC data for decision-making. Although the current study provided a more comprehensive presentation of participant behavior by providing DBRC data as opposed to relying solely on ODR data, as the greater part of CICO research has, DBRC data continues to have limitations as compared to direct observation data. The major limitation of DBRC data is the delay in rating until the completion of the rating interval as opposed to direct observation data in which behaviors are rated as they occur or shortly thereafter. In this way, DBRC continues to share with ODRs a reliance on teacher perceptions of student behavior. Although DBRC data may align with direct observation data, additional research is needed to lend support to this notion.

As the current study evaluated methods of criterion setting across two groups of participants, the ABAB designs did not allow for direct comparison across these methods of criterion setting. As this serves as a pilot study, future studies should directly compare these two methods of criterion setting through manipulation of the research design or through an increased number of participants. Additionally, the results are limited by the number of participants as well as the grade level (e.g., elementary) of the participants, as it is unclear whether results would have generalized to other grade levels (i.e., middle school, high school). An increased number of participants would have allowed for an increased level of confidence in the assumptions gleaned. It would have also been desirable for the second phase of CICO implementation to have been extended for all participants; however, the end of the school year prevented this from occurring. This resulted in variable length of intervention phases with second CICO phases for some participants being discrepant in length from the initial CICO phase and thereby much shorter in duration.
Finally, the outcomes of the current study are limited as a number of participants’ performance did not decrease during withdrawal phases, which would have increased the level of experimental control. These results were similar to those found in the only other study (i.e., Campbell & Anderson, 2011) to use a withdrawal design to evaluate the effectiveness of CICO. Participants in Campbell and Anderson (2011) exhibited levels of directly observed academic engagement during the withdrawal phase that was within range of the CICO phase, suggesting that maintenance of appropriate behaviors may occur to some extent after CICO is withdrawn. However, it is unclear how long these gains would persist in the absence of CICO. Nonetheless, had less overlap occurred between phases, more experimental control would have been evidenced thereby increasing the reader’s confidence that outcomes achieved were produced by CICO as opposed to extraneous factors.

Future Research

As more studies are now demonstrating the effectiveness of CICO, future research should begin to refine CICO procedures and decision-making to enhance CICO effectiveness. A fruitful line of future research should evaluate likely outcomes associated with various levels of teacher ratings of appropriate behavior in baseline with CICO implementation. In line with this, combinations of CICO and other interventions should be evaluated to increase the likelihood that CICO will affect change in most students, as has not been the case across most studies. In this way, treatment planning can be more empirically guided, thereby aiding school-based interventionists who have varying levels of experience in training to increase the effectiveness and efficiency of interventions.
Implications

This study contributes to the literature in a number of ways. The current study demonstrates that CICO is a viable intervention effective at improving the appropriate behavior and decreasing the inappropriate behavior of students evaluated in a study that combined evaluation of direct observations of participants’ behavior and evaluation of treatment integrity. The current study has implications for treatment planning based on baseline levels of teacher ratings of appropriate behavior in order to develop more effective intervention packages for students with low levels of appropriate behavior during baseline. The study highlights that CICO is likely more easily implemented with higher levels of treatment integrity in the elementary setting as compared to the high school setting as was evaluated in Harpole (2010). Collectively, this study supports the notion that CICO is a robust intervention that serves as an effective method to provide ongoing assessment and treatment of elementary student’s problem behavior.
APPENDIX A

PARENT/GUARDIAN CONSENT FOR STUDENT PARTICIPATION

University of Southern Mississippi
Consent Document for Research Participants

Title of Study:
Evaluation of Pre-Set Conventional Versus Performance-Based Criterion for Reinforcement in Check In-Check Out

Purpose
Your child is being asked to participate in a study that is studying the effects of an intervention in increasing appropriate behavior. This study is important because it will evaluate the effectiveness of an efficient intervention for schools to implement in order to address the behavioral needs of at-risk students.

Participants:
Your child was selected for participation because he or she received a problematic number of office discipline referrals during the current school year, he or she was recommended by a teacher or administrator due to presenting social behavior concerns, and because his problem behaviors do not include serious, dangerous, or infrequently occurring behaviors.

Procedure:
If you agree to allow your child to participate in this study, your child will participate in the intervention. Through the intervention, your student will receive structured feedback about his or her behavior throughout the school day.

Benefits/Risks to Participant:
Your child’s participation in the study will provide him or her with additional teacher and staff attention and feedback, in an attempt to improve his behavior at school. Participation in the study will provide the student with clear behavior expectations, increase the students’ routine throughout the day, provide additional adult attention and feedback, and increase home-school communication. Although not evaluated, students may benefit from this study in that increases in appropriate behavior may result in increased academic performance and social functioning. Few risks are anticipated as a result of participation on the study. The potential risks include a possible increase in the student’s inappropriate behavior as the use of these procedures could increase inappropriate behavior. If the student does not respond to the intervention or his behavior worsens, the primary investigator will provide or refer for additional intervention services. Participants may also refuse to complete the study at any point during the experiment.
**Voluntary Nature of the Study/Confidentiality:**
Your child’s participation in this study is entirely voluntary and you may refuse to complete the study at any point during the experiment. In addition, all information obtained during the study will be kept confidential. All information that may identify your child will be withheld. Your child’s name and other identifying information will not be used in the research papers, any submission to a professional journal for publication, or presentation. The only circumstances in which we would release information about you or your child would be if he or she tells us he or she is a harm to self or others, if your child is abused, if the release of information is court ordered, or if there is a medical emergency in which release of information is important for your child’s safety.

**Contacts and Questions:**
At any time you may withdraw from the study or ask any questions you may have regarding this study. Questions concerning the research should be directed at Lauren Harpole or Dr. Joe Olmi at (601) 266-5255 or via email at laurenharpole@gmail.com or d.olmi@usm.edu. This project has been reviewed by the Human Subjects Protection Review Committee, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research subject should be directed to the chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5147, Hattiesburg, MS 39406-0001, (601) 266-6820.

**Parental Consent:**
I have had the purposes and procedures of this study explained to me and have had the opportunity to ask questions. My questions have been answered to my satisfaction, and I am voluntarily signing this form to participate in this research study. My signature shows my willingness to allow my child to participate in this study under the conditions stated.

This section to be completed by Parent/Guardian.

__________________________________
Signature of Parent/Guardian

______________________________
Date

Best time to be contacted: ____________________________________________

Phone #: __________________________________________________________

Email Address (Optional): ____________________________________________

Would you prefer being provided an orientation to this intervention, which will occur at a later date, in a meeting at the school or over the telephone? School ____  Phone ____
Title of Study:
Evaluation of Pre-Set Conventional Versus Performance-Based Criterion for Reinforcement in Check In-Check Out

Purpose
You are being asked to participate in a study that is studying the effects of an intervention in increasing student appropriate behavior. This study is important because it will evaluate the effectiveness of an efficient intervention for schools to implement in order to address the behavioral needs of at-risk students.

Participation:
You are being asked to participate because one of your students is participating in the study, or you have been nominated to serve as the coordinator of the intervention for a student.

Procedure:
If you agree to participate in this study, you will be participating in an intervention that provides increased attention and feedback to an at-risk student in an attempt to increase his appropriate behaviors.

Benefits/Risks to Participant:
Teachers and participating staff will be provided with increased consultation related to behavior strategies used in the study which can be generalized to other students. Few risks are anticipated as a result of participation on the study. Participants may also refuse to complete the study at any point during the experiment.

Voluntary Nature of the Study/Confidentiality:
Your participation in this study is entirely voluntary and you may refuse to complete the study at any point during the experiment. In addition, all information obtained during the study will be kept confidential. All information that may identify you will be withheld. Your name and other identifying information will not be used in the research papers, any submissions to a professional journal for publication, or presentation. The only circumstances in which we would release information about you would be if there is there is a threat of harm to self or others, abuse, if the release of information is court ordered, or if there is a medical emergency in which release of information is important for someone’s safety.
Contacts and Questions:
At any time you may withdraw from the study or ask any questions you may have regarding this study. Questions concerning the research should be directed at Lauren Harpole or Dr. Joe Olmi at (601) 266-5255 or via email at laurenharpole@gmail.com or d.olmi@usm.edu. This project has been reviewed by the Human Subjects Protection Review Committee, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research subject should be directed to the chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5147, Hattiesburg, MS 39406-0001, (601) 266-6820. A copy of this form will be given to the participant.

Participant Consent:
I have had the purposes and procedures of this study explained to me and have had the opportunity to ask questions. My questions have been answered to my satisfaction, and I am voluntarily signing this form to participate in this research study. My signature shows my willingness to participate in this study under the conditions stated.

This section to be completed by teacher/staff.

Signature of Teacher/Staff  Date

Planning Period Time: ________________________________

Phone #: ______________________________________

Email Address: ________________________________

Best method to contact you (phone or email): _____________
### APPENDIX C

#### CICO OBSERVATION FORM

<table>
<thead>
<tr>
<th>Behavior</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
<th>2.1</th>
<th>2.2</th>
<th>2.3</th>
<th>2.4</th>
<th>2.5</th>
<th>2.6</th>
<th>3.1</th>
<th>3.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>3.3</th>
<th>3.4</th>
<th>3.5</th>
<th>3.6</th>
<th>4.1</th>
<th>4.2</th>
<th>4.3</th>
<th>4.4</th>
<th>4.5</th>
<th>4.6</th>
<th>5.1</th>
<th>5.2</th>
<th>5.3</th>
<th>5.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>5.5</th>
<th>5.6</th>
<th>6.1</th>
<th>6.2</th>
<th>6.3</th>
<th>6.4</th>
<th>6.5</th>
<th>6.6</th>
<th>7.1</th>
<th>7.2</th>
<th>7.3</th>
<th>7.4</th>
<th>7.5</th>
<th>7.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>8.1</th>
<th>8.2</th>
<th>8.3</th>
<th>8.4</th>
<th>8.5</th>
<th>8.6</th>
<th>9.1</th>
<th>9.2</th>
<th>9.3</th>
<th>9.4</th>
<th>9.5</th>
<th>9.6</th>
<th>10.1</th>
<th>10.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>10.3</th>
<th>10.4</th>
<th>10.5</th>
<th>10.6</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
<th>2.1</th>
<th>2.2</th>
<th>2.3</th>
<th>2.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>2.5</th>
<th>2.6</th>
<th>3.1</th>
<th>3.2</th>
<th>3.3</th>
<th>3.4</th>
<th>3.5</th>
<th>3.6</th>
<th>4.1</th>
<th>4.2</th>
<th>4.3</th>
<th>4.4</th>
<th>4.5</th>
<th>4.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>5.1</th>
<th>5.2</th>
<th>5.3</th>
<th>5.4</th>
<th>5.5</th>
<th>5.6</th>
<th>6.1</th>
<th>6.2</th>
<th>6.3</th>
<th>6.4</th>
<th>6.5</th>
<th>6.6</th>
<th>7.1</th>
<th>7.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>7.3</th>
<th>7.4</th>
<th>7.5</th>
<th>7.6</th>
<th>8.1</th>
<th>8.2</th>
<th>8.3</th>
<th>8.4</th>
<th>8.5</th>
<th>8.6</th>
<th>9.0</th>
<th>9.2</th>
<th>9.3</th>
<th>9.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>9.5</th>
<th>9.6</th>
<th>10.1</th>
<th>10.2</th>
<th>10.3</th>
<th>10.4</th>
<th>10.5</th>
<th>10.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Student Name: ____________________________
Date: __________________________
Observer: __________________________
Notes: ____________________________
____________________________________
____________________________________

---
APPENDIX D

DAILY BEHAVIOR REPORT CARD

Name: _______________________________  Date: _________________________

**TEACHER PROCEDURES**

**Beginning of Class:**
- □ Teacher provided precorrections about student’s behavior expectations

**End of Class:**
- □ 1. Rate student below
- □ 2. Explain ratings to student
- □ 3. Provide praise
- □ 4. Provide corrective feedback, if needed

<table>
<thead>
<tr>
<th>Teacher Completes Section Below</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject/Class Period</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

*CICO Coordinator Completes Section Below:*

Pts. Earned = _____ ÷ Pts. Possible = _____ * 100

**Percentage Earned = ____________ %**

Met Goal of ________ % = Yes or No

Reward(s) Earned & Provided ________________

*Parent/Guardian Completes Section Below:*

Parent/Guardian Signature ___________________________

Comments: ______________________________________

_____________________________________________

________________________________________________
APPENDIX E

CHECK IN & OUT PROCEDURES

Student: ________________________  Date: ______________________

Coordinator: Place a check in the box as each step is conducted or N/A if not applicable.

Check In:

_____ Student attended check in.

_____ Signed DBRC from previous day was collected, if applicable

_____ Student was provided new DBRC.

_____ Student was precorrected about behavior expectations.

Check Out:

_____ Student attended check-out.

_____ Student was provided praise for compliance with expectations.

_____ Student was provided corrective feedback for noncompliance with expectations.

_____ Percentage of points earned on DBRC was calculated.

_____ Determined whether student met goal.

_____ Student provided incentive for meeting goal, if earned.

_____ Send DBRC home with student
APPENDIX F

PARENT/GUARDIAN AND STUDENT ORIENTATION

☐ Explain purpose of CICO
☐ Explain student’s goal percentage of points earned on the DBRC
☐ Explain teachers’ role in CICO
☐ Explain how points are scored on the DBRC
☐ Explain how students will know if they are meeting their point goals
☐ Explain how student checks in
☐ Explain how student checks out
☐ Explain how student gives DBRC to teacher
☐ Explain procedure for getting DBRC signed by parent/guardian and returning to school
☐ Complete CICO Contract
APPENDIX G
CHECK IN-CHECK OUT CONTRACT

I, ________________________________, agree to work on these things.

1. ________________________________
2. ________________________________
3. ________________________________

I will work with ____________________________ to keep track of my progress. I understand that I will have an opportunity to earn a reward each day when I meet my goals. A list of rewards I would like to earn includes:

1. ________________________________
2. ________________________________
3. ________________________________

I will try hard to do my best to meet these goals every day.

____________________________
Signature of Student

APPENDIX H

THE INTERVENTION RATING PROFILE (IRP-15)

Completed by: __________________

The purpose of this questionnaire is to obtain information that will aid in the evaluation of the intervention for __________________. Please circle the number which best describes your agreement or disagreement with each statement.

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

1. This was an acceptable procedure for the child's problem behavior.

2. Most teachers would find this procedure appropriate for problem behaviors.

3. This procedure was effective in changing the child's problem behavior.

4. I would suggest the use of this procedure to other teachers.

5. The child's problem behavior was severe enough to warrant use of this procedure.

6. Most teachers would find this procedure suitable for dealing with the child's problem behaviors.

7. I would be willing to use this procedure again.

8. This procedure did **NOT** result in any negative side-effects for the child.
9. This procedure would be appropriate for a variety of children.

10. This procedure was consistent with those I have used in the past.

11. This procedure was a fair way to deal with the child's problem behavior.

12. This was reasonable for the child's problem behavior.

13. I liked the procedure.

14. This procedure was beneficial in understanding this child's problem behavior.

15. Overall, this procedure was beneficial for the child.

Adapted from Martens, Witt, Elliott, & Darveaux (1985).
APPENDIX I

ADAPTED VERSION OF THE CHILDHOOD INTERVENTION RATING PROFILE

Completed by: _____________________

<table>
<thead>
<tr>
<th></th>
<th>I Definitely Agree</th>
<th>I Agree</th>
<th>I Somewhat Agree</th>
<th>I Somewhat Disagree</th>
<th>I Disagree</th>
<th>I Definitely Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The program used was fair.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I think my teacher was too harsh on me.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Being in this program caused problems with my friends.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. There were better ways to help me with my behavior.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. This program could help other kids too.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I liked the program we used.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Being in this program helped me do better in school.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Witt & Elliott (1985).
THE UNIVERSITY OF SOUTHERN MISSISSIPPI

Institutional Review Board
118 College Drive #5147
Hattiesburg, MS 39406-0001
Tel: 601.266.6820
Fax: 601.266.5509
www.usm.edu/trb

HUMAN SUBJECTS PROTECTION REVIEW COMMITTEE
NOTICE OF COMMITTEE ACTION

The project has been reviewed by The University of Southern Mississippi Human Subjects Protection Review Committee in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services (45 CFR Part 46), and university guidelines to ensure adherence to the following criteria:

- The risks to subjects are minimized.
- The risks to subjects are reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered regarding risks to subjects must be reported immediately, but not later than 10 days following the event. This should be reported to the IRB Office via the "Adverse Effect Report Form".
- If approved, the maximum period of approval is limited to twelve months. Projects that exceed this period must submit an application for renewal or continuation.

PROTOCOL NUMBER: 10100603
PROJECT TITLE: Evaluation of Performance-Based Versus Pre-Set Conventional Criterion for Reinforcement in Check In - Check Out
PROPOSED PROJECT DATES: 10/10/2010 to 10/10/2011
PROJECT TYPE: Dissertation
PRINCIPAL INVESTIGATORS: Lauren Lestremay Harpole
COLLEGE/DIVISION: College of Education & Psychology
DEPARTMENT: Psychology
FUNDING AGENCY: N/A
HSPRC COMMITTEE ACTION: Expedited Review Approval
PERIOD OF APPROVAL: 10/18/2010 to 10/17/2011

Lawrence A. Hosman, Ph.D.
HSPRC Chair

Date
REFERENCES


Harpole, L. L. (2010). Application of Check In-Check Out as a targeted intervention to increase appropriate behavior in at-risk high school students. (Unpublished master’s thesis). University of Southern Mississippi, Hattiesburg, Mississippi


