I, Elaine M Wahl, hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in School Psychology.

It is entitled:
Comparing Versions of the Good Behavior Game: Can a Positive Spin Enhance Effectiveness?

Student’s name: Elaine M Wahl

This work and its defense approved by:

Committee chair: Renee Oliver Hawkins, Ph.D.
Committee member: Todd Haydon, Ph.D.
Committee member: Julie Morrison, Ph.D.
Comparing Versions of the Good Behavior Game:
Can a Positive Spin Enhance Effectiveness?

A dissertation submitted to the:
Graduate School of the University of Cincinnati
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy (Ph.D.) in the
School Psychology Program of the
School of Human Services of the
College of Education, Criminal Justice, and Human Services

By
Elaine Wahl, M.Ed.
B.A., University of Evansville, 2008
M.Ed., University of Cincinnati, 2010

March 2014

Dissertation Committee:
Renee O. Hawkins, Ph.D. (Chair), School Psychology
Julie Morrison, Ph.D., School Psychology
Todd Haydon, Ph.D., Special Education
Abstract
Disruptions can prevent the classroom from being an effective learning environment. The Good Behavior Game (GBG) is a group contingency that has been proven to effectively prevent disruptions and increase engagement. However, the traditional methods of the GBG include teacher scanning for negative student behaviors, and may not align with Positive Behavioral Interventions and Supports (PBIS), which is becoming widely adopted in many schools. The study extends the findings of Wright and McCurdy (2011), who compared the GBG to a positive version of the GBG, called the Caught Being Good Game (CBGG). The study examined the effects on student behavior, including engagement and off-task behaviors, as well as teachers’ use of positive and negative statements. Results showed both the GBG and the CBGG improved student behavior, with data not clearly indicating one was superior. Neither intervention led to an increase in positive teacher statements. Implications for teachers and suggestions for further research are discussed.
Acknowledgements

I would like to extend many thanks to everyone that has supported me throughout the dissertation process. To my committee members, Dr. Renee Hawkins, Dr. Julie Morrison, and Dr. Todd Haydon, thank you for your expertise, guidance, and encouragement. I would like to thank Dr. Renee Hawkins for all of her insight and feedback. She has mentored me through multiple research and practical experiences that have helped me to become a well-rounded student and future school psychologist. I would also like to thank Dr. Julie Morrison for all of her support and for encouraging me to further my education. Dr. Todd Haydon has given me valuable feedback and a new perspective.

I also would like to thank my family for their constant love and support. Thank you to my husband, Kasey, who has been so supportive and has made many sacrifices without hesitation. To my parents, thank you for always being there for me. My mom has edited too many papers to count over the years. To all of my siblings, your support means the world.
# Table of Contents

Abstract ................................................................................................................................. ii

Acknowledgements ............................................................................................................... iv

Table of Contents ................................................................................................................. v

List of Tables ............................................................................................................................ vii

List of Figures .......................................................................................................................... viii

Introduction .............................................................................................................................. 1

Group Contingencies .............................................................................................................. 2

The Good Behavior Game ....................................................................................................... 5

Randomizing Contingency Components ............................................................................. 6

Positive Variations of Group Contingencies ......................................................................... 8

Current Study ............................................................................................................................ 11

Method ..................................................................................................................................... 12

Materials ................................................................................................................................ 14

Dependent Measures and Data Collection .......................................................................... 15

Student Engagement .............................................................................................................. 16

Disruptions ............................................................................................................................... 16

Teacher Statements ................................................................................................................. 16

Inter-Observer Agreement .................................................................................................... 18

Research Design .................................................................................................................... 19

Experimental Conditions ..................................................................................................... 19

Baseline .................................................................................................................................. 19

Training .................................................................................................................................. 19
List of Tables

Table 1. Inter-Observer Agreement Results ................................................................. 18
Table 2. Adherence Results by Classroom ................................................................. 24
Table 3. Classroom 1 PND ....................................................................................... 29
Table 4. Classroom 2 PND ....................................................................................... 33
Table 5. Classroom 3 PND ....................................................................................... 36
Table 6. Classroom 4 PND ....................................................................................... 39
Table 7. Classrooms 1 and 2 Teacher Statements PND ......................................... 44
Table 8. Classrooms 3 and 4 Teacher Statements PND ......................................... 45
Table 9. Teacher Social Validity Survey Results .................................................... 47
Table 10. Student Social Validity Survey Results ..................................................... 48
List of Figures

Figure 1. Classroom 1 Engagement and Disruptions.....................................................28
Figure 2. Classroom 1 Off-Task Motor ........................................................................29
Figure 3. Classroom 2 Engagement and Disruptions.....................................................31
Figure 4. Classroom 2 Off-Task Motor ........................................................................32
Figure 5. Classroom 3 Engagement and Disruptions.....................................................34
Figure 6. Classroom 3 Off-Task Motor ........................................................................35
Figure 7. Classroom 4 Engagement and Disruptions.....................................................37
Figure 8. Classroom 4 Off-Task Verbal .......................................................................38
Figure 9. Classroom 1 Teacher Statements.................................................................41
Figure 10. Classroom 2 Teacher Statements...............................................................42
Figure 11. Classroom 3 Teacher Statements...............................................................43
Figure 12. Classroom 4 Teacher Statements...............................................................44
Introduction

Teachers are expected to ensure their students succeed academically and, to achieve this goal, teachers are specifically trained in methods and strategies to teach information that will help their students become well educated. In addition, positive and effective proactive discipline measures are essential for creating an environment more conducive to learning and academic success. In fact, the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) requires that Positive Behavioral Interventions and Supports (PBIS) be considered when addressing behavioral concerns. A key characteristic of PBIS is that not only are disruptive and problem behaviors reduced, but positive and prosocial behaviors are also taught and reinforced (Sailor, Dunlap, Horner, & Sugai, 2009). Schools that successfully implement PBIS incorporate a tiered approach to behavior management that focuses on prevention of problem behaviors. The tiers represent a continuum of support, with the first tier representing primary prevention across the school or classroom, the second tier representing secondary support for those at-risk, and the third tier representing more intense and individualized support (Sugai & Horner, 2002). With strong and evidence-based strategies at tier one, less individualized and intensive interventions are needed and resources may be used more appropriately. PBIS shows promise in addressing challenging behaviors and is associated with an increase in perceived school safety as well as increased academic outcomes (Fox, Dunlap, & Powell, 2002; Horner et al., 2009; Lassen, Steele, & Sailor, 2006; Luiselli, Putnam, & Sunderland, 2002; Todd, Haugen, Anderson, & Spriggs, 2002).

Classwide tier one supports are especially critical for classrooms that have
multiple students exhibiting disruptive behaviors. Student disruptions such as not listening or talking at inappropriate times are a common classroom problem (Luiselli et al., 2002) and are found by teachers to be one of the most common obstacles to promoting successful learning (Arbuckle & Little, 2004). Student disruptions, even those that are mild, are often managed ineffectively through negative disciplinary methods, such as reprimands, office referrals, and suspensions (Skiba, Peterson, & Williams, 1997). Unfortunately, these methods may reinforce the students’ negative behavior (Mayer, 2002; Nelson & Roberts, 2000) and do not teach students how to behave appropriately (Skiba & Peterson, 2000).

Evidence-based tier one classwide behavior management strategies should be used in classrooms that have high rates of disruption in order to reduce the need for intense, individualized interventions (Crone & Horner, 2003). Group contingencies are one type of proactive classwide supports that have been found to effectively improve student behavior (Coogan, Kehle, Bray, & Chafouleas, 2007; Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Murphy, Theodore, Aloiso, Alric-Edwards, & Hughes, 2007; Schanding & Sterling-Turner, 2010; Theodore, Bray, Kehle, & Jenson, 2001) and can save valuable instructional time.

**Group Contingencies**

With the use of a group contingency, the entire class shares the same target behaviors, criteria, and reinforcement (Kelshaw-Levering et al., 2000). There are three types of group contingencies identified in the literature; independent, dependent, and interdependent (Litlow & Pumroy, 1975). With an independent group contingency, each student’s access to the reward depends on his or her own behavior (Cooper, Heron,
Heward, 2007). A peer’s performance has no impact on one’s ability to access a reinforcer. For example, each student who receives an 80% on a quiz receives a pencil. Independent group contingencies may save teacher time because, rather than creating individual behavior plans for each student, all students share the same target behaviors, criteria, and rewards. However, there are some disadvantages to their use. With independent group contingencies, some students’ behavior may significantly improve yet still may never reach the criteria, and thus, never earn the reward (Theodore, Bray, Kehle, & Dioguardi, 2004). Using the previous example, a student who typically earns a 50% on the quiz, may increase to a 75% when the contingency is introduced. Despite the improvement, the student would not earn a reward because he or she did not meet the criteria of 80%. Also, as a result of only some students earning the reinforcer, teachers may be required to spend time managing the negative behaviors of students who do not earn the reward such as jealous comments, crying, or tantrums. Teachers are also required to manage students’ access to the reinforcers by tracking which students met the criteria and delivering the reward to only those students (Skinner, Cashwell, & Dunn, 1996).

With a dependent contingency, the whole group or class receives the reward depending on the behavior of an individual student or small group (Cooper et al., 2007). Therefore, only the selected individual(s) need to meet the criteria in order for the whole class to earn the reward. A dependent group contingency might be used to improve the behavior of one or two students. For example, one student in the classroom may often call out without permission. A teacher may set the criteria that if the student calls out two times or less during the class period, the class will earn free time. Disadvantages to dependent group contingencies include peer pressure on the target student(s) as well as
possible negative interactions toward the target student when the criterion is not met (Theodore et al., 2004).

Finally, an interdependent group contingency is one in which the whole class or group receives the reward based on the entire group’s behavior (Cooper et al., 2007). For example, Pappas, Skinner, and Skinner (2010) increased performance on reading quizzes by randomly choosing the number of students that had to pass the quiz in order for the class to earn the reward. As another example, Lannie and McCurdy (2007) divided the class into teams. Teams received checkmarks when any student from the team displayed a disruptive behavior. At the end of the period, the teams that received less checkmarks than the set criterion received a reward. Interdependent group contingencies have been used to increase positive social interactions (Cashwell, Skinner, & Smith, 2001; Kohler et al., 1995) and homework accuracy (Reinhardt, Theodore, Bray, & Kehle, 2009), as well as improve disruptive behaviors (Kelshaw-Levering et al., 2000; Murphy et al., 2007) and general academic performance (Popkin & Skinner, 2003). They have been used at the preschool (Kohler et al., 1995; Murphy et al., 2007), elementary (Kelshaw-Levering et al., 2000; Lohrmann & Talerico, 2004), middle school (Popkin & Skinner, 2003), and high school (Christ & Christ, 2006; Theodore, Bray, & Kehle, 2004) levels. They have also successfully been used to help a variety of student populations including those with serious emotional disorder (Popkin & Skinner, 2003), autism spectrum disorders (Kohler et al., 1995), learning disabilities (Lohrmann & Talerico, 2004), and those in general education classrooms (Sharp & Skinner, 2004).

There are several advantages associated with the use of interdependent and dependent group contingencies as compared to the use of independent group
contingencies. Interdependent and dependent group contingencies are easy to manage (Salend, Reynolds, & Coyle, 1989; Schanding & Sterling-Turner, 2010) and time efficient because the reward is either delivered to the whole group or none of the members. Potentially, there are also more reinforcers to choose from because the teacher may provide access to an activity for the whole class (Skinner et al., 1996). For example, listening to music during math class could be chosen as a reward for the whole group. Also, because all the members receive a reward based on some aspect of the group’s behavior, students participating in an interdependent or dependent group contingency are more likely to encourage each other’s appropriate behavior (Coogan et al., 2007; Kelshaw-Levering et al., 2000).

**The Good Behavior Game**

The Good Behavior Game (GBG) is a type of interdependent group contingency that has been well established as an effective management tool. The GBG was first implemented by Barrish, Saunders, and Wolf (1969) to reduce out-of-seat behavior and inappropriate vocalizations in a fourth-grade classroom. In their study, the teacher explained that they would be playing a game during math class, divided the class into teams, and explained that teams would be given special privileges for following the classroom rules. The teams were given points when a member broke a classroom rule. At the end of the math period, the team that had the fewest points or, both teams if they each received less than five points, were given special privileges. Since then, many teachers and researchers have used the GBG to improve classroom behaviors such as out-of-seat behavior (Harris & Sherman, 1973), aggressive behavior, (Kellam, Rebok, Ialongo, & Mayer, 1994); inappropriate vocalizations (Davies & Witte, 2000); on-task
behavior (Lannie & McCurdy, 2007; Leflot, van Lier, Onghena, & Colpin, 2010), and assignment completion (Darveaux, 1984). In their meta-analysis of the GBG, Tingstrom, Sterling-Turner, and Wilczynski (2006) explain many advantages of the GBG including an increase in prosocial behaviors and encouragement of peers and decrease in attention from peers for disruptive behaviors so that they do not earn additional points. Teachers often find the GBG easy to manage and efficient because all students are under the same contingency. In addition, follow-up studies have found that the GBG is associated with positive long-term effects for students including reduction of disruptions, aggressive behaviors, and substance abuse (Embry, 2002) as well as an increased likelihood for females to attend college (Hemelt, Roth, & Eaton, 2013).

**Randomizing Contingency Components**

Despite their numerous advantages, some possible limitations of group contingencies have been identified by researchers. First, because all students share the same rewards, a consequence that may reinforce the behavior of one student, may not be reinforcing to another student. Consequences intended to be reinforcers might not be effective for all students in a classroom (Kelshaw-Levering et al., 2000; Skinner et al., 1996). For example, a teacher may intend to use extra playground time as a reinforcer. Although playground time may be reinforcing for many students, a student may find playground time not rewarding or even aversive. As a result, the student may not work for the reward or at extremes; try to sabotage the contingency so that the class does not earn the reward (Barrish et al., 1969). One way to make a group-earned activity more broadly reinforcing is to randomize the reinforcers so that students are unaware of what the reinforcer will be (Murphy et al., 2007).
When randomizing the reward, at the end of the class period in which the group contingency is used, the teacher will often choose a slip of paper from a container that specifies the reward earned (i.e., Murphy et al., 2007; Theodore et al., 2001). For example, Murphy et al. (2007) studied the effects of an interdependent group contingency and mystery motivators, or randomized reinforcers, on disruptive behavior in a Head Start classroom during a morning activity. During the intervention, rules were established and the teacher gave students checkmarks for disruptive behaviors. At the end of the morning activity, if all students had five or fewer checkmarks, a reward was selected from a mystery motivator box. Using an ABAB reversal design, the findings suggest the interdependent group contingency with randomized reinforcers resulted in a reduction of disruptive behaviors.

Another potential limitation of group contingencies is that problematic behaviors may increase during the intervention when students are no longer able to meet the criteria to earn the reward (Skinner et al., 1996). For example, when a classroom whose criteria is to receive less than five checkmarks for disruptive behavior receives the fifth checkmark by lunch, students may display an increase in disruptive behaviors the rest of the day because they are no longer able to earn a reward. Skinner et al. (1996) suggest using randomized criteria so that student behaviors are always under control of the group contingency. Given that students are unaware of the criteria required to earn a reward, they cannot determine that they failed to meet the criteria. For example, when using a group contingency to increase reading comprehension and math homework scores, Reinhardt et al. (2009) randomized the criteria by choosing a random score out of a jar and rewarding the class when they met the score.
Similarly, another potential limitation of group contingencies is that students may improve one targeted behavior; however, another inappropriate or unacceptable behavior may increase. For example, a classroom that targets out-of-seat behavior may see improvements in that behavior, but another behavior, such as calling out without permission, may increase. Randomizing the target behavior may lead to broader improvements because all behaviors are under control of the contingency (Skinner et al., 1996). For example, specific behaviors, such as calling out, out of seat, and kicking/hitting others, may be written on slips of paper, chosen at the end of the period, and students are rewarded if they meet the criteria for the target behavior.

Positive Variations of Group Contingencies

Positive attention and praise for appropriate classroom behavior have proven to be effective behavior management strategies for preventing disruptive behavior and increasing on-task behavior (Craft, Alber, & Heward, 1998; Hall, Lund, & Jackson, 1968; Madsen, Becker, & Thomas, 1968; Mesa, Lewis-Palmer, & Reinke, 2005). Though the traditional methods of group contingencies, including the GBG, involve rewarding a group for exhibiting desired behavior, the procedures often require teachers to scan for negative student behaviors (i.e., Christ & Christ, 2006; Kelshaw-Levering et al., 2000; Murphy et al., 2007). For instance, Kelshaw-Levering et al. (2000) used an interdependent group contingency in a second-grade classroom where the teacher gave points for off-task behaviors, inappropriate vocalizations, out-of-area behavior, and noncompliance. In a classic study by Wolf, Hanley, King, Lachowicz, and Giles (1970), the procedures involved taking points away for out-of-seat behavior. In a modification of the GBG during lunch, McCurdy, Lannie, and Barnabas (2009) also gave points for rule
violations. Although these methods have proven effective, they direct teachers’ attention toward the undesirable behaviors. In fact, the procedures of the GBG have not been found to lead to an increase in teachers’ use of positive attention or praise (Lannie & McCurdy, 2007). This is important to note as teachers often use negative statements more often than positive statements (Strain, Lambert, Kerr, Stagg, & Lenkner, 1983; White, 1975), especially towards children with challenging behavior (Van Acker, Grant, & Henry, 1996).

As PBIS becomes more supported, using a positive variation of group contingencies may align more closely with the desired school climate as well as with the ideas that teachers and staff are being trained (Wright & McCurdy, 2011). Positive variations of group contingencies are less common but have been used throughout the years. Hansen and Lignugaris/Kraft (2005) and Morrison and Jones (2006) used a group contingency to increase students’ positive peer statements. In both studies, data were collected on both positive and negative peer statements, but only positive peer statements were publically praised. The GBG has also been modified from traditional methods of negative scanning to positive scanning or a combination. For example, Fishbein and Wasik (1981) modified the GBG by awarding points to teams for engaging in desired behaviors during library time. In another modification of the GBG, Patrick, Ward, and Crouch (1998) awarded points for positive social behaviors (i.e., high fives, “Good job!”) and removed points for negative social behaviors (i.e., arguing, laughing at another student). McGoey, Schneider, Rezzetano, Prodan, and Tankersley (2010) used a version of GBG, called Good Behavior Game Plus Merit. The traditional methods of awarding points to students for breaking classroom rules were used; however, the teachers also
awarded bonus points to students for appropriate behavior that could cancel these points.

Recent studies have begun to compare the effectiveness of positive variations of group contingencies to the traditional methods. Tanol, Johnson, McComas, and Cote (2010) compared two versions of the GBG with kindergarten students. Students were divided into teams and the classroom rules were explained to them. During the first version, also called the response cost version, the teams began with stars that were removed when a rule was broken. The teacher was trained to explain the rule that was broken and then deliver a praise statement to a member on another team following the rules. During the reinforcement version, teams started with no stars and were given stars and teacher praise for following the rules. Teams who were not following the rules were not given attention. At the end of the period, teams who got three or more stars were given a reward. Student and teacher behaviors were measured. Both variations were effective in improving student behavior, with the reinforcement version being slightly more effective for some students. Additionally, the teachers were more likely to praise students during the reinforcement sessions than the response cost sessions. However, it is important to note that the teachers were trained specifically when and how to react to rule violation and rule following during each version. Teachers also preferred the reinforcement version because it created a more positive environment.

Another recent study by Wright and McCurdy (2011) examined a positive variation of the GBG, called the Caught Being Good Game (CBGG) on kindergarten and fourth-grade students’ on-task and off-task behaviors. Students were divided into four teams. During the GBG version, the teacher tallied points for disruptive behavior. Teams earned a reward for receiving fewer points than a mystery criterion and also had
an opportunity to earn a weekly reward if total points for the week were less than the criterion. During the CBGG version, the teacher was signaled by an audio cue at random intervals (VI-4-min) to give teams points if all members were following the classroom rules. They also had the opportunity to earn weekly rewards if they earned more points than the mystery criterion. Using a withdrawal design, they found both contingencies were effective in decreasing disruptive behavior and increasing on-task behavior. Neither contingency was more effective than the other. Additionally, both contingencies were found equally acceptable by the students and teachers.

**Current Study**

The current study sought to extend the findings of Wright and McCurdy (2011) in several ways. First, the study examined the differential effects of the two types of contingencies (i.e., monitoring positive behavior vs. monitoring negative behavior) on teacher behavior, including positive and negative statements, in addition to student behavior. It was hypothesized that scanning for appropriate behaviors during the CBGG version, rather than scanning for negative behaviors as done with the GBG, would lead to an increase in the use of positive teacher statements.

As Wright and McCurdy (2011) stated in their limitations, the different versions (GBG versus CGBB) create different schedules of reinforcement and punishment, which may produce different rates of responding over time. The GBG utilizes a fixed ratio (FR) schedule of punishment. For each incident of misbehavior, the teacher delivers a point. The GBG also utilizes differential reinforcement of low rates (DRL) by delivering a reinforcer at the end of the class period when target behaviors occur less than a set criterion. The CBGG utilizes a variable interval (VI) momentary differential
reinforcement of other behaviors (DRO) schedule where points are delivered on a variable interval schedule when students are not displaying the undesired target behaviors. The VI schedule means that reinforcement is delivered at random intervals. This type of schedule leads to a consistent rate of responding over time. In contrast, the FR schedule of punishment yields a high rate of responding that may not be as consistent over time (Cooper et al., 2007). While the Wright and McCurdy (2011) study implemented each condition once, the current study replicated each condition to allow for the comparison of the effects of the contingencies and thus, different reinforcement schedules over time. Additionally, the study utilized four classrooms, as opposed to two classrooms, to allow for counterbalancing.

Furthermore, Wright and McCurdy (2011) utilized an audible computer signal during the CBGG to prompt the teachers to scan the classroom for points. As stated in their limitations, this may have served as a stimulus for students to follow the rules, possibly limiting the effects of the intervention. The current study used a private signal so that students were unaware of when their behavior was being recorded. Overall, it was hypothesized that the CBGG would lead to sustained positive student behavior over time as well as an increase in positive statements delivered by the teacher.

Method

Participants and Setting

The study was conducted at two elementary schools in the Midwestern United States in four classrooms during a time of the day during which the teachers experienced problem behaviors most often. Both schools had high economic disadvantage levels (77.6% and 99.5%) and a large number of students with disabilities (16% and 16.1%,
respectively). In order to ensure each classroom would benefit from the intervention, the primary researcher completed three preliminary classroom observations. Across observations, the median level of disruptions, including off-task motor and off-task verbal, must have occurred at least 20% of the intervals to be included in the study. Additionally, the Council for Exceptional Children (1987) has established a 90% criterion as indicative of meaningful engagement for classroom learning. Across preliminary observations, median levels of classroom engagement must have occurred for no more than 80% of intervals for inclusion.

Four classrooms met the inclusionary criteria, including two kindergarten classes, one class combined of both first and second graders, and one second grade class. The two kindergarten classrooms were taught by the same teacher (AM and PM). Before the study began, a letter was sent home with each student. The letter explained the purpose of the study. Given that the interventions were classwide, all students participated in the interventions and earned the rewards. However, parents had the option to sign and return the letter if they did not want observers to collect data on their students’ behavior. Of the four classrooms, one student returned a signed form. There was a total of 19, 17, 16, and 20 student participants, for classrooms 1, 2, 3, and 4 respectively.

The primary researcher, an advanced doctoral student in school psychology, served as the behavioral consultant, provided teacher support for implementation of the interventions, monitored data, graphed progress, and trained graduate students to use appropriate data collection tools. Graduate students in school psychology served as research assistants by collecting data and supporting the teacher with implementation as needed. Teachers served as the implementation agents in order to make the interventions
as natural as possible and allow them to develop the implementation skills.

**Materials**

Materials needed to implement the interventions included a rules poster, MotivAider timer, intervention script, and a Teacher Collected Data sheet. A MotivAider is a small timer that is set to fixed or random intervals and can be worn on a belt loop or placed in a pocket. At the end of each interval, the timer vibrates. Prior to the intervention, a script for the intervention that was implemented in the classroom first was provided to the teacher to introduce the contingency the first day as well as a faded script to introduce the contingency on the following days (Appendix B). The teacher was also given a script with step-by-step directions for daily implementation (Appendix C). The consultant reviewed the steps with the teacher and answered any questions. The script for the second intervention was introduced and reviewed with the teacher before its implementation so that teacher was clear about the distinction between the steps of the two interventions. The teacher was also trained to complete a self-report method of adherence by using the Teacher Collected Data sheet (Appendix D).

The teacher and consultant met before implementation to divide the classroom into two to five teams, depending on teacher’s preference and class size. Similar to other studies using the GBG procedures (e.g., Donaldson, Vollmer, Krous, Downs, & Berard, 2011) teachers identified students who are more likely to engage in problem behaviors and those students were placed on different teams in order to prevent a team from having a greater probability of engaging in disruptive behavior.

Rules included three to five positive statements and were developed collaboratively by the primary researcher and the teacher based on preliminary
observation data to reflect the classroom concerns. For example, if out-of-seat behavior was a common classroom disruption, a rule may state that students must raise their hand and ask for permission to get out of their seats.

Also, before the intervention began, the consultant met with the teacher in order to create a list of acceptable rewards. Examples included pencils, tickets to turn in for no morning work, and certificates. The consultant presented the list to the class and, if acceptable to the teacher, the students were able to add rewards to the list with the teacher’s input (Appendix E). Students ranked the rewards on this list with one representing the reward they would like to earn the most. The consultant then reviewed each list and gave three points to the reward that was chosen as number one, two points to the reward chosen as number two, and one point to the reward chosen as number three. The three or four rewards with the most points were used as the rewards during the intervention. Additionally, the primary researcher reviewed the reward assessments to ensure all students have one of their top three choices as possible rewards. The primary researcher added one of the student’s top three choices if one of his or her preferred rewards was not already in the mix.

**Dependent Measures and Data Collection**

Both student and teacher behaviors were observed throughout the targeted class periods every day. Observations took place throughout the entire class period; however, because there were occasional interruptions to the usual schedule (i.e., fire drill and holiday parties) that prevented observations from lasting the entire class period, observations were a minimum of 10 min. Observers were trained in a modified version (Appendix A) of the Behavioral Observation of Students in Schools (BOSS; Shapiro,
2004), which was designed to measure student engagement and off-task behaviors. Teacher-directed instruction was removed and teacher positive and negative statements were added in order to observe the effects on teacher behavior. The student variables remained the same; however, the system was modified in order to monitor classwide behavior. Rather than observing one student and a peer comparison each fifth interval, a different student was observed each interval. The primary researcher and research assistants observed and coded the behavior of the first student and moved to the next student at the end of the interval. This pattern continued until all students were observed and repeated until the end of the observation. When two observers were coding behavior for inter-observer agreement, before they began the observation, they discussed which student they would observe first as well as the pattern they would follow (i.e., row by row, front to back). The student not included in the study was skipped during observations.

**Student Engagement.** Engagement was coded as either active engaged time (AET) or passive engaged time (PET). Active engagement was defined as actively responding to academic task. Examples of active engagement included writing and reading aloud and talking to a teacher or peer about the assigned task. Passive engagement occurred when the student was oriented to the academic activity and not off-task or actively engaged. Examples of passive engaged behavior included reading a book or looking at the board. Engagement was coded using momentary time sampling, meaning that at the beginning of each 15-second interval, the observer coded whether or not the student was engaged at that moment (Shapiro, 2004).

**Disruptions.** Disruptions included two categories of off-task behavior including
off-task verbal and off-task motor. Off-task verbal included any verbalization that is made without teacher permission. This included humming, singing, and talking to peers. Off-task motor included any movement not toward an assigned task. Examples included being out of one’s seat and playing with a pencil or other materials inside or on one’s desk. Disruptions were observed using partial interval recording, meaning that the observer recorded the behavior if the target student was off-task at any point during the 15-second interval (Shapiro, 2004).

Through an interview as well as teacher and observer data obtained through frequency counts during preliminary observations before the intervention began, the teacher and consultant determined which off-task category, off-task verbal or off-task motor, was the highest priority classroom concern. This variable was graphed separately and used to determine phase changes. Classrooms 1, 2, and 3 prioritized off-task motor while classroom 4 prioritized off-task verbal.

**Teacher Statements.** The modified observation code included teacher behavior, including positive and negative statements. The definitions were similar to those used by Tanol et al. (2010). Positive statements were defined as any verbal statement that acknowledges rule following. For example, positive statements were coded when a teacher acknowledged a student for raising his or her hand. Negative statements were defined as any verbal statement made by teacher that expresses disapproval of a student’s behavior or lack of rule following, including redirecting the student. For instance, when a teacher told a student to stop hitting his pencil against his desk, negative statement was coded. Teacher behaviors were recorded throughout the class period using partial interval sampling during all intervals.
Inter-Observer Agreement

The primary researcher also served as the primary observer for data collection, introducing the possibility of observer bias. In order to monitor the reliability and validity of data, inter-observer agreement data were collected. Graduate students in school psychology, including the primary researcher, served as data collectors. Prior to the implementation of baseline, observers were trained in data collection. The primary researcher reviewed the definitions as well as the time sampling methods for each variable. Each observer co-observed with the primary researcher until an inter-observer agreement (IOA) of at least 90% was reached. Percentage agreement was calculated for 22% of observed sessions using interval-by-interval method by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100.

Table 1 represents the IOA means and ranges for each variable. Inter-observer agreement was high for each variable, validating the consistency of data collection.

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>94.37%</td>
<td>84-100%</td>
</tr>
<tr>
<td>Disruptions</td>
<td>94.82%</td>
<td>83-100%</td>
</tr>
<tr>
<td>Off-Task Motor</td>
<td>96.03%</td>
<td>88-100%</td>
</tr>
<tr>
<td>Off-Task Verbal</td>
<td>97.38%</td>
<td>91-100%</td>
</tr>
<tr>
<td>Positive Statements</td>
<td>99.17%</td>
<td>97-100%</td>
</tr>
<tr>
<td>Negative Statements</td>
<td>95.72%</td>
<td>83-100%</td>
</tr>
</tbody>
</table>
Research Design

An ABCBC design was used in classrooms 1 and 2 and an ACBCB design was used in classrooms 3 and 4 in order to counterbalance the order of condition presentation. For the ABCBC design, data were first collected during the baseline (A) phase where typical classroom management procedures were followed. After baseline, the GBG was introduced (B). Then, during the third phase, the GBG was removed and the CBGG is introduced (C). After the CBGG was implemented, the teacher then reintroduced the GBG (B). During the last phase, the CGBB was reintroduced. After the completion of the last phase, the teacher was told that he or she may implement the desired intervention. Classrooms 3 and 4 followed a similar pattern except the GBGG (C) was implemented during the second and fourth phases and the GBG (B) was implemented during the third and fifth phases. Data collection procedures were the same during all phases (A, B, and C). The off-task variable determined to be highest classroom concern was used to determine phase changes. Stable trends, or as stable as possible, indicated phase changes. A minimum of five and a maximum of eight data points were collected in each phase, or as dictated by the data.

Experimental Conditions

Baseline (A). During baseline, observations were conducted according to the observation procedures and definitions described above. The teacher was asked to conduct class using typical management strategies.

Training. The primary researcher reviewed the intervention with each teacher individually before implementation for about 20 min. She read through the script and
answered any questions. The consultant also reviewed the rules that were established as well as examples and non-examples of each rule to make sure the teacher understood when he or she should give points to teams for rule following (CBGG) or rule violation (GBG). The teacher participated in training sessions, in which he or she was given feedback after each session regarding his or her implementation, until he or she reached 100% adherence during the first phase of each intervention. The feedback session involved either a short discussion or an email to the teacher regarding the step that was not administered appropriately. If the teacher implemented the intervention with 100% adherence on the first day of implementation, no training sessions were needed. The teacher for classroom 2 had two feedback sessions for the CBGG, and the teachers for classrooms 3 and 4 had feedback sessions for both the GBG and the CBGG. For example, when adherence was not met for the step involving administering the points for rule violation during the GBG, an email was sent to a teacher to remind her of the criteria for receiving a point. The primary researcher reviewed examples and non-examples of behaviors that earned points.

**Intervention.** The teacher introduced the intervention as a game. On the first day, the teacher utilized the introduction script (Appendix B). The teacher introduced the rules, described the procedures of the ‘game’, and explained the teams. The classroom rules were posted in the front of the classroom and students were told they needed to comply with the rules in order to receive special rewards. After reviewing each rule, the teacher discussed specific examples. The students were informed of the possible rewards. The students were also given the opportunity to ask questions. A common question included asking whether the goal was to earn points or not. Daily, the teachers
took a few minutes to introduce the game, remind students of the rules, explain which
game was going to be played, and then began the game immediately afterwards
(Appendix C).

**Good Behavior Game (B).** During the GBG, whenever the teacher noticed a
student violating a rule, he or she marked a point for the student’s team on the tally sheet
found on Teacher Collected Data sheet (Appendix D). Teachers had a laminated tally
sheet attached to their lanyard or in their pocket so they could access it as quickly as
possible. Points were assigned privately in order to minimize potential disruption caused
by public posting. At the end of the period, the teacher chose a slip of paper indicating
the maximum number of points the teams could receive. The teams that had fewer points
than this mystery criterion earned a reward. The teacher explained which teams met the
criterion, immediately chose a random slip of paper indicating the mystery reward, and
delivered the reward to the appropriate teams. Each team also had a chance to earn a
weekly reward if they received fewer points than the mystery weekly criterion (Lannie &
McCurdy, 2007; McCurdy et al., 2009). The teams’ daily points were posted in the
classroom as well as the cumulative total for the week.

**Caught Being Good Game (C).** During the CBGG, the teacher wore a vibrating
timer (i.e., MotivAider) set to random intervals determined by the consultant and the
teacher. At random times throughout the period, the teacher was prompted by the
MotivAider to award points to teams for following classroom rules. For example, if the
intervals were set to 5 min (VI = 5 min), the timer vibrated at one, two, three, four, or five
minutes. When it vibrated, the teacher scanned the classroom and each team with all
members on-task and following all classroom rules, earned a point. Similarly to the
GBG, points were assigned privately on the tally sheet located on the Teacher Collected Data sheet (Appendix D). At the end of the class period, the teacher chose a slip, indicating the number of points needed to earn the reward. The teacher chose a slip from the reward envelope or jar and delivered the reward to the teams that met the criterion. The teams also had the opportunity to earn a weekly reward if they received more points than the weekly mystery criterion.

**Criterion for Rewards**

Before implementation, the teacher and consultant met to discuss the maximum number of rule violations and minimum number of points awarded for following classroom rules that was acceptable in order for a team to earn a reward. Using initial baseline data, the teacher and consultant discussed what level of reduction or increase that was required. For example, during the GBG, the teacher of classrooms 1 and 2 decided that five rule violations was acceptable. In this case, the numbers one, two, three, four, and five were written on slips of paper that were pulled from the mystery criterion jar. Teachers of classrooms 3 and 4 decided that ten and four violations were acceptable, respectively. During the CGBB, each teacher decided teams need to earn at least two points to earn a reward. In this case, the numbers two, three, and four were written on the slips of paper.

**Adherence**

Treatment adherence data were collected using two methods. First, the teachers were asked to complete the Teacher Collected Data sheet (Appendix D) each day. For each intervention component implemented, the teacher placed a checkmark next to that
step. The number of steps implemented was divided by the total number of steps and multiplied by 100 to determine the percentage of adherence.

Also, the primary researcher or research assistants completed an intervention checklist (Appendix F) during at least 20% of observed intervention sessions across phases. The observer checked each component off the list as it is completed. In order to ensure the teacher gave points accurately for following the rules, the observer simultaneously administered points to students for rule violation (GBG) or rule following (CGBB). At the end of the class, the observer compared his/her points the teacher’s points and only checked this step when 80% or higher agreement was met. Whenever adherence was not delivered with 100% accuracy, the observer reviewed the intervention steps with the teacher to ensure comprehension. Although this step was completed for most adherence checks, there were a couple observations when this was not possible (i.e., researcher was not able to collect data during complete class period, and thus, comparison of points was not applicable).

Although we aimed to have teachers complete the Teacher Collected Data sheet each day, there were some days this did not happen for various reasons (i.e., the teacher forgot or the researcher did not give enough copies to the teachers). The teachers completed the sheets during 77%, 79%, 90%, and 58% of sessions for classrooms 1, 2, 3, and 4, respectively. Each teacher indicated 100% adherence across sessions, with the exception of classroom 4, who averaged 99%.

The researcher collected adherence data using the intervention checklists during 40% of sessions across classrooms. Table 2 shows adherence levels for all classrooms. Overall, adherence was high for both interventions, averaging 91-94% across classrooms.
There were no significant differences in adherence between the two interventions, indicating both interventions were implemented with integrity.

Table 2

*Adherence Results by Classroom*

<table>
<thead>
<tr>
<th>Classroom</th>
<th>% of sessions</th>
<th>Caught Being Good Game</th>
<th>Good Behavior Game</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>43%</td>
<td>33%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>100%</td>
<td>89%</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>100%</td>
<td>60-100%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>36%</td>
<td>40%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>92%</td>
<td>97%</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>83-100%</td>
<td>83-100%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>50%</td>
<td>40%</td>
<td>45%</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>90%</td>
<td>91%</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>60-100%</td>
<td>60-100%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>38%</td>
<td>41%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>92%</td>
<td>93%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>80-100%</td>
<td>80-100%</td>
<td></td>
</tr>
</tbody>
</table>

**Intervention Acceptability**

At the end of the study, the teacher and students evaluated the interventions by completing a scale adapted from Erhardt, Barnett, Lentz, Stollar, and Reifin (1996) and Martens, Witt, Elliot, and Darveaux (1985). Results were used to examine the acceptability of the interventions and which, if either, version was more acceptable. Using a 5-point scale with 1 meaning “strongly disagree”, 2 meaning “disagree”, 3 meaning “not sure”, 4 meaning “agree”, and 5 meaning “strongly agree”, the teacher was
asked to rate items regarding the ease of implementation, willingness to use the interventions in the future, and the impact of the interventions on the students (Teacher Social Validity Survey, Appendix G). Students evaluated the interventions on items regarding the classroom rules, the intervention procedures, and the reinforcers (Student Social Validity Survey, Appendix H). Teachers and students were asked which version they prefer.

While the surveys reveal meaningful information about the teachers’ preferences and intervention feasibility, they are subjective and have the potential to be biased (Kennedy, 2005). Sustainability may be a more objective source for gauging social validity. Kennedy (2005) defined sustainability as “an index of whether the procedures and outcomes of an experiment continue once the research is completed and the researchers are no longer involved” (p. 229). Therefore, after the research was complete, teachers had the opportunity to choose which intervention they would like to implement in their classroom. The primary researcher continued to supply any materials needed for at least two weeks following the study and observed the classroom to see which, if any, intervention continued to be implemented in the classroom. Continued implementation of an intervention provided additional evidence to the acceptability of the interventions.

Results

In order to examine the effectiveness of the interventions, student behavior (including engagement, disruptions, and the classroom prioritized target variable) and teacher statements (positive and negative) were examined using several methods. Visual analysis was the primary method utilized, which involves analyzing three dimensions of graphed, time series data: trend, level, and variability (Kennedy, 2005). Percentage of
non-overlapping data (PND) was used a secondary method. PND was calculated using the methods suggested by Scruggs and Mastropieri (1998), which involves determining the highest (if desired effects are an increase in behavior) or lowest (if desired effects are a decrease in behavior) point in baseline and the number of points that are above or below that point in the subsequent phase. For example, if a classroom’s highest engagement point during baseline is 75% of intervals, then one would divide the number of points above 75% during the first intervention phase by the total number of intervention points in that phase. Then, one would continue this pattern with subsequent adjacent phases (i.e., B-C, C-B, B-C). The higher the PND, the more effective the intervention.

Specifically Scruggs and Mastropieri (1998) describe interventions with PND higher than 90% as “very effective”, 70%-90% as “effective”, 50%-70% as “questionable”, and below 50% as “ineffective”. PND reveals meaningful information; however, there are several limitations including the lack of sensitivity to trends and the intensity of effects. Student and teacher social validity results are also reported.

**Student Behavior**

Graphs of student engagement and disruptions in classrooms as well as each classroom’s prioritized target variable are included. Visual analysis shows low levels of engagement across classrooms during baseline. Each classroom had an engagement average significantly less than the 90% criterion from the Council for Exceptional Children (1987). All classrooms showed an increase in the level of engagement when the GBG and CBGG were in place. Also, each classroom’s level of disruptions, including the prioritized classroom variable, showed drastic improvements with the interventions in place.
Classroom 1. Engagement and disruption data for classroom 1 are shown in Figure 1. During baseline, disruptions were variable and the average was 31.29% (SD = 11.31) and engagement level was low (M = 52.86%, SD = 8.11), based on criterion set forth by previous research (Council for Exception Children, 1987). When the GBG was introduced, there was an immediate decrease in disruptions (M = 8.00%, SD = 4.20) and an increase in engagement (M = 78.67%, SD = 6.44). When the CBGG was implemented, the engagement level slightly decreased (M = 70.63%, SD = 3.58) and disruption level increased (M = 18.38%, SD = 4.78) with a stable trend. The GBG was reintroduced and data remained consistent for both engagement (M = 76.33%, SD = 3.83) and disruptions (M = 14.67%, SD = 4.37). The CBGG was reinstated and again, data remained consistent (engagement, M = 74.33%, SD = 4.13; disruptions, M = 14.17%; SD = 1.47). Data seemed to slightly worsen after the first intervention phase, but remained improved from baseline.
Two weeks after the study ended, follow-up data were gathered. The classroom teacher was told that it was her choice to implement one of the games or to discontinue using a contingency system in her classroom. The researcher observed the classroom according to the same criteria used to evaluate adherence to the interventions. The only step implemented during this observation was the posting of the classroom rules. Data reflected levels similar to baseline levels (engagement, 57%; disruptions, 35%), emphasizing the need for the interventions.

Classroom 1 prioritized off-task motor (shown in Figure 2). Baseline data were variable ($M = 19.14\%$, $SD = 11.32$). When the GBG was implemented, there was an immediate decrease in level ($M = 5.50\%$, $SD = 2.88$) and a stable trend. The level increased ($M = 10.25\%$, $SD = 4.30$) when the CBGG began and remained improved from

---

Figure 1. Classroom 1 engagement and disruptions.
baseline throughout the rest of the phases (GBG, $M = 7.17\%, SD = 2.64$; CBGG, $M = 7.00\%, SD = 2.97$). During follow-up, the off-task motor disruptions increased to 21% of intervals, a steep increase from the intervention phases.

![Figure 2. Classroom 1 off-task motor.](image)

Table 3 shows the PND for classroom 1. PND from baseline to the first phase, GBG, was high for each variable (engagement, disruptions, and off-task motor), indicating the intervention was successful in improving student behavior (Scruggs & Mastropieri, 1998). During each of the following phases, PND was low, indicating little change in the on-going effectiveness of the interventions.

Table 3

*Classroom 1 PND*
Classroom 2. Visual results for classroom 2 behavior are shown in Figure 3. During baseline, engagement was variable and averaged 52.17% ($SD = 8.04$) and disruptions trend was stable and averaged 29.00% ($SD = 3.46$). When the first intervention, the GBG, was introduced there was an immediate improvement in both engagement and disruptions. Engagement level increased ($M = 77.40\%, SD = 7.83$) and disruption level decreased ($M = 10.60\%, SD = 3.71$) both with a relatively stable trend. The CBGG was then introduced and the engagement level slightly decreased ($M = 65.25\%, SD = 4.17$) and disruption level slightly increased ($M = 19.63\%, SD = 3.74$). The GBG was then reintroduced and data remained relatively stable from the previous phase (engagement, $M = 68.20\%, SD = 5.22$; disruptions, $M = 15.00\%, SD = 3.74$).

During the last phase, the CBGG was reinstated and the level remained consistent for engagement ($M = 69.67\%, SD = 3.61$) and disruptions ($M = 15.33\%, SD = 3.61$). Overall, similar to classroom 1, the first phase had the most dramatic effects, followed by sustained, but less dramatic, improvement from baseline throughout the rest of the phases.
The teacher of classroom 2 was the same teacher as classroom 1. During follow-up, the same step, posting of the classroom rules was implemented, but no other steps. Again, data revealed levels similar to baseline, emphasizing the need for intervention.

Classroom 2 prioritized off-task motor. During baseline the trend was stable and averaged 20.50% (SD = 3.73). When the GBG was introduced, there was an immediate decrease in level ($M = 5.60\%, SD = 2.07$) with a stable trend. The CBGG was then put in place and a slight increase was seen ($M = 9.63\%, SD = 2.45$) with a slight increasing trend. Level and trend remained consistent throughout the final phases (GBG, $M = 8.20\%, SD = 2.39$; CBGG, $M = 9.17\%, SD = 1.17$). Similar to engagement and disruption data, during follow-up, motor disruptions increased to baseline levels.

Figure 3. Classroom 2 engagement and disruptions.
The PND statistics again add evidence for the effectiveness of both interventions, GBG and CBGG, on students’ engagement and disruptions. When the GBG was first implemented after baseline, PND was 100% for both engagement and disruptions, indicating improvement during all intervention sessions, reflecting highly effective treatment (Scruggs and Mastropieri, 1998). When the class transitioned from GBG to CBGG, PND was 0% for both engagement and disruptions, indicating the CBGG did not further improve behavior. When the GBG was reinstated, PND was low for engagement and disruptions (40% and 20%, respectively), indicating little change from the previous phase of the CBGG. Finally, PND was again 0% when the GBG was reintroduced. PND for off-task motor was similar to engagement and disruptions, very high when the GBG was introduced (100%), and very low throughout the rest of the phases (range, 0-20%).

Figure 4. Classroom 2 off-task motor.
The high PND during the first phase and low PND during the following phases add further evidence that the effects remained consistent throughout the phases.

Table 4

*Classroom 2 PND*

<table>
<thead>
<tr>
<th></th>
<th>Baseline &gt; GBG</th>
<th>GBG &gt; CBGG</th>
<th>CBGG &gt; GBG</th>
<th>GBG &gt; CBGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>100.00%</td>
<td>0.00%</td>
<td>40.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Disruptions</td>
<td>100.00%</td>
<td>0.00%</td>
<td>20.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Off-Task</td>
<td>100.00%</td>
<td>0.00%</td>
<td>20.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Motor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In summary, the first phase, the GBG, seemed to produce very immediate and intense effects, perhaps due to the newness of the intervention. Afterwards, the results seemed to level-off while still having improvement from baseline.

*Classroom 3.* Figure 5 shows the results of the interventions for classroom 3 on student engagement and disruptions. During baseline, classroom 3 had an engagement average of 46.90% ($SD = 6.46$) with some variability and a disruption average of 43.71% ($SD = 5.88$) with an increasing trend. When the CBGG was implemented, there was an immediate increase in the level of engagement ($M = 68.63\%, SD = 4.10$) and a stable trend. The level of disruptions also decreased ($M = 28.38\%, SD = 4.69$) and had a stable trend. The GBG was then implemented and the data were slightly more variable; however, the level of engagement further increased ($M = 81.00\%, SD = 10.31$) and the level of disruptions further decreased ($M = 18.33\%, SD = 7.79$). Upon reinstatement of the CBGG, the level of engagement decreased ($M = 73.00\%, SD = 7.04$) with a stable
trend and the level of disruptions increased ($M = 23.00\%, SD = 4.52$) with a slight increasing trend. While data worsened from the previous GBG phase, the data remained at improved levels from baseline. The GBG was implemented for the final phase and data slightly improved from the prior phase for engagement ($M = 77.17\%, SD = 6.37$) and disruptions ($M = 17.17\%, SD = 4.07$).

During follow-up, the only step still implemented was the posted classroom rules. The rest of the intervention steps were not implemented. The data show engagement decreased and disruptions increased, adding evidence to the need and effectiveness of the interventions. However, after the observation was complete, the researcher observed the teacher carrying out the GBG procedures when the students were gathering their backpacks and other materials to go home. Data were not collected during this time.

Figure 5. Classroom 3 engagement and disruptions.
Classroom 3 prioritized the off-task motor variable. During baseline, the class averaged 31.99% ($SD = 5.21$) of intervals, meaning the students were either out of their seats, playing with objects on their desks, or making a movement that is not related to their assignment. When the first CBGG was introduced, there was an immediate improvement and the level decreased ($M = 18.75\%, SD = 5.75$). Upon implementation of the first GBG further improvement was noticed ($M = 12.67\%, SD = 6.00$) with a stable trend. After the second CBGG was implemented, the level remained about the same ($M = 13.67\%, SD = 3.93$) with a slight increasing trend. The second implementation of the GBG led a similar level ($M = 11.67\%, SD = 6.12$) with more variability and no discernible trend. During the follow-up observation, off-task motor increased to 23% intervals, a level higher than the mean during each intervention phase.

*Figure 6.* Classroom 3 off-task motor.
The effectiveness of the interventions on student engagement and disruptions is also supported by the PND statistics outlined in Table 5. PND during the first phase of the CBGG was 100.00%, indicating it was “very effective” (Scruggs & Mastropieri, 1998) and improvement was demonstrated during all sessions for both engagement and disruptions. When the GBG was implemented, PND remained relatively high (engagement, 66.67%; disruptions, 77.78%), adding evidence that the GBG further improved student behavior. PND for the last two phases remained below 50%, suggesting little change in the effectiveness. PND for off-task motor, the prioritized classroom variable, was relatively high during the first intervention phase (75.00%), indicating the CBGG was “effective” (Scruggs & Mastropieri, 1998). The subsequent phase changes had very low PND levels, indicating little change in the effectiveness.

### Table 5

**Classroom 3 PND**

<table>
<thead>
<tr>
<th></th>
<th>Baseline &gt; CBGG</th>
<th>CBGG &gt; GBG</th>
<th>GBG &gt; CBGG</th>
<th>CBGG &gt; GBG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>100.00%</td>
<td>66.67%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Disruptions</td>
<td>100.00%</td>
<td>77.78%</td>
<td>0.00%</td>
<td>33.33%</td>
</tr>
<tr>
<td>Off-Task Motor</td>
<td>75.00%</td>
<td>33.33%</td>
<td>0.00%</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

Visual analysis and PND statistics reveal that both interventions improved student behavior for classroom 3; however, the GBG seemed to be slightly more effective in increasing student engagement and decreasing disruptions.

**Classroom 4.** Similar to the other classrooms, Figure 7 shows that during
baseline, classroom 4 had an engagement level below the criteria set by the Council for Exceptional Children (1987; \( M = 63.50\% \), \( SD = 4.37 \)) and a disruption level averaging 29.00\% (\( SD = 3.10 \)). When the CBGG was put in place, there was an immediate increase in the level of engagement (\( M = 83.00\% \), \( SD = 4.55 \)) and decrease in level of disruptions (\( M = 15.57\% \), \( SD = 3.78 \)). When the GBG was put in place, the engagement level remained high (\( M = 84.67\% \), \( SD = 5.32 \)), but there was a decreasing trend. The level of disruptions remained similar to the CBGG phase (\( M = 12.83\% \), \( SD = 3.60 \)). The CBGG was reintroduced and an increasing trend for engagement was noticed (\( M = 81.50\% \), \( SD = 6.16 \)). The level of disruptions remained stable (\( M = 14.17\% \), \( SD = 2.64 \)). Finally, the GBG was reintroduced and engagement level remained stable (\( M = 80.17\% \), \( SD = 3.43 \)) while the level of disruptions had a slightly increasing trend (\( M = 17.17\% \), \( SD = 4.40 \)).

![Figure 7. Classroom 4 engagement and disruptions.](image-url)
During the follow-up observation, the teacher was implementing the CBGG with 100% adherence. Data reveal that the engagement level remained at 88% of intervals and the disruption level remained at 14% of intervals. Students continued to respond positively to the CBGG over time.

Classroom 4 prioritized off-task verbal, which included the students calling out without permission or making any noises. During baseline, the level was 29.00% (SD = 3.25). During the intervention phases, the level decreased and remained relatively consistent (CBGG, $M = 5.71\%$, $SD = 2.43$; GBG, $M = 7.00\%$, $SD = 4.15$; CBGG, $M = 7.83\%$, $SD = 0.98$; GBG, $M = 9.17\%$, $SD = 3.76$). During the follow-up phase, off-task verbal remained at a similar level, 7% of intervals.

![Graph](image)

Figure 8. Classroom 4 off-task verbal.

PND statistics again add evidence to the effectiveness of the interventions on
student behavior. PND was 100% from baseline to the first phase for all three student variables (e.g., engagement, disruptions, and off-task verbal), meaning no intervention points overlapped with baseline. Throughout the rest of the phases, PND remained low from one phase to the next, indicating little change in the effectiveness (Scruggs & Mastropieri, 1998).

Table 6

Classroom 4 PND

<table>
<thead>
<tr>
<th></th>
<th>Baseline &gt; CBGG</th>
<th>CBGG &gt; GBG</th>
<th>GBG &gt; CBGG</th>
<th>CBGG &gt; GBG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>100.00%</td>
<td>33.33%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Disruptions</td>
<td>100.00%</td>
<td>16.67%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Off-Task Verbal</td>
<td>100.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>16.67%</td>
</tr>
</tbody>
</table>

Overall, both interventions improved the behavior of the students in classroom 4.

Visual analysis shows the students responded well to the intervention over time as demonstrated by the follow-up phase.

The GBG and the CBGG both improved student behavior for all classrooms. When the interventions were in place, student engagement increased and disruptions decreased. Additionally, each classroom’s prioritized target variable (i.e., off-task motor or off-task verbal) showed improvements with the use of the CBGG and GBG.

Teacher Statements

The impact of the interventions on teacher statements was also measured. Similar to the research found by Strain et al. (1983) and White (1975), the teachers in the study
used more negative statements, which included any statement of disapproval toward a student for student behavior, than positive statements which included any statement acknowledging rule following. Included graphs display the level of teachers’ positive and negative statements during baseline, each of the intervention phases, and follow-up phase for each classroom.

Classroom 1. Figure 9 shows the level of positive and negative statements for classroom 1. The average level of negative statements for classroom 1 was at 19.17% (SD = 6.87) during baseline. When the GBG was introduced, the level of negative statements decreased (M = 9.00%, SD = 5.37), but had an increasing trend. Throughout the rest of the phases, the level of negative statements remained fairly consistent (CBGG, M = 13.38%, SD = 7.23; GBG, M = 9.67%, SD = 3.67; CBGG, M = 8.83%, SD = 4.17). The average level of positive statements during baseline was 3.86% (SD = 2.41). Throughout the rest of the phases, this level remained consistent. During the follow-up observation during which only posting of the classroom rules was observed, the negative statements rose to 19% of intervals and no positive statements were observed. Overall, there appeared to be a decrease in all teacher statements regarding student behavior. This may have been reflective of the improved behavior. There does not appear to be a noticeable difference in teacher statements between interventions.
Figure 9. Classroom 1 teacher statements.

**Classroom 2.** For classroom 2, the level of negative statements during baseline was much higher ($M = 16.50\%; SD = 4.72$) than the level of positive statements ($M = 2.33\%; SD = 2.16$). When the GBG was implemented, there was a slight decrease in the level of negative statements ($M = 10.00\%; SD = 3.67$); however, positive statements remained very low or nonexistent. During the next three phases, the level of positive statements (CBGG, $M = 1.38\%, SD = 2.13$; GBG, $M = 1.60\%, SD = 1.52$; CBGG, $M = 0.83\%, SD = 0.75$) and negative statements (CBGG, $M = 10.75\%; GBG, M = 9.00\%;$ CBGG, $M = 6.67\%$) remained consistent and variability was low. During the follow-up observation, when only posting of the classroom rules was observed, negative statements went up to 24\% of intervals and there were no positive statements used. Again, it is interesting to note the decrease in overall teacher managerial statements.
Figure 10. Classroom 2 teacher statements.

**Classroom 3.** In classroom 3, the level of negative teacher statements variable throughout baseline (M = 19.99%; SD = 9.55) and this pattern continued, although there were slight decreases, throughout intervention phases (M = 16.00%, 13.78%, 20.50%, 16.50%, in CBGG, GBG, CBGG, and GBG, respectively; SD = 6.07, 7.17, 6.60, and 7.89 in CBGG, GBG, CBGG, and GBG, respectively). With the exception of a few days, the teacher used more negative statements than positive statements. The level of positive statements was low throughout all phases.
During the follow-up observation, when there was no intervention in place, negative statements occurred during 34% of intervals while positive statements did not occur at all.

**Classroom 4.** Figure 12 shows the level of positive and negative statements for classroom 4. Negative statements during baseline averaged 11.83% ($SD = 5.23$) with a slightly increasing trend. When the CBGG was introduced, there was an immediate decrease ($M = 6.86\%, SD = 2.67$) and this level remained consistent during the rest of the phases (GBG, $M = 8.50\%, SD = 3.27$; CBGG, $M = 5.83\%, SD = 1.17$; GBG, $M = 6.83\%, SD = 3.66$) with low variability. During baseline, positive statements were variable. With interventions in place, the level of positive statements was low and, in some cases, nonexistent.
Overall, for each classroom, the level of positive statements was consistently lower than the level of negative statements. Intervention phases did not result in differences in statements. Unfortunately, there was generally a slight decrease in the use of positive statements. One possible explanation is that an improvement in students’ behavior led teachers to use fewer managerial statements in general.

The low PND levels also suggest the interventions did lead to an increase in positive statements. There were slight decreases in negative statements, but these PND levels were also low. Table 7 reveals the PND statistics for classrooms 1 and 2.

Table 7

<table>
<thead>
<tr>
<th>Classrooms 1 and 2 Teacher Statements PND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline &gt; GBG&gt;CBGG CBGG&gt;GBG GBG&gt;CBGG</td>
</tr>
</tbody>
</table>

*Figure 12. Classroom 4 teacher statements.*
Classroom 1 PND for positive statements had little or no change during intervention phases, indicating the interventions were “ineffective” in increasing teacher’s use of positive statements. PND for negative statements was in the 70-90% range from baseline to the GBG, indicating an “effective” change from baseline to intervention, with PND levels below 50% for the rest, indicating little change. Classroom 2 PND levels for positive and negative statements were below 50%, indicating the intervention was “ineffective” (Scruggs & Mastropieri, 1998).

Table 8 reveals the PND statistics for classrooms 3 and 4.

Table 8

*Classrooms 3 and 4 Teacher Statements PND*

<table>
<thead>
<tr>
<th></th>
<th>Baseline &gt;</th>
<th>GBGG &gt; GBG</th>
<th>GBG &gt; CBGG</th>
<th>CBGG &gt; GBG</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom 3</td>
<td>Positive Statements</td>
<td>Negative Statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>12.50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>33.33%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Classroom 4</th>
<th>Positive Statements</th>
<th>Negative Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00%</td>
<td>14.28%</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>16.67%</td>
<td>16.67%</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>16.67%</td>
</tr>
</tbody>
</table>

All PND levels for classrooms 3 and 4 were below 50%; thus, using the criteria set forth by Scruggs and Mastropieri (1999), the interventions were deemed “ineffective” in creating change in teacher statements.

**Acceptability**

Table 9 shows the results of the Teacher Social Validity Survey. Overall, the teachers found both interventions acceptable, with most statement ratings ranging from “agree” to “strongly agree.” The teacher of classrooms 1 and 2 rated both interventions the same for all statements and indicated she had no preference for one intervention over the other. The teacher of classroom 3 rated the interventions the same for all statements, except she “agreed” the CBGG script was easy to follow, but “strongly agreed” the GBG script was easy to follow. She indicated no preference for the interventions. The teacher of classroom 4 liked the procedures of the CBGG more than the GBG, found it easier to include in her routine, and would be more willing to use it in the future. This teacher
commented, “The Good Behavior Game seemed to interrupt my teaching a lot more than the Caught Being Good Game” and noted that she would use the CBGG in the future. All teachers stated they planned to continue to use the interventions in the future.

Table 9

*Teacher Social Validity Survey Results*

<table>
<thead>
<tr>
<th></th>
<th>Class 1 and 2</th>
<th>Class 3</th>
<th>Class 4</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBGG</td>
<td>5 5</td>
<td>4 5</td>
<td>5 5</td>
<td>4.7</td>
</tr>
<tr>
<td>GBG</td>
<td>5 5</td>
<td>4 4</td>
<td>4 3</td>
<td>3.7</td>
</tr>
<tr>
<td>The intervention was easy to follow</td>
<td>4 4</td>
<td>5 5</td>
<td>4 3</td>
<td>4.3</td>
</tr>
<tr>
<td>I liked the procedures used</td>
<td>4 4</td>
<td>5 5</td>
<td>4 2</td>
<td>3.7</td>
</tr>
<tr>
<td>in this intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The intervention was easy to include</td>
<td>4 4</td>
<td>5 5</td>
<td>4 4</td>
<td>4.3</td>
</tr>
<tr>
<td>in my daily routine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would be willing to use this</td>
<td>4 4</td>
<td>5 5</td>
<td>4 4</td>
<td>4.3</td>
</tr>
<tr>
<td>intervention in the future</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall, this intervention was</td>
<td>4 4</td>
<td>5 5</td>
<td>4 4</td>
<td>4.3</td>
</tr>
<tr>
<td>beneficial for the student(s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The game I would prefer to use in</td>
<td>No Preference</td>
<td>No Preference</td>
<td>Caught Being Good Game</td>
<td></td>
</tr>
<tr>
<td>the future is (circle one):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I plan to continue to use this</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>game in my classroom (Yes/No):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In addition to the survey, a follow-up observation was conducted at least two weeks after the study ended. During this observation the teachers were told they could implement either intervention and were given necessary supplies, but could also choose not to implement the interventions. Classrooms 1 and 2, which were taught by the same teacher, had the classroom rules posted as well, but did not implement any other steps. Classroom 3 continued to have the rules posted in the front of the classroom. During the typical observation time, which occurred at the end of the day, neither intervention was implemented. However, during the last transition when the students were gathering items to go home, which tended to be a difficult time for the students, the teacher implemented the GBG. The observer was not able to administer points at the same time as the teacher because she had already ended data collection; however, all other steps were implemented with 100% adherence. Classroom 4 continued to implement the CBGG with 100% adherence. All three teachers noted in the questionnaire they planned to continue using an intervention; however, only two of the three teachers continued to implement the interventions.

Table 10 shows the results for the Student Validity Survey. They were asked to rate statements about the interventions with “disagree”, “not sure”, and “agree.” The researcher assigned 1 to “disagree”, 2 to “not sure”, and 3 to “agree.”

Table 10

<table>
<thead>
<tr>
<th>Classroom</th>
</tr>
</thead>
</table>

48
<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>The classroom rules were clear and easy to understand.</td>
<td>3</td>
<td>3</td>
<td>2.7</td>
<td>2.7</td>
<td>2.9</td>
</tr>
<tr>
<td>I liked the game we used in class.</td>
<td>3</td>
<td>2.8</td>
<td>2.4</td>
<td>2.3</td>
<td>2.6</td>
</tr>
<tr>
<td>I did not mind being watched by the observers in the classroom.</td>
<td>3</td>
<td>2.9</td>
<td>2.6</td>
<td>2.3</td>
<td>2.7</td>
</tr>
<tr>
<td>I would recommend this game to other classroom teachers.</td>
<td>2.6</td>
<td>2.1</td>
<td>2.4</td>
<td>2.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Overall, this game was helpful for my class.</td>
<td>3</td>
<td>2.9</td>
<td>2.4</td>
<td>2.6</td>
<td>2.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Do you like one game better than the other?</th>
<th>Y – 42%</th>
<th>Y – 60%</th>
<th>Y – 100%</th>
<th>Y – 78%</th>
<th>Y – 70%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N – 58%</td>
<td>N – 40%</td>
<td>N – 0%</td>
<td>N – 22%</td>
<td>N – 30%</td>
</tr>
<tr>
<td>If yes, which game do you like better?</td>
<td>GBGG – 100%</td>
<td>CBGG – 89%</td>
<td>CBGG – 62.5%</td>
<td>CBGG – 71%</td>
<td>CBGG – 81%</td>
</tr>
<tr>
<td></td>
<td>GBG – 0%</td>
<td>GBG – 11%</td>
<td>GBG – 37.5%</td>
<td>GBG – 29%</td>
<td>GBG – 19%</td>
</tr>
</tbody>
</table>

Student results ranged from “not sure” to “agree” for each statement. Overall, students liked the interventions, thought they were helpful, and did not mind being watched by the observers. About 70% of the students had a preference for one intervention over the other, and of those, 81% preferred the positive version, the CBGG. Students were also asked if they would like to say anything else about the games and
many students had positive feedback including, “It’s fun when you get a ticket, candy, or sticker,” “It is one of my favorite games,” “I want to keep playing because it helps me,” and “I really like the timer.” One student suggested, “Maybe we can switch teams”. Teams remained the same throughout the study because there were no teams that consistently did not meet the goal compared to other teams; however, switching teams could have helped keep the game novel and exciting.

There were some students that stated they did not like the GBG. One student stated, “I really did not like the game where we got points for breaking rules”, and another stated, “When we got points for being good, I liked it. I dislike B” (which was the GBG). However, on the other hand, one student stated, “I like B (GBG) better because if you did something wrong you can fix it next time,” likely indicating the student benefited from receiving the immediate feedback when the teacher marked a point after each incident of rule breaking.

**Discussion**

One purpose of the current study was to determine if either intervention, the GBG or the CBGG, would be more effective in improving student behavior. Overall, both the CBGG and the GBG improved student behavior, by increasing student engagement and decreasing disruptions as well as decreasing each classroom’s prioritized target variable (i.e., off-task motor or off-task verbal). The results from the GBG and CBGG interventions did not show a lot of differentiation for student behavior.

Classrooms 1 and 2 implemented the GBG first. The most dramatic improvements were seen during this first implementation phase, but improvements were still seen during all intervention phases from baseline levels. Perhaps the newness of the
intervention during the first phase led to more dramatic results. Classrooms 3 and 4 implemented the CBGG first. For classroom 3, generally, the GBG led to more improved behavior than the CBGG. The first implementation of the CBGG led to improved behavior from baseline, then the GBG was implemented and led to further improved behavior. Similar results were seen during the reinstatement of the CBGG and GBG, where GBG levels were better than CBGG. Both interventions were effective for classroom 4, while the CBGG was slightly more effective.

These results reiterate the well-established effectiveness of a classroom management strategy, the GBG (Tingstrom et al., 2006), as well as a way to effectively modify it using a positive approach. The results from this study are comparable to the Wright and McCurdy (2011) study which found that both the interventions, the GBG and the CBGG, were effective at improving student behavior. The replication of the phases in the current study demonstrate the sustained effectiveness of both interventions over time. While favorable results are indicated, results must be interpreted with caution. Phase changes were determined using the prioritized classroom variable (i.e., off-task motor or off-task verbal) and the researchers aimed to have a stable trend before switching interventions; however, there was sometimes a decreasing or increasing trend in overall engagement and disruptions. Therefore, it is difficult to determine if one would be more effective yet.

An additional purpose of the study was to determine the effects of the interventions on teacher behavior, specifically teachers’ use of positive and negative statements. Overall, similar to the results on student behavior, the results for teacher behavior did not indicate noticeable changes between interventions. The results align
with previous research on teacher behavior which indicates the GBG does not lead to an increase in positive statements (Lannie & McCurdy, 2007). It was hypothesized the CBGG may lead to some increase in teachers’ use of positive statements due to scanning for positive behaviors rather than negative behaviors. Data indicate there was not an increase in positive statements and, for the most part, the level of negative statements remained higher than the positive statements. Overall, there seemed to be a slight decrease in all teacher statements, both positive and negative. Given the impact of student behavior on teacher behavior (Sherman & Cormier, 1974), it is likely the improved behavior led to the teacher using less managerial statements.

Additionally, the study sought to gain information regarding the acceptability of the interventions through teacher and student surveys as well as through follow-up observations. Teacher and student social validity results indicate both interventions were acceptable to teachers and students. Teachers of classrooms 1, 2, and 3 indicated they did not have a preference while the teacher of classroom 4 indicated she preferred the CBGG. Similar to previous research on the GBG (i.e., Barrish et al., 1969; Davies & White, 2000) and the recent research on the CBGG (Wright & McCurdy, 2011) teachers agreed the interventions were acceptable and easy to include in their routine. Adherence data show that the teachers also implemented the interventions with high degree of fidelity. Teachers had varying levels of education including one teacher who had a bachelors degree and another teacher who had a masters degree. This range of education level adds evidence that the interventions can be delivered with high fidelity, regardless of educational background.

It is interesting to note the differences between the teacher social validity survey
results and the follow-up adherence data. All teachers stated they agreed that the interventions were easy to implement and thought it led to improvements in student behavior; however, only classroom 4 continued to implement the intervention after the study was complete. Classroom 3 implemented the GBG for a portion of the observation period. The reasons for discontinuing the interventions despite the effectiveness and the teacher ratings are not known; however, sustainability may be more objective (Kennedy, 2005) and bear more weight than the survey results.

Student survey results indicated most students preferred the CBGG. During observations that occurred during the GBG intervention, the researcher noticed that some students who frequently rule broke rules were easily frustrated and commented they didn’t like the game. Although points were delivered on a private sheet, students were still able to see the teachers administer them. Also, through the student acceptability survey, one student suggested switching teams. Future researchers may periodically switch teams to keep the games novel and exciting, while keeping them heterogeneous.

The results suggest both the GBG and CBGG may be effective for improving student behavior. However, there were several limitations in the present study that should be noted. First, it is possible that some students were not differentiating between the two intervention conditions. Although the teacher introduced the game that would be played each day, the GBG or CBGG, the points were delivered privately. Some students, especially those in kindergarten, may not have discerned between the two. Therefore, while there were two different schedules of reinforcement (GBG, fixed ratio schedule of punishment; CBGG, variable interval momentary differential reinforcement of other behaviors), which tend to lead to different rates of responding, the private nature of the
point administration may have led to less noticeable differentiation. Future studies may utilize visual signals to specify the game being implemented or may administer points publicly.

In addition, one purpose of this study was to further examine the effects of the interventions on teacher behavior. During baseline, the levels of teacher statements, especially positive teacher statements, were often near zero. These low levels limit the amount of change possible. Additionally, although the study sought to examine the effects on teacher behavior, it was not explicitly targeted. Future research of the GBG and CBGG may choose to control teacher statements and study the impact on student behavior.

During the CBGG, teachers were prompted by a Motivaider to scan the room and administer points for teams that had all members following all the rules. When determining adherence for delivering points during the CBGG, both the teacher and observer administered the points simultaneously, and the number of agreements were divided by the number of disagreements. The observers noted that while it was typically easy to notice when the teacher began to deliver points, there was sometimes a slight delay. In the future, it may be beneficial to develop a signal for scanning the class or another system to ensure both teacher and observer are administering the points simultaneously.

Additionally, a variable interval schedule of reinforcement (VI = 5-6 minutes) was used during the CBGG. This schedule was determined in consultation with the classroom teacher. Classrooms with higher levels of disruptions during baseline, such as classroom 3, may have benefited from a thicker schedule of reinforcement.
The GBG and the CBGG were designed as classwide behavior management strategies and continued to be implemented as a tier 1 strategies in the current study. However, modifying components of the contingencies may successfully improve the behavior of students who are at-risk and continue to display challenging behaviors (Ling, Hawkins, & Weber, 2011). For example, pre-correction and/or positive attention directed toward the at-risk students may further improve students’ behavior. Future research may study the impact of such modifications. When implementing any intervention, it is important to keep in mind the function of students’ behavior. Although many students respond positively to praise, research shows some students do not respond well to positive teacher attention (March & Horner, 2002; Moore & Edwards, 2003). When deciding the most appropriate intervention, considering the function of behavior is important.

Given the effectiveness and acceptability of these strategies, teachers and school psychologists may find it advantageous to focus on classwide behavior rather than the behavior of an individual. When considering the classroom behavior management strategies, both the GBG and CBGG interventions may be viable options and it may be a matter of teacher preference, while keeping in mind that the procedures of the CBGG may align more closely with a positive school climate.
References


Individuals with Disabilities Education Improvement Act, H.R. 1350, 108th Congress (2004).


Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37*(6), 523-533.


group contingencies: Smoothing the rough waters. *Special Services in the Schools, 12*, 61-78.


### Observation Code

#### A negative statement (-1) is coded when a teacher makes a verbal
- Telling the class to be quiet when they are talking.
- Telling the class they are too loud.
- Telling a student to stop talking while the teacher is speaking.
- Working on student behavior and other vocal input.

#### A positive statement (+1) is coded when a teacher makes a verbal
- Acknowledging a student for working on an assignment.
- Acknowledging the following examples include:
  - Student loaned a student a good of work.
  - Student loaned a student a good of work.

<table>
<thead>
<tr>
<th>Partially Engaged (PTE)</th>
<th>Engaged (AET)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening to teacher and a question.</td>
<td>Listening to a piece of teacher about the assigned material.</td>
</tr>
<tr>
<td>Listing to teacher and a question.</td>
<td>Listening to a piece of teacher about the assigned material.</td>
</tr>
<tr>
<td>Reading assigned material silently.</td>
<td>Reading assigned material silently.</td>
</tr>
<tr>
<td>Reading assigned material silently.</td>
<td>Reading assigned material silently.</td>
</tr>
<tr>
<td>Listening to a lecture.</td>
<td>Listening to a lecture.</td>
</tr>
<tr>
<td>Writing.</td>
<td>Writing.</td>
</tr>
</tbody>
</table>

### Notes
- Examples include:
  - Acknowledging a student for working on an assignment.
  - Acknowledging the following examples include:
    - Student loaned a student a good of work.
    - Student loaned a student a good of work.
Appendix B

Script for introducing the contingency

Caught Being Good Game

Introduction on first day:
☐ Today we will begin a new game. We will see how well everyone can follow the rules. (Review rules and give examples and nonexamples.) During each (instructional period), we will play this game.

☐ I will be paying attention to how well all of you follow each of these rules. We have divided the class into teams. (Explain teams on the first day.) I have a timer that is set to go off at random times. Each time I feel it buzz, I will look around the classroom to see which teams have all its members following the rules. The teams that are following the rules will earn a point.

☐ At the end of the period, I will pick a slip from the ‘points’ jar. If the number of points your team earns is equal to or greater than the number on the slip, your team will earn a prize. I will pick a slip from the ‘prize’ jar, which will say the name of the prize you will earn. If your team does not reach the number of points on the slip, you can try again the next day.

☐ Also, each day I will post the number of points you earn on the board. At the end of the week, I will choose a slip, and if you receive more points than the number on the slip, your team will win a prize for the week.

☐ Are there any questions?

Faded Introduction:
☐ Today we are going to play the game again. Remember, during this game, I will look for teams who are following the rules to give points to when the buzzer goes off. You want to earn as many points as possible. At the end of the period, I will choose slips that say how many points your team needs and we will see which teams earn a prize.

Good Behavior Game

Introduction on first day:
☐ Today we will begin a new game. We will see well everyone can follow the rules. (Review rules and give examples and nonexamples.) During each (instructional period), we will play this game.

☐ We have divided the class into teams. (Explain teams on the first day.) Each time I see a student breaking a rule, I will assign a point to that team. For example, if a student from Team 1 calls out, I will assign team 1 a point. In this game, you do not want your team to earn points.

☐ At the end of the period, I will pick a slip from the ‘points’ jar. If the number of
points your team earns is equal to or less than the number on the slip, your team will earn a prize. I will pick a slip from the ‘prize’ jar, which will say the name of the prize you will earn. If your team does not earn a prize, you can try again the next day.

☐ Also, each day I will post the number of points you earn on the board. At the end of the week, I will choose a slip, and if you receive equal or fewer points than the number on the slip, your team will win a prize for the week.

☐ Are there any questions?

**Faded Introduction:**

☐ Today we are going to play a game again. During this game, I will be giving points to teams who break the rules. You want to earn as few points as possible to earn the reward. At the end of the period, I will choose slips and we will see which teams earn a prize.
Appendix C

Scripts

Script for Caught Being Good Game

1. Rules are posted in the front of the classroom.

2. Introduce the contingency (using introduction script or faded script).

3. At random times (prompted by timer) during the class period, give a point to each team that has all members following all rules.

4. Choose slip from the ‘points’ envelope/jar.

5. Tell them the number of points each team received and the number of points they needed to reach.

6. Choose slip from the ‘prize’ envelope/jar (if applicable).

7. Deliver the reward to the teams that met the goal (if applicable).

Script for Good Behavior Game

1. Rules are posted in the front of the classroom.

2. Introduce the contingency (using introduction script or faded script).

3. Each time a student breaks a rule during the class period, give a point to the team.

4. Choose slip from the ‘points’ envelope/jar.

5. Tell them the points each team received and the criterion for the maximum number of points.

6. Choose slip from the ‘prize’ envelope/jar (if applicable).

7. Deliver the reward to the teams that met the goal (if applicable).
Appendix D

Teacher Collected Data

Caught Being Good Game

Date: _______________________________ Classroom: ______

<table>
<thead>
<tr>
<th>Steps</th>
<th>Please check each step that is completed. If a step is not applicable, please mark N/A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rules are posted</td>
<td>☐</td>
</tr>
<tr>
<td>Introduce contingency (using introduction or faded script)</td>
<td>☐</td>
</tr>
<tr>
<td>Points are given random times during class period each time the timer vibrates</td>
<td>☐</td>
</tr>
<tr>
<td>At end of period, slip is chosen from ‘points’ envelope/jar</td>
<td>☐</td>
</tr>
<tr>
<td>Tell students how many points they received and how many they needed</td>
<td>☐</td>
</tr>
<tr>
<td>Reward is chosen from envelope/jar (if applicable)</td>
<td>☐</td>
</tr>
<tr>
<td>Reward is given to teams that meet the criterion (if applicable)</td>
<td>☐</td>
</tr>
</tbody>
</table>
| Please circle teams that earned the reward                           | Team 1
|                                                                        | Team 2
|                                                                        | Team 3
|                                                                        | Team 4

Good Behavior Game

When the timer vibrates, place a checkmark next to the teams that have all members following all classroom rules.

<table>
<thead>
<tr>
<th>Team 1</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team 2</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team 3</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team 4</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Date: _______________________________  Classroom: ______

<table>
<thead>
<tr>
<th>Steps</th>
<th>Please check each step that is completed. If a step is not applicable, please mark N/A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rules are posted</td>
<td>☐</td>
</tr>
<tr>
<td>Introduce contingency (using introduction or faded script)</td>
<td>☐</td>
</tr>
<tr>
<td>Points are given each time a student breaks a rule</td>
<td>☐</td>
</tr>
<tr>
<td>At end of period, slip is chosen from ‘points’ envelope/jar</td>
<td>☐</td>
</tr>
<tr>
<td>Tell students how many points they received and the maximum number allowed</td>
<td>☐</td>
</tr>
<tr>
<td>Reward is chosen from envelope (if applicable)</td>
<td>☐</td>
</tr>
<tr>
<td>Reward is given to teams that meet the criterion (if applicable)</td>
<td>☐</td>
</tr>
</tbody>
</table>
| Please circle teams that earned the reward                          | Team 1  
Team 2  
Team 3  
Team 4                                           |

Place a checkmark under the team each time a member breaks a classroom rule.

<table>
<thead>
<tr>
<th>Team 1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix E

Reward Assessment

Please rank these in order with 1 being the one you would like to earn the most and ___ being the one you would like to earn the least.

_____

_____

_____

_____

_____

_____

Appendix F
Treatment Integrity Checklists

Caught Being Good Game

☐ Rules are posted in the front of the classroom.
☐ Teacher introduces the contingency (using introduction script or faded script).
☐ At random times (prompted by timer) during the class period, teacher gives points for following rules. For each team that is following all rules, teacher marks a point.

<table>
<thead>
<tr>
<th></th>
<th>1st Check</th>
<th>2nd Check</th>
<th>3rd Check</th>
<th>4th Check</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Points</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>Observer Points</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

Total Agreements: ____
(i.e., both teacher and observer had box marked or unmarked)
Total Disagreements + Agreements: ____
(i.e., total number of boxes observed)
Agreement %: ____
(Divide agreements by total disagreements plus agreements. Check box when agreement is greater than 80%.)
☐ Chooses slip from the ‘point’ jar.
☐ The teacher tells the teams how many points they received, how many points were necessary, and which teams earned the reward.
☐ Chooses slip from the ‘reward’ jar (if applicable).
☐ Delivers the reward to the teams that met the criterion (if applicable).

Total Checked: ____
Total Available: ____
% Adherence: ____

Good Behavior Game

☐ Rules are posted in the front of the classroom.
☐ Teacher introduces the contingency (using introduction script or faded script).
☐ Each time a student engages in a rule violation, the teacher marks a point

<table>
<thead>
<tr>
<th></th>
<th>Team 1</th>
<th>Team 2</th>
<th>Team 3</th>
<th>Team 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observer Points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Teacher Points: ______
Total Observer Points: ______
Agreement %: ______
(Divide smaller total by larger total and multiply by 100. Check box when agreement is greater than 80%.)

☐ Chooses slip from the ‘point’ jar.
☐ The teacher tells how many points they received and the maximum number they could receive
☐ Chooses slip from the ‘reward’ jar (if applicable).
☐ Delivers the reward to the teams that were at or below the criterion (if applicable).

Total Checked: ______
Total Available: ______
% Adherence: ______
Appendix G

Teacher Social Validity Survey

Purpose: The purpose of this questionnaire is to get feedback concerning your overall satisfaction with the intervention(s) implemented in your classroom.
Directions: Please read the following statements and circle the number (1-5) that best describes your agreement or disagreement with each statement.

### Caught Being Good Game

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Not sure</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The intervention script was easy to follow</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I liked the procedures used in this intervention</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>The intervention was easy to include in my daily routine</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I would be willing to use this intervention in the future</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Overall, this intervention was beneficial for the student(s)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

### Good Behavior Game

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Not sure</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The intervention script was easy to follow</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I liked the procedures used in this intervention</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>The intervention was easy to include in my daily routine</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I would be willing to use this intervention in the future</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Overall, this intervention was beneficial for the student(s)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

The game I would prefer to use in the future is (circle one):
- Good Behavior Game
- Caught Being Good Game
- No Preference

I plan to continue implementing this game in my classroom (circle one):
- Yes
- No

Please make any additional comments: _______________________________________

(adapted from Erhardt et al., 1996 and Martens, Witt, Elliot, & Darveaux, 1985)
Appendix H

Student Social Validity Survey

Student Intervention Acceptability Questionnaire

Directions: Please read the following statements and circle the answer that best describes your agreement or disagreement with each statement.

1) The classroom rules were clear and easy to understand.
   Agree  Not Sure  Disagree

2) I liked the game we used in class.
   Agree  Not Sure  Disagree

3) The game helped me behave better in class.
   Agree  Not Sure  Disagree

4) I did not mind being watched by the observers in the classroom.
   Agree  Not Sure  Disagree

5) I would recommend this game to other classroom teachers.
   Agree  Not Sure  Disagree

6) Overall, this game was helpful for my class.
   Agree  Not Sure  Disagree

7) Do you like one game better than the other?  Yes  No
   If yes, which game do you like better?
   A. The game where our team gets points for following rules when the teacher’s timer vibrates.
   B. The game where our team gets points whenever we break the rules.

Is there anything else you would like to tell me about the games?

__________________________________________________________________________________________

(adapted from Erhardt et al., 1996 and Martens, Witt, Elliot, & Darveaux, 1985)