2012

Mystery Motivator Calendar: An Interdependent Group Contingency, Variable Ratio, Classroom Intervention

Eva Aleksandra Kowalewicz
Loyola University Chicago

Recommended Citation
https://ecommons.luc.edu/luc_diss/361

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.
Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 2012 Eva Aleksandra Kowalewicz
LOYOLA UNIVERSITY CHICAGO

MYSTERY MOTIVATOR CALENDAR: AN INTERDEPENDENT GROUP
CONTINGENCY, VARIABLE RATIO, CLASSROOM INTERVENTION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN SCHOOL PSYCHOLOGY

BY
EVA ALEKSANDRA KOWALEWICZ
CHICAGO, ILLINOIS
DECEMBER 2012
ACKNOWLEDGEMENTS

First of all, I would like to thank all the wonderful teachers and principals who welcomed me into their classrooms and their schools. I will be forever grateful for their openness, flexibility, enthusiasm for this project and support throughout the research process. I also greatly appreciate having had the opportunity to learn from the amazing work they do.

I would also like to acknowledge my committee – Dr. Gina Coffee (co-chair), Dr. Pamela Fenning (co-chair), and Dr. Rosario Pesce – for their encouragement and professional guidance at each step of the research process. I am very grateful to Dr. Gina Coffee for her mentorship and emotional support. I am especially thankful that she was willing to drive to meet me to discuss this project and its progress, braving Friday traffic in Chicago on numerous occasions. I would also like to thank Dr. Pamela Fenning for being the first to spark the idea for this project in her school based consultation class and for her extremely timely and valuable feedback. I am very thankful to Dr. Rosario Pesce for helping me find one of my most wonderful research sites. The staff members at this school were truly inspirational. I feel lucky to have had the opportunity to work with such an amazingly supportive principal and two inspiring kindergarten teachers. In addition, I deeply appreciate Dr. Michael Boyle’s and Rockeya Wilson’s help in classroom recruitment. Without them this project would have never come to fruition.
I would also like to acknowledge my family and friends. First, I would like to thank my mother, Agnes Kowalewicz for all that she did to make my life easier throughout the duration of this project. I would especially like to thank her for drawing over 60 poster Mystery Motivator Calendar templates, for spending her entire Memorial Day weekend proof reading the first draft of my dissertation, for doing my laundry, and for supporting me in far too many ways to list here. I would like to thank both my parents, Richard and Agnes Kowalewicz, for allowing me to take over their kitchen that served as my workspace, for keeping me well fed and for their emotional and financial support.

I would also like to thank my sweetheart, Matthew Puchalski, for his support. I am especially grateful that he added a touch of balance to my life in a time consumed by research and writing. Thanks are also in order for my favorite feline friend of 19 years for keeping me company through hours of writing. She helped relieve some of my stress by simply sitting by my side. Finally, I would like to thank my friends for being patient with my absence as I focused on completing this body of work.
This dissertation is dedicated to my father and mother, Ryszard and Agnieszka Kowalewicz, who instilled in me a love for learning and a tireless work ethic, taught me to be thoughtful in everything I do and who have supported me in every endeavor I have ever undertaken in my life. Without their support this project would not have been possible.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ iii

LIST OF TABLES .................................................................................................................... ix

LIST OF FIGURES ................................................................................................................ x

ABSTRACT ............................................................................................................................ xi

CHAPTER ONE: INTRODUCTION ......................................................................................... 1
  Background ....................................................................................................................... 3
  Purpose of this Study ....................................................................................................... 7
  Research Questions ......................................................................................................... 11
  Hypotheses ....................................................................................................................... 12
  Organization of the Paper ............................................................................................... 13

CHAPTER TWO: REVIEW OF THE LITERATURE .................................................................. 15
  Classroom Behavior Management .................................................................................. 15
  Students, Behavior and Motivation .............................................................................. 16
  Positive Behavior Support ............................................................................................. 18
  Behavior Management Through Contingencies ........................................................... 22
    Independent Group Contingencies ............................................................................. 23
    Dependent Group Contingencies ................................................................................ 24
    Interdependent Group Contingencies ......................................................................... 24
  Reinforcement Schedules ............................................................................................... 25
  Mystery Motivator .......................................................................................................... 27
    Independent Contingency Mystery Motivator Studies ............................................. 28
    Dependent Contingency Mystery Motivator Studies ................................................. 35
    Combined Dependent and Interdependent Mystery Motivator Study .................... 36
    Interdependent Mystery Motivator Studies .............................................................. 37
  Intervention Acceptability .............................................................................................. 42
  Summary .......................................................................................................................... 43

CHAPTER THREE: METHODOLOGY ..................................................................................... 49
  Procedures ......................................................................................................................... 49
    Recruitment .................................................................................................................. 49
    Inclusion Criteria for Participant Selection ............................................................... 49
    Consent ......................................................................................................................... 49
  Participants and Setting .................................................................................................. 50
    Classroom A .................................................................................................................. 53
    Classroom B .................................................................................................................. 54
    Classroom C .................................................................................................................. 54
    Classroom D .................................................................................................................. 55
    Classroom E .................................................................................................................. 56
CHAPTER FOUR: RESULTS .................................................................................................................. 81
Presentation of Results .................................................................................................................. 81
Research Question One .................................................................................................................. 82
  Classroom A ................................................................................................................................. 82
  Classroom B ................................................................................................................................. 86
  Classroom C ................................................................................................................................. 89
  Classroom D ................................................................................................................................. 92
  Classroom E ................................................................................................................................. 96
  Classroom F ................................................................................................................................. 100
  Classroom G ............................................................................................................................... 103
  Classroom H ............................................................................................................................... 107
Research Question Two ............................................................................................................... 111
Research Question Three ............................................................................................................ 114
Research Question Four .............................................................................................................. 117
Summary of Results ............................................................................................................. 120

CHAPTER FIVE: DISCUSSION ............................................................................................. 124
  Introduction ......................................................................................................................... 124
  Theoretical Orientation ....................................................................................................... 124
  Purpose ................................................................................................................................ 126
  Research Question One ...................................................................................................... 127
  Research Question Two ....................................................................................................... 134
  Research Question Three .................................................................................................... 137
  Research Question Four ...................................................................................................... 140
  Limitations ......................................................................................................................... 144
    Research Design Issues/Methodological Considerations .................................................. 145
    Issues Pertaining to History ............................................................................................. 147
    Instrumentation/Measurement Issues .............................................................................. 149
    Issues Specific to Individual Classrooms .......................................................................... 151
    Intervention Transportability ............................................................................................ 153
  Directions for Future Research .......................................................................................... 154
  Conclusions ......................................................................................................................... 157

APPENDIX A: RECRUITMENT EMAIL ................................................................................. 159

APPENDIX B: CONSENT DOCUMENTATION ....................................................................... 161

APPENDIX C: PROBLEM IDENTIFICATION INTERVIEW ..................................................... 171

APPENDIX D: INTERVENTION PROTOCOL ........................................................................ 173

APPENDIX E: MYSTERY MOTIVATOR CALENDAR IMAGES ............................................. 178

APPENDIX F: TALLY COUNTER IMAGE ............................................................................ 183

APPENDIX G: MYSTERY PRIZES PER CLASSROOM ......................................................... 185

APPENDIX H: CLASSROOM BEHAVIORAL GOALS HANDOUTS ..................................... 189

APPENDIX I: CHILDREN’S INTERVENTION RATING PROFILE ....................................... 198

APPENDIX J: INTERVENTION RATING PROFILE-20 ......................................................... 203

REFERENCES ......................................................................................................................... 206

VITA .......................................................................................................................................... 215
LIST OF TABLES

Table 1: Classroom Participant Demographic Data .......................................................... 51
Table 2: School-Wide Demographic Data from Illinois Interactive Report Card .......... 52
Table 3: Targeted Disruptive and Replacement Behavior per Classroom ..................... 64
Table 4: Inter-Observer Agreement per Classroom ........................................................ 78
Table 5: Classroom A Intervention Phase Data .............................................................. 85
Table 6: Classroom B Intervention Phase Data .............................................................. 88
Table 7: Classroom C Intervention Phase Data .............................................................. 91
Table 8: Classroom D Intervention Phase Data .............................................................. 95
Table 9: Classroom E Intervention Phase Data .............................................................. 99
Table 10: Classroom F Intervention Phase Data ............................................................ 102
Table 11: Classroom G Intervention Phase Data ............................................................ 106
Table 12: Classroom H Intervention Phase Data ............................................................ 111
Table 13: IRP-20 Question Mean Scores and Question Scores per Teacher ............... 116
Table 14: Overall Teacher Acceptability Score and Mean Item Score per Classroom .......................................................... 117
Table 15: Mean CIRP Item Scores and Mean Classroom Scores for Kindergarten and First Grade .................................................................................. 119
Table 16: Mean CIRP Item Scores and Mean Classroom Scores for Third and Fourth Grades ................................................................................................. 120
Table 17: Summary of Disruptive Classroom Behavior Data .......................................... 121
LIST OF FIGURES

Figure 1: Disruptive Behavior Classroom A ................................................................. 84
Figure 2: Disruptive Behavior Classroom B ................................................................. 88
Figure 3: Disruptive Behavior Classroom C ................................................................. 91
Figure 4: Disruptive Behavior Classroom D ................................................................. 94
Figure 5: Disruptive Behavior Classroom E ............................................................... 98
Figure 6: Disruptive Behavior Classroom F ............................................................... 102
Figure 7: Disruptive Behavior Classroom G ............................................................... 105
Figure 8: Disruptive Behavior Classroom H ............................................................... 110
ABSTRACT

Classroom behavior management problems are often seen as one of the primary barriers to achieving an educational environment conducive for academic and social-emotional growth of students. The literature indicates there is a need for evidence-based, easy-to-implement classroom behavioral interventions that align with the positive behavioral support philosophy. This study examined the effectiveness of the Mystery Motivator Calendar, an interdependent group contingency, variable-ratio, classwide intervention as a tool for reducing disruptive classroom behavior in eight diverse, general-education elementary school classrooms across seven different schools. The study employed a single-case, ABAB (baseline, intervention, withdrawal, reinstatement and follow-up), changing criterion design. The effectiveness of the intervention on disruptive classroom behavior was assessed for an eight-week period. The Mystery Motivator Calendar was found to decrease the frequency of disruptive behavior in all classrooms. The change in the effectiveness of the intervention with the passage of time was also examined. In the majority of classrooms, disruptive behavior continued to decrease, even while the criteria for intervention became more difficult to attain. Teacher intervention acceptability data indicated that seven of the eight teachers who participated in this study found the intervention to be acceptable. Results were somewhat mixed regarding student intervention satisfaction. Five classrooms generally indicated they “liked” the intervention, one classroom indicated the intervention was “somewhat liked”
and the remaining two classrooms generally rated the intervention more favorably than unfavorably, however ratings were not high enough to indicate satisfaction. Limitations and implications are presented in the discussion.
CHAPTER ONE
INTRODUCTION

In a national survey of teachers, 76% reported they would be better able to teach if student behavior problems were not so prevalent (Public Agenda, 2004). Another 33% of teachers reported having seriously considered quitting the profession of teaching due to student behavior problems (Public Agenda, 2004). The American Psychological Association also concluded that assistance with classroom management was one of the greatest needs identified by teachers (Coalition for Psychology in Schools and Education, 2006).

The aforementioned statistics underscore the importance of classroom behavioral management techniques. Effective classroom management is critical for an educational environment conducive for academic and social-emotional growth of students. Behavior problems in the classroom can lead to decreased academic learning time, decreased academic performance and even lower standardized test scores (Canter, Paige, Roth, Romero, & Carroll, 2004). Teachers with poor classroom management skills enable students’ continued behavior problems (Kellam, Ling, Merisca, Brown & Ialongo, 1998). Unfortunately, most teachers believe they have not been adequately trained to address disruptive behavior in the classroom (Kauffman, Wong, Lloyd, Hung, & Pullen, 1991). In particular, they often do not feel prepared to address the behavioral challenges of mainstreamed special education students (Chandler, 2000; Kehle & Bray, 2000;
Stephenson, Linfoot, & Martin, 1999), which can result in resisting inclusionary experiences for children with special needs (Scruggs & Mastropieri, 1996).

In summary, the aforementioned research indicates there is a great need for evidence-based, easy-to-implement classroom behavior interventions that are economic, flexible, and effective. Behavioral interventions applied in the classroom also need to be easily accessible and time efficient. Extensive protocols, complicated data gathering tools, and demanding and rigid curricula inherent in many behavioral intervention programs can overwhelm teachers. They may find such interventions to be too difficult and too time consuming to implement. As a result, they may not implement any proactive classroom behavior management interventions or may resort to punitive measures when presented with classroom discipline problems.

For teachers to feel comfortable addressing problem behaviors through any intervention, they must first have an understanding of why students behave the way they do. When students engage in inappropriate or disruptive behavior, it is possible they have not learned how to behave appropriately, they have not seen the benefits of appropriate behavior, or they have learned to get their needs met through engaging in inappropriate behavior (O’Neill et al., 1997; Sprick, 2009). The underlying forces of disruptive behavior must be understood and taken into consideration when designing and implementing school based behavioral interventions.

In order for teachers to implement interventions they must view them to be socially valid (Walker, 2004). Social validity refers to judgments about the social importance of interventions on three levels: (1) treatment must be socially significant, so
the behaviors targeted are deemed important and relevant; (2) treatment procedures must be considered socially appropriate; and (3) the effects of the treatment must be socially important or have meaningful significance (Wolfe, 1978). It is important to examine social validity in research studies because socially valid interventions have greater transportability from research settings to applied settings. Essentially, in order for teachers to implement any intervention, they must find it acceptable for use in their classrooms and deem it to be appropriate, effective, and fair.

**Background**

In the recent years, the literature on Positive Behavior Support (PBS), a systemic approach to behavior management in schools, has helped shift the focus of behavior management from a reactive and negative approach to a more proactive and positive approach (Office of Special Education Programs Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010). The research behind this philosophical shift has consistently demonstrated PBS is more effective in achieving long-term behavioral change and in teaching appropriate behavioral skills than the traditional reactive and punishment oriented model (Office of Special Education Programs Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010). PBS includes primary (Tier 1), secondary (Tier 2), and tertiary (Tier 3) prevention strategies. Primary prevention includes universal interventions focused on school-wide and class-wide systems. Secondary prevention includes specialized group interventions and focuses on at-risk students. Tertiary prevention includes specialized individual
interventions and focuses on students with chronic, intense problem behavior (OSEP Technical Assistance Center for Positive Behavioral Interventions & Supports, 2005). One proactive, effective and systemic way of addressing behavior management in classrooms, which aligns with PBS practices, involves the use of contingencies or contingency contracting. A contingency contract is a written description of dependent relationships involving student performance, teacher performance and reinforcing consequences (Schloss & Smith, 1998). Both contingency contracts and PBIS include clear behavioral expectations/rules, procedures for reinforcing expected behavior, procedures for discouraging problem behavior, procedures for data collection and decision-making, and enforcement of rules and expectations.

Teachers can develop contingency contracts for either individual students or for groups of students. Many teachers prefer to use group contingencies for managing behavior because individual contingencies are more time intensive in implementation (Litow & Pumroy, 1975; Maag, 1999). In a group-oriented contingency, the entire class is reinforced based on the behavior of one student, a group of students, or the entire class (Maag, 2001a). There are three primary types of group contingencies: independent group contingencies, dependent group contingencies, and interdependent group contingencies. Independent group contingencies employ the same behaviors, criteria, and consequences for all the students in a classroom, but reinforcement is delivered individually, based on each student’s behavior (Litow & Pumroy, 1975; Rathvon, 2008). In a dependent group contingency, one student or a small group of students may earn the reward for the entire class (Litow & Pumroy, 1975; Skinner, Skinner, & Sterling Turner, 2002). Finally,
interdependent group contingencies reinforce the group based upon the entire class meeting a specified criterion (Litow & Pumroy, 1975; Rathvon, 2008). The literature reports many advantages associated with the use of interdependent group contingencies.

Individual and group contingency plans can vary according to the type of reinforcement schedule used. Schedules of reinforcement refer to patterns of timing for the delivery of reinforcement (Alberto & Troutman, 2009). There are four primary types of reinforcement schedules: fixed-interval, variable-interval, fixed-ratio and variable-ratio. All reinforcement schedules will be described in greater detail in Chapter Two. In a variable ratio schedule reinforcement is delivered after the completion of a variable number of responses (Cooper, Heron & Heward, 2007). Target behaviors under this schedule are high in frequency and consistent (Alberto & Troutman, 2009). A variable schedule of reinforcement heightens the child’s anticipation for the reward, and decreases the likelihood of a post-reinforcement pause (Cooper et al., 2007; Davis & Blankenship, 1996).

Mystery Motivators are a type of positive reinforcement technique described by Jenson, Rhode and Reavis (1992), which use a variable ratio reinforcement schedule. The Mystery Motivator involves the use of unknown reinforcer(s) with an unpredictable reinforcement schedule. Mystery Motivator studies have utilized independent, dependent, and interdependent contingencies in an attempt to improve student’s social-emotional, behavioral and academic functioning. The Mystery Motivator can be used with individual students, groups of students or entire classrooms to increase and/or decrease specific behaviors (Wright, 2004). The majority of studies conducted on the efficacy of the
Mystery Motivator employ independent group contingencies for both behavioral and academic changes (DeMartini-Scully, Bray, & Kehle, 2000; Kehle, Madaus, Baratta, & Bray, 1998; LeBlanc, 1998; Madaus, Kehle, Madaus, & Bray, 2003; Matovic, 2010; Moore, Waguespack, Wickstrom, Witt & Gaydos, 1994; Motram, Bray, Kehle, & Jenson, 2000; Musser, Bray, Kehle, & Jenson, 2001; Robinson & Sheridan, 2000). Because reinforcement is delivered individually, based on each student’s behavior such interventions can be complicated and time intensive to implement. Only one study has utilized both a dependent group contingency plan in combination with a Mystery Motivator (Kehle & Bray, 2000). Also, only one study has assessed the effectiveness of a combined dependent and interdependent group contingency with a Mystery Motivator (Kelshaw, Sterling-Turner, Henry, & Skinner, 2000).

Similarly, few studies have assessed the effects of the interdependent group contingency combined with a Mystery Motivator on student behavior. Murphy, Theodore, Aloiso, Alric-Edwards and Hughes (2007) assessed the effectiveness of the interdependent group contingency combined with a Mystery Motivator in reducing disruptive behavior in a preschool classroom setting. In 2010, Hoag conducted another preschool study utilizing an interdependent group contingency with a Mystery Motivator. She compared the effectiveness of the Mystery Motivator to a known reinforcer. Schanding and Sterling-Turner (2010) conducted a study utilizing an interdependent group contingency plan in combination with a Mystery Motivator in a general education high school classroom. In this study, researchers assessed the mystery motivator, as it affected three students identified as exhibiting problem behaviors, as well as the effects
on non-identified students in a ninth-grade high school biology class. Effects on non-identified students were assessed through a sampling methodology. In Kraemer, Davies, Arndt, and Sawyer (2012) researchers compared the efficacy of the Mystery Motivator intervention with the Get ‘Em on Task intervention, a computer-signaling program that helps teachers reward individual students based on individualized auditory signals, in addressing off-task classroom behavior in two fifth grade classrooms.

Almost all of the Mystery Motivator interventions proved to be effective, regardless of whether the contingency plan was independent, dependent or interdependent. Teachers (Madaus et al., 2003; Moore et al., 1994; Murphy, Theodore, Aloiso, Alric-Edwards & Hughes, 2007; Musser et al., 2001) and students (LeBlanc, 1998; Madaus et al., 2003; Musser et al., 2001; Robinson & Sheridan, 2000) also deemed most of the interventions as acceptable or satisfactory. Teachers’ acceptance is particularly important since they are more likely to implement and utilize empirically supported interventions if they find them socially valid. This is why many of the Mystery Motivator studies also assessed the social validity of the intervention from a teacher’s perspective (and sometimes the students’ as well) (Madaus et al., 2003; Moore et al., 1994; Murphy et al., 2007; Musser et al., 2001).

**Purpose of this Study**

The bulk of the existing research on the effectiveness of Mystery Motivators, evaluates the effect of the intervention either on targeted individual students or targeted groups of students (DeMartini et al., 2000; Kehle & Bray, 2000; Kehle et al., 1998; LeBlanc, 1998; Madaus et al., 1994; Matovic, 2010; Moore et al., 1994; Motram et al.,
2000; Musser et al., 2001; Robinson & Sheridan, 2000; Theodore et al., 2001), and often these students have already been identified for special education services (Kehle & Bray, 2000; Motram et al., 2000; Musser et al., 2001; Theodore et al., 2001) or as an at-risk group (Murphy et al., 2007). Only a handful of studies have evaluated the effectiveness of the Mystery Motivator in an interdependent context on the performance of all the students in a classroom (Bennett, 2010; Hoag, 2010; Kraemer et al., 2012; Murphy et al., 2007; Schanding & Sterling-Turner, 2010). Of these studies, only four examined the effects of the Mystery Motivator intervention on disruptive classroom behavior (Kraemer, 2012; Murphy et al., 2007; Schanding & Sterling-Turner, 2010).

Of these studies, only one was conducted in a general education elementary school classroom with the aim of addressing disruptive classroom behavior (Kraemer et al., 2012). This study compared the effects of the Mystery Motivator and the Get ‘Em on Task interventions on disruptive classroom behavior using an alternating treatment design. Significant limitations were identified. First, the study was only conducted in two classrooms in one school with a very homogenous student population. Secondly, the study did not control for carry over effects in the alternating treatment design. Third, the study did not take into account counterbalancing or presenting the interventions in different orders in the classrooms that participated. Interventions were presented in the same sequence in both classrooms. Thus, the results of the study can only be generalized to those exposed to the same treatments in the same order of presentation (Kazdin, 2011). Furthermore, while the Mystery Motivator intervention was in place every day, objective, quantitative data were collected only two times per a week using a time sampling
methodology by an outside observer. Finally, despite the author’s use of the term ‘Mystery Motivator’, the reinforcers did not remain unknown, instead, prizes were chosen from a reinforcement menu.

Another study that examined the effects of the Mystery Motivator on classroom behavior, conducted in a general education biology high school classroom (Schanding & Sterling-Turner, 2010), also had important limitations. In this study, a sampling methodology was used to evaluate the effectiveness of the intervention on non-target students. Sampling data was collected by an outside observer and extrapolated to paint a picture of disruptive behavior for the whole class. Such extrapolations can only provide an estimate of disruptive classroom behavior (Snowden, 2004).

In the two preschool studies described earlier (Hoag, 2010; Murphy et al., 2007), the teachers were required to keep track of the behavior of individual students within the classroom by marking checks next to individual student’s names when they did not meet behavioral goals. Teachers in both studies identified the data recording procedures as cumbersome.

Another limitation of many of the prior studies is the short duration of the intervention. The Mystery Motivator was only implemented for a short time period, usually around two to three weeks (DeMartini-Scully et al., 2000; Kehle et al., 2000; Matovic, 2010; Madaus et al. 2003; Musser et al., 2001). As a result, the studies did not ascertain whether the efficacy of the intervention decreased over time as the novelty wore off for students.

This study sought to address the limitations identified in existing studies. The
effectiveness of an interdependent group contingency Mystery Motivator intervention as a classroom behavior management tool for general education elementary school classrooms was assessed. First, this study was conducted across seven different schools in eight classrooms, which led to a diverse student sample. Second, a frequency count as opposed to a time sampling methodology was utilized to gather data, which was a more accurate measure of disruptive behavior (Snowden, 2004). Third, care was taken to ensure that potential reinforcers remained a mystery for students thus ensuring anticipation and interest remained heightened (Skinner et al., 1996).

To address the limitation of cumbersome data recording procedures (Hoag, 2010; Matovic, 2010; Murphy et al., 2007) identified in many studies, classroom teachers were provided with a tally counter to record the frequency of disruptive behavior in the classroom. Teachers did not have to stop instruction to record data. They simply added a tally to the counter, which they held in their hand by pressing on a small button. Furthermore, the tally counter provided a discriminative stimulus to notify students of disruptive behavior by making a clicking sound. This stimulus was much less intrusive to the flow of classroom activities than verbal redirection.

In the present study, the intervention was implemented for a period of eight weeks, which allowed for a better understanding of how the passage of time impacts the effectiveness of the intervention. In addition to addressing limitations, this study was expanded to include individual student handouts to serve as reminders of intervention goals. Also, a monthly Mystery Motivator Calendar was posted for the whole class to see and monitor progress.
Also, because teachers’ perceptions about the utility, ease and effectiveness of interventions are crucial for social validity, this study also explored teacher acceptability of the Mystery Motivator intervention. The opinions of the recipients of the intervention, the students, were also taken into consideration when evaluating social validity.

Continued research in the area of group contingency interdependent Mystery Motivator interventions targeting disruptive classroom behavior is still needed because the four existing studies utilized single case design. Often these studies were only conducted in one or two classrooms (Kraemer et al., 2012; Murphy et al., 2007; Schanding & Sterling Turner, 2010) resulting in limitations to external validity (Kazdin, 2011). Replications in multiple contexts, with populations composed of varying ages, grades, ethnic and socio-economic backgrounds are necessary to provide a stronger basis for observed relationship(s). The specific research questions in the present study were as follows:

**Research Questions**

1) How will the interdependent group contingency Mystery Motivator Calendar intervention affect the frequency of disruptive behavioral events in general education elementary school classrooms, as identified by each teacher?

2) How will the effectiveness of the Mystery Motivator Calendar intervention change over time?

3) How will teachers rate the acceptability of the Mystery Motivator Calendar intervention as measured by a modified version of the Intervention Rating Profile-20 (Witt & Martens, 1983)?
4) On average, how will classroom students rate their satisfaction with the Mystery Motivator Calendar intervention as measured by two modified versions the Children’s Intervention Rating Profile (Witt & Elliott, 1985)?

**Hypotheses**

1) It was hypothesized the proposed study would replicate the results of prior studies, which documented the effectiveness of the Mystery Motivator in eliciting behavioral change in classrooms as an interdependent group contingency intervention. Thus, it was hypothesized that the Mystery Motivator Calendar would lead to a decrease in disruptive classroom behavior.

2) It was hypothesized that a slight decrease in efficacy would be observed as students became satiated with the intervention and the novelty decreased.

3) Given the results of previous studies that examined the acceptability of the Mystery Motivator intervention, it was hypothesized teachers would rate the intervention acceptable as measured by a modified version of the Intervention Rating Profile (IRP-20) (Witt & Martens, 1983). The intervention was considered acceptable if a score of 100 or higher was obtained from teachers on the IRP-20.

4) Given the results of previous studies that examined students’ opinions about the Mystery Motivator intervention, it was hypothesized that average classroom ratings on two modified versions of the Children’s Intervention Rating Profile (Witt & Elliot, 1985), would indicate student intervention satisfaction.
Organization of the Paper

This paper is organized into five chapters. The first chapter is the introduction, which has been presented above. Chapter Two examines the existing literature relevant to the topic. First, classroom behavior management and its importance are discussed. Next, some of the potential reasons students engage in disruptive behavior and misbehavior are reviewed followed by an overview of the Positive Behavior Support approach to managing behavior in schools. This is followed by a discussion of behavior management through the use of various types of group contingencies. After the discussion of various group contingencies, reinforcement schedules are reviewed. There is a particular focus on the variable-ratio reinforcement schedule because this schedule is utilized in all Mystery Motivator interventions. Subsequently, a thorough review of the existing literature on Mystery Motivator interventions, most notably interdependent group contingency Mystery Motivator interventions is presented. The importance of teacher and student intervention acceptability is discussed.

In Chapter Three, Methodology, procedures, research design and data analysis are discussed in detail. This includes a detailed description of the participants, intervention sites, consent procedures, and intervention procedures. This is followed by a description of the research design, which includes dependent and independent variables, materials, instrumentation and treatment integrity. Finally, the procedures for data analysis are described.

In Chapter Four, the results of the study are presented by research question. First, the effects of the Mystery Motivator Calendar on disruptive classroom behavior are
examined. Next, the paper examines whether the effects of the Mystery Motivator Calendar changed over time. Then, teacher intervention acceptability is discussed through an analysis of responses to a modified version of the Intervention Rating Profile-20 (Witt & Marten, 1983). Finally, student intervention satisfaction is examined through an analysis of average classroom responses to two modified versions of the Children’s Intervention Rating Profile (Witt & Elliott, 1985).

The final chapter is devoted to a discussion of the study results in the context of the existing literature within a behavioral theoretical orientation. First, a brief summary of the purpose of the study is outlined. Next, the results are discussed in the context of relevant literature along with potential implications. Then, the study’s limitations are discussed followed by suggestions for future research. Finally, overall conclusions are drawn.
CHAPTER TWO

REVIEW OF THE LITERATURE

Classroom Behavior Management

Effective classroom management is critical for an educational environment conducive to academic and social-emotional growth of students. Behavior problems in the classroom require teachers to spend a substantial amount of time managing disruptive behaviors, thus taking time away from academic instruction (Kauffman et al., 1999). Behavior problems can lead to decreased academic learning time, decreased academic performance and lower standardized test scores (Canter et al., 2004). In fact, 76% of teachers stated they would be better able to teach if discipline problems were not so prevalent (Public Agenda, 2004). Additionally, a 2006 survey of teachers, conducted by the American Psychological Association, found assistance with classroom management as one of the greatest needs identified by teachers (Coalition for Psychology in Schools and Education, 2006). Over 2,000 teacher respondents stated they wanted assistance with classroom management due to worries about student safety and a desire for strategies to deal with disruptive behaviors.

The ability of teachers to manage the behavior of their students is essential in achieving positive educational outcomes. Teachers with poor classroom management skills enable students’ continued behavior problems (Kellam et al., 1998). Unfortunately, most teachers do not believe they have been adequately trained to address disruptive
behavior in the classroom (Kauffman et al., 1991). General education teachers often do not feel prepared to address the behavioral challenges mainstreamed special education students present (Kehle & Bray, 2000; University of Kansas, 2005). This results in resistance to providing inclusionary experiences for children with special education needs (Kehle & Bray, 2000). Disruptive behavior has also been found to be especially problematic in classrooms of economically disadvantaged students (Kellam et al., 1998). As a result, teachers’ inability to manage behavior has the largest impact on students with the most significant needs.

Teachers who feel that they are ineffective report high levels of stress and symptoms of burnout (Browers & Tomic, 2000). In fact, a sense of inefficacy in managing disruptive classroom behavior has been found to be one of the most significant reasons teachers provide for leaving the profession (Ingersoll & Smith, 2003). In one study, 33% of teachers reported having seriously considered quitting teaching due to student discipline and behavior problems (Public Agenda, 2004). In order to retain teachers and increase teachers’ confidence in their ability to successfully deal with a wide range of behavior related classroom challenges, easy to implement, evidenced-based classroom behavioral interventions need to be available.

**Students, Behavior and Motivation**

For teachers to feel comfortable addressing problem behaviors through interventions, they must first have an understanding of why students behave the way they do. According to the behavioral theoretical orientation, the majority of human behavior is learned. This means that it can be unlearned and shaped (Sprick, 2009). Behavior is
repeated when it is reinforced. Furthermore, all behavior that is repeated is motivated; individuals do not continue to engage in behavior without motivation (Katzell & Thompson, 1990). Motivation can be better understood through the following model, “Expectancy x Value = Motivation” (Feather, 1982). In this model “expectancy” refers to the degree to which an individual expects to be successful and “value” refers to the degree to which an individual values the rewards that accompany success (Feather, 1982). This model recognizes that an individual’s motivation is the product of both “expectancy” and “value.” A decrease in either one of these factors will result in a decrease in overall motivation (Feather, 1982). Hall and Hall (1980) remind teachers that when using techniques and reinforcers to increase students’ motivation, variety produces better effects than relying on only one tool. Repetition can lead to boredom and satiation thus lessening the motivating effectiveness of reinforcing consequences, both positive and negative.

Motivation can be categorized as extrinsic or intrinsic. The “value” factor in the aforementioned model can include both intrinsic and extrinsic motivation (Sprick, 2009). Intrinsic motivation occurs when the pleasant consequences of a behavior are natural or directly related to the nature of the behavior. For most children play is intrinsically motivating as it is pleasurable by nature. Extrinsic motivation occurs when someone engages in a behavior because of pleasant consequences not directly related to the essential nature of the behavior (Sprick, p. 27). For most children, completing homework, which is not pleasurable by nature, is usually extrinsically motivated to avoid the negative consequences of poor grades. Most people are motivated to engage in a
particular behavior by a complex mix of intrinsic and extrinsic factors (p. 26). While some teachers may believe the only kind of valid motivation is intrinsic motivation, and children should not receive rewards of any kind (Cameron, Banko, Pierce, 2001), the principles of PBS with respect to acknowledgment/reinforcement do not adhere to this notion.

Regardless of the type of motivation, motivation must be present in order for behavior to occur. When students continue to engage in inappropriate or disruptive behavior, it is possible they have not learned how to behave appropriately, they have not seen the benefits of appropriate behavior, nor have they learned to get their needs met through engaging in appropriate behavior (O’Neill et al., 1997). Students who do not know how to engage in appropriate behavior can be said to have skill deficits, while students who know how to engage in appropriate behavior, but do not have the motivation to do so, can be said to have performance deficits. Such factors must be taken into consideration when designing and implementing school based behavioral interventions. Schools should strive to ensure students know how to behave appropriately, they are motivated to behave appropriately, and the school environment does not support or reinforce inappropriate behavior.

**Positive Behavior Support**

An approach to behavior management in schools that takes all the aforementioned factors into account is known as Positive Behavior Support (PBS). Over the past 20 years, the educational system has changed greatly. The norms for disciplining problem behavior through punishment have been gradually replaced with more proactive and
positive approaches (Sprick, 2009). PBS has been the driving force of this shift (Office of Special Education Programs Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010). PBS seeks to impact the larger ecological context in which the student functions (Alberto & Troutman, 2009). It is similar to prevention focused consultation models (Meyers, Meyers, & Grogg, 2004) and is rooted in applied behavior analysis (Sugai & Horner, 1999) and humanistic psychology (Carr et al., 2002). Many educators and researchers believe PBS is the best model for delivering behavioral support to all students (Sprague, 2006). PBS strives to make “problem behavior less effective, efficient, and relevant, and desired behavior more functional” (OSEP Technical Assistance Center for Positive Behavioral Interventions & Supports, 2005). Key components of the PBS include: clear system-wide behavioral expectations/rules, procedures for teaching expected behavior (including ‘telling’ students what is expected, ‘showing’ students what is expected, and allowing students to have the opportunity to ‘practice’ what is expected), procedures for encouraging expected behavior, procedures for discouraging problem behavior, procedures for data collection and decision making, consistent and positive enforcement of rules and expectations, proactive correction of rule violations and social behavior errors, collaboration among school professionals (Lewis & Sugai, 1999; Sprick, 2009) and the use of a problem-solving team process (Office of Special Education Programs Technical Assistance Center for Positive Behavioral Interventions & Supports, 2005).

PBS includes primary (Tier 1), secondary (Tier 2), and tertiary (Tier 3) supports. Primary support includes universal interventions focused on school-wide and class-wide
systems. Secondary support includes specialized group interventions focused on at-risk students who need more support than is offered at Tier 1. Tertiary support includes specialized individual interventions and focuses on students with chronic, intense problem behavior (Office of Special Education Programs Technical Assistance Center for Positive Behavioral Interventions & Supports, 2005).

The research behind this philosophical shift has consistently demonstrated that PBS is more effective in achieving long-term behavioral change and teaching appropriate behavioral skills than the traditional reactive and punishment oriented model (Office of Special Education Programs Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010). Schools that have implemented PBS have reported a 20% to 60% reduction in discipline referrals coupled with academic gains and a more pro-social school climate (Cushing, 2000). In fact, a review of over 100 articles that investigated behavior interventions and behavioral outcomes concluded that PBS was successful in addressing up to 80% of problem behaviors for two-thirds of the behaviors studied (Carr et al., 1999). In addition, the Individuals with Disabilities Education Act (IDEA) mandate the consideration of positive behavioral interventions and supports when a student’s behavior impedes his/her learning or that of others (Turnbull & Turnbull, 2000).

Despite the strong research base supporting the use of the PBS model in schools, there many teachers who resist the implementation of PBS practices. They do not think students should be rewarded for doing what is expected (Horn, 1991). Techniques based on positive reinforcement are also perceived as a threat to individual freedom, as
externally applied reinforcement is viewed as coercive in nature (Maag, 2001b). Society expects individuals to be intrinsically motivated, however, the educators who criticize positive reinforcement use punishment as a means of behavior management, although it is also externally applied (Maag, 2001b) and limits individual freedom. Punishment is much more widely used for several reasons. It is quickly and easily administered, it stops the problem behavior quickly, albeit temporarily, and it can be reinforcing to the teacher (Maag, 2001b). However, with punishment often come serious negative side effects (Newsome, Favell, & Rincover, 1983), which are undesirable in the school setting. First, punishment does not help students learn and develop positive behaviors; rather it has been shown to cause avoidance behavior (Newsome, Favell, & Rincover, 1983). Second, the recipient of punishment may start to avoid the person administering punishment, as well as the situation in which punishment occurred (Newsome, Favell, & Rincover, 1983). Furthermore, the punished individual may find the best way to avoid punishment is to avoid doing anything that resembles the punished action, which may lead to learned helplessness (Newsome, Favell, & Rincover, 1983). Additionally, the punished individual may only stop engaging in undesirable behavior until the aversive situation is avoided or until pleasant results outweigh the potential punishment (Newsome, Favell, & Rincover, 1983). Also, a child may model punishment if he/she perceives adults solve problems through the use of punishment (Newsome, Favell, & Rincover, 1983). In addition, punishment may lead to retaliatory behavior because a punished individual is likely to be upset (Newsome, Favell, & Rincover, 1983). Finally, punishment can lead to negative self evaluations (Newsome, Favell, & Rincover, 1983). The aforementioned negative side
effects of punishment present many good reasons for teachers to use different strategies, such as positive behavior support, to control unwanted behaviors (LaVigna & Donnellan, 1986). In order for teachers to begin to employ positive behavior interventions on a greater scale, they need to be assured these interventions are effective, time efficient, and easy to implement (University of Kansas, 2005), thus socially valid (Wolfe, 2004). This is particularly important in the current climate of budgetary constraints, growing class sizes, and increased pressure for accountability based on student outcomes.

**Behavior Management Through Contingencies**

An effective way of addressing behavior management in classrooms, which aligns with PBS practices, is through contingencies or contingency contracting. A contingency contract is a description of contingency relationships incorporating student performance, teacher performance and the reinforcing consequences (Schloss & Smith, 1998). Contingency contracts should encompass the following: a precise definition of expected behaviors, delivery of positive consequences following the established criteria for expected behavior, a statement of the adverse consequences for not meeting the established behavior criteria, a statement of adult responsibilities, and an objective behavior recording system (Schloss & Smith, 1998). Teachers can develop contingency contracts for either individual students or for groups of students. In an individual contingency, the intervention and its components are individualized. Individual contingencies often require a great deal of time and resources for correct implementation and may be unrealistic and impractical for busy teachers (Litow & Pumroy, 1975).
Many teachers prefer to use group contingencies for managing classroom behavior. In a group-oriented contingency, the entire class is reinforced based on the behavior of one student, a group of students, or the entire class (Maag, 2001a). Group contingencies can save time and resources in program development and implementation (Maag, 1999). Another advantage of group contingencies is their potential to facilitate positive social interactions (Maag, 1999). A meta-analysis investigating the efficacy of interventions designed to reduce disruptive classroom behaviors in school-aged children found group contingencies to be the most effective treatment (Stage & Quiroz, 1997).

There are three types of classroom-based group contingencies: independent, dependent, and interdependent (Litow & Pomroy, 1975).

**Independent Group Contingencies**

Independent group contingencies employ the same behaviors, criteria, and consequences for all the students in a classroom, but reinforcement is delivered individually, based on each student’s behavior (Litow & Pumroy, 1975). Each student’s access to reinforcers is individually determined, and each student is responsible only for his or her own behavior, independent of the class’s performance (Litow & Pumroy, 1975; Maag, 1999; Rathvon, 2008). The only reason this intervention is classified as a ‘group’ contingency is because every student participates in it and every student has the same access to the reinforcers. An example of an independent group contingency is a token economy in which students earn individual points based on their performance. Some advantages of this type of contingency are that no student is penalized for the behavior of anyone else and each student has access to rewards under exactly the same terms,
similarly, a disadvantage is the power of peer pressure is unlikely to be harnessed (Maag, 1999). Another disadvantage of this type of contingency from the teacher perspective is the cumbersome and time-consuming data recording procedures for participating students are required (Skinner, Skinner, & Sterling-Turner, 2002)

**Dependent Group Contingencies**

In a dependent group oriented contingency, one student or a small group of students may earn the reward for the entire class (Litow & Pumroy, 1975; Skinner et al., 2002). For example, if one student has difficulty with behavior, then the teacher could develop a behavior contract with that student, if the student is able to meet the requirements of the behavior contract then the entire class would earn a pizza party. An advantage of a dependent group contingency is the target student becomes the “hero” and his/her peers may root him/her on. However, a related disadvantage is the target student may receive negative attention if he/she fails to earn the reward (Maag, 1999). A study by Gresham and Greshman (1982) found dependent group contingencies to be more effective than independent group contingencies in decreasing disruptive classroom behavior.

**Interdependent Group Contingencies**

Finally, interdependent group contingencies reinforce the group based upon the entire class meeting a specified criterion (Litow & Pumroy, 1975; Rathvon, 2008). In an interdependent group contingency, all the students in a defined group must meet the set standard for any of the group members to earn reinforcement. An example of the interdependent contingency is using a jar of marbles and an intermittent beep tape to keep
track of appropriate behavior during class silent reading time (University of Kansas, 2005). If all group members are exhibiting appropriate behavior when the beep sounds then a marble is added to the jar. When the jar is full, the entire group earns the reward.

The literature reports many advantages associated with the use of interdependent group contingencies. The use of this type of contingency can help the teacher avoid jealousy and peer rejection because the entire class rather than the individual student can receive a reward if criteria are met (Skinner, Cashwell & Dunn, 1996). Also, interdependent group contingencies have been found to promote increased social interactions and group cooperation as a result of the class working toward a mutual goal (Skinner et al., 1996; Skinner et al., 2002). Also, by employing the same target behaviors, criteria, and reinforcement for all group members, the teacher can save a great deal of time (Skinner et al., 2002). Finally, this type of group contingency has also been found to be more effective than independent contingencies in decreasing disruptive classroom behavior (Gresham & Gresham, 1982).

**Reinforcement Schedules**

Individual and group contingency plans can vary according to the type of reinforcement schedule used. Schedules of reinforcement refer to patterns of timing for the delivery of reinforcement (Alberto & Troutman, 2009). There are four primary types of reinforcement schedules: fixed-interval, variable-interval, fixed-ratio and variable-ratio. Under fixed-interval schedules a specific amount of time (interval of time) must pass during which at least one appropriate response occurs for reinforcement to be delivered (Alberto & Troutman, 2009). Behaviors under this schedule of reinforcement
often have certain characteristics including: behaviors occur at a low rate compared to other reinforcement schedules, the length of the interval will affect the response rate (Skinner, 1953) and the rate of responding is eventually noticeably lower or stops completely after reinforcement is delivered (Alberto & Troutman, 2009).

Another type of reinforcement schedule is the variable-interval schedule, under which the intervals, after which reinforcement is delivered, are of variable length, although, their average length is consistent (Alberto & Troutman, 2009). The distinguishing characteristic of this type of schedule is the reinforcement occurs in a random or nearly random order (Cooper et al., 2007). This unpredictability helps improve consistency in the rate of response, and the behavior occurs at a more consistent and steady rate across intervals than in the fixed-interval reinforcement schedule (Alberto & Troutman, 2009). Variable interval schedules of reinforcement tend to produce low to moderate rates of response and similar to fixed interval schedules, the larger the average ratio of time, the lower the overall response rate (Cooper et al., 2007).

A third type of reinforcement schedule is the fixed-ratio schedule. In this type of reinforcement schedule, the completion of a certain number of responses is required for reinforcement to be delivered (Alberto & Troutman, 2009). Behaviors under this schedule of reinforcement often have the following characteristics: high rates of response and quick responding (Alberto & Troutman, 2009; Cooper et al., 2007). A higher ratio is likely to produce a higher rate of response, although too large of a ratio can decrease the rate of response (Cooper et al., 2007). Two potential problems may arise in using this type of reinforcement schedule. The first problem is working too quickly may lead to
problems with fluency or accuracy of responses (Alberto & Troutman, 2009). The second problem is termed post-reinforcement pause, or a lowered response rate following the delivery of reinforcement (Alberto & Troutman, 2009; Cooper et al., 2007).

Both of the aforementioned problems are eliminated with the use of the fourth type of reinforcement schedule, known as a variable ratio schedule. In a variable ratio schedule, reinforcement is delivered after the completion of a variable number of responses (Cooper et al., 2007). Behaviors under this schedule are high and consistent (Alberto & Troutman, 2009). A variable schedule of reinforcement heightens anticipation for the reward, and decreases the likelihood of a post-reinforcement pause (Cooper et al., 2007; Davis & Blankenship, 1996). Ferster and Skinner (1957) described this schedule of reinforcement as the most reinforcing schedule when compared to other schedules of reinforcement. Beaman, Stoffer, Woods, and Stoffe (1983) found behaviors placed on a variable reinforcement schedule were more resistant to extinction than behaviors placed on a regular schedule of reinforcement. The probability of reinforcement at any moment remains essentially constant, and the student adjusts by holding to a constant rate (Skinner, 1953).

**Mystery Motivator**

The Mystery Motivators is a positive reinforcement technique described by Jenson et al. (1994), which uses a variable ratio reinforcement schedule. The Mystery Motivator involves the use of unknown reinforcer(s) with an unpredictable reinforcement schedule. A growing body of research demonstrates that unpredictable or random rewards in contingency-based-interventions can significantly enhance the power of the
intervention (Rathvon, 2008). Further, randomizing contingency interventions can produce immediate and dramatic improvement in academic performance and classroom behavior (Popkin & Skinner, 2003). Students are less likely to sabotage the class’s performance because the reward (Mystery Motivator) is not known and because they are not working for a specific reinforcer (Skinner et al., 1996). The Mystery Motivator intervention aligns with the PBS philosophy because it is proactive; it involves teaching and rewarding positive behavior; it can be delivered in a group format or an individual format; and objective, measureable data are used to evaluate the intervention. Studies utilizing the Mystery Motivator have utilized independent, dependent and interdependent contingencies in an attempt to improve student’s social-emotional, behavioral and academic functioning.

**Independent Contingency Mystery Motivator Studies**

Independent group contingencies employ the same behaviors, criteria, and consequences for all the students in a classroom, but reinforcement is delivered individually, based on the individual student’s behavior or performance (Litow & Pumroy, 1975). The majority of studies conducted on the efficacy of the Mystery Motivator employ independent group contingencies for both behavioral and academic changes. Studies have assessed the effectiveness of this intervention as a tool to improve both academic and behavioral performance.

Two published studies, one unpublished dissertation, and one unpublished thesis used independent group contingency plans to analyze the effects of the Mystery Motivator on homework completion. In 1994, Moore et al. used the Mystery Motivator
intervention in an attempt to increase homework completion and accuracy in third- and fifth-grade students. Eight out of the nine participants showed an increase in daily homework completion and demonstrated an increase in homework accuracy. However, the researchers in this study did not ensure the reinforcer remained unknown. Instead, they publicly posted a reinforcer menu, thus not implementing a necessary component of the original Mystery Motivator design. This may have had an impact on the anticipation and interest a ‘mystery’ reinforcer is meant to arouse. The teacher participating in the intervention rated it as both highly acceptable and easy to implement in the classroom.

Similarly, Madaus et al. (2003) utilized an ABAB reversal with multiple baselines design, to assess the effectiveness of the Mystery Motivator as an intervention to remediate mathematics homework accuracy and completion problems in five fifth-grade students. Each student had his or her own Mystery Motivator chart for each treatment phase. The students and teacher measured homework completion and accuracy percentages daily. Results indicated the Mystery Motivator intervention was effective for all students on one or both outcome criteria, however, complete returns to baseline were seen during follow-up. Additionally, both teachers and students rated the intervention positively.

In Teta’s dissertation (2008), the effects of the Mystery Motivator intervention on homework completion and academic achievement in students diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) were assessed. The study utilized an ABAB reversal design. There were seven student participants. Results indicated the Mystery Motivator intervention was successful in improving homework completion rates for all
students. The effect sizes measuring the homework completion rates were large for six of the students and medium for one of the students. Furthermore, intervention effects were maintained through follow-up. The effect sizes measuring academic performance indicated large improvements for two students, no significance for four of the students, and a negative effect for one student.

In Deutscher’s thesis (2004), the Mystery Motivator intervention was used to increase homework completion in fifth grade students. Of the four students who participated in the study, one had a learning disability and three did not. All participants demonstrated increased rates of homework completion and accuracy from baseline to the intervention and follow-up phases. Students and their teachers also found the Mystery Motivator intervention acceptable. It is important to note, Deutscher used a reinforcement menu from which the students chose prizes, omitting a component of the original Mystery Motivator design.

Studies have also examined the effect of the independent group contingency Mystery Motivator intervention on behavior. LeBlanc (1998) compared the effectiveness of two different home-based reinforcement delivery packages for increasing appropriate classroom behavior in four kindergarten students. One package included the home-school note (a note sent home describing the child’s behavior) with the Mystery Motivator. The other package included the home-school note used with a reward menu as the home-based reinforcement delivery system. Treatment effects of the two reinforcement systems were assessed using an ABAC multiple baseline design with crossover. Results suggested both reinforcement delivery packages effectively increased appropriate classroom
behavior for each student, differential treatment effects between the interventions were not observed. However, it is important to point out that similar to Moore et al. (1994), Leblanc did not ensure the reinforcer was unknown when using the Mystery Motivator intervention. The days the student would receive reinforcement were unknown and variable, but on the days the student met the criteria and would receive reinforcement, the student chose a prize from the reinforcement menu, which could have impacted anticipation and interest.

Cowart (1997) completed a study very similar to Leblanc’s (1994) for her thesis. She examined the additive effects of praise and variable reinforcement in reducing disruptive behaviors in three elementary-aged children with a multiple baseline design. During the first part of the intervention a home school note combined with praise was utilized. The second part of the intervention combined praise and a home school note with a Mystery Motivator. Results indicated praise and a home school note alone were not effective in reducing disruptive behavior for most children. However, praise and a home school note with a Mystery Motivator were found to be effective in reducing disruptive behavior. Similar to the LaBlanc, Cowart did not ensure the reinforcers were unknown when using the Mystery Motivator intervention. The days the students would receive reinforcement were unknown and variable, but on the days the students met the criteria for reinforcement, they would choose a prize from a reinforcement menu, which could have impacted the interest that a ‘mystery’ reinforcer is meant to arouse.

Matovic (2010) examined the effectiveness of the Mystery Motivator intervention on the disruptive behaviors of two kindergarten students and two fourth grade students.
Academic performance was also monitored throughout the study to determine whether a decrease in disruptive behavior would lead to an increase in academic performance. Teacher reports and direct student observations were used to measure the effectiveness of the intervention. Results indicated the Mystery Motivator significantly reduced disruptive behavior according to teacher reports on a behavior rating scale. Academically, only the fourth grade student showed an improvement in his academic performance after the intervention phase. Interestingly, only the fourth grade teacher rated the intervention as acceptable. It is possible the individual nature of the intervention made the intervention unacceptable and difficult to manage for the kindergarten teacher. Furthermore, slight increases in disruptive behavior were noted during follow-up when compared to disruptive behavior during the intervention period.

In 1998, again utilizing an ABAB design, Kehle et al. studied the effectiveness of an intervention package that included the Mystery Motivator. They combined self-modeling, Mystery Motivators, self-reinforcement, stimulus fading, spacing, and antidepressant medication as an intervention for three children with selective mutism. In this study, the desired behavior was speech within the school environment. Participants were required to audibly ask for the contents of the envelope to receive reinforcement. Each child showed a complete cessation of selective mutism and maintained treatment gains at follow-up.

Similarly, Robinson and Sheridan (2000) targeted the remediation of undesirable behaviors in their study. They assessed the effectiveness of the Mystery Motivator intervention with a behavioral contract in remediating bedtime compliance and time spent
out of bed. Three adults and their four children, aged 3-5 years, participated in this study. Three of the four children showed substantial changes between baseline and treatment phases. The fourth child showed only slight improvement. Both, the parents and children who participated in this study rated the intervention as acceptable.

Much like Kehle et al. (1998), DeMartini et al. (2000) developed an intervention package that included a Mystery Motivator. They combined the Mystery Motivator with precision requests, antecedent strategies and response cost to remediate disruptive behaviors in two eight-year-old girls in the general education setting. A combination ABAB multiple-baseline reversal design was used to investigate treatment effects. The occurrence of disruptive behaviors was recorded. During baseline, the percentage of intervals of disruptive behaviors averaged 41%. This was reduced to an average of 20% during intervention. During the withdrawal of intervention, intervals that had incidents of disruptive behaviors increased to an average of 25%. Reinstatement of the intervention led to a further reduction of disruptive intervals with an average of 20%, showing the intervention package was effective. However, increases in disruptive behavior were again noted during follow-up.

Musser et al. completed another study that utilized a multicomponent intervention including the Mystery Motivator in 2001. The intervention was composed of a precision request program, Mystery Motivators, token economy with response cost, and antecedent strategies. The goal of the intervention was to reduce disruptive classroom behavior in three African American students, all of whom had previously received a diagnosis of Oppositional Defiant Disorder (ODD) and ADHD. Two of the student participants were
male, ages eight and ten; and one was female, age nine. The researchers wanted to investigate if the aforementioned strategies, each of which had been shown to be effective on their own, were more effective when used in combination. The intervention was implemented for a period of two weeks. Mystery Motivators were exchanged for a specific number of stickers. Response cost was also employed; one sticker was taken away if a student failed to comply with a request. All students showed a decrease in disruptive behavior throughout the intervention and at follow-up, while the control students’ behavior remained unchanged. However, disruptive behavior slightly increased during follow-up when compared to disruptive behavior during the intervention. In addition, all three students went from being in the clinically significant range pre-intervention to within the normal range at follow-up according to the Achenbach Child Behavior Checklist. The results indicated the students’ mean percentage of disruptive intervals decreased from 37% during baseline to 10% during intervention. At follow-up, the students maintained their respective treatment effects, evidencing a mean of 11% of disruptive intervals. Student and teacher satisfaction data were also positive.

Motram et al. (2000) also utilized a multi-component intervention that included a Mystery Motivator, but did not include the precision request program, with three male second-grade students diagnosed with ODD, who were mainstreamed in a private school setting. General education male students served as a control. Dependent data consisted of students’ disruptive behavior defined by noncompliance, calling out, making noises, out-of-seat, playing with objects, and orienting in a direction other than the teacher or assignment. During the intervention phase, the treatment effect was immediate and
substantial so much that the three target students’ behavior was indistinguishable from their male classmates. The treatment effect was maintained at a five-week follow-up.

**Dependent Contingency Mystery Motivator Studies**

Dependent group contingencies reinforce the entire group based on the performance of one or a couple target students meeting the criteria (Litow & Pumroy, 1975). Only one study utilized a dependent group contingency plan in combination with a Mystery Motivator. In 2000, Kehle and Bray studied the effectiveness of the Mystery Motivator on the reduction of disruptive classroom behavior in an elementary general education classroom of 23 students including three male mainstreamed students diagnosed as Seriously Emotionally Disturbed (SED). It was the teacher’s first year teaching, and the disruptive and noncompliant behavior of the three aforementioned males was serious enough that the teacher contemplated quitting. In this study, a Mystery Motivator for each of the boys was prominently displayed in front of the class. When one or more of the boys met behavior criteria in accordance with classroom rules, they would receive a point. After accumulating a predetermined number of points, the three students received their respective Mystery Motivators and each time this occurred all the students in the class also received the same reward. The intervention produced dramatic results with a 50% reduction of disruptive intervals from baseline to follow-up. However, some increases in disruptive behavior were seen during follow-up conditions as compared to intervention conditions.
Combined Dependent and Interdependent Mystery Motivator Study

Interdependent group contingencies reinforce the group, based upon the entire class meeting a specified criterion (Litow & Pumroy, 1975). Reinforcement of the group is contingent on the behavior of the whole class. Only one study has assessed the effectiveness of a combined dependent and interdependent group contingency with a Mystery Motivator (Kelshaw et al., 2000). In this study, the aforementioned strategy was used to address behavior in a special education classroom composed of five students. The dependent component was defined as the group being eligible for reinforcers based on the behavior of one randomly determined student whose behavior was evaluated. The interdependent component was defined as the entire group’s eligibility for reinforcers, based on levels of the group’s behavior. How the class would be judged to receive reinforcement for the day was based on a random selection procedure. Possible criterion for judging included a range of options: the behavior of the class as a whole, the student with greatest number of checks, the use of one particular student’s behavior, or the average number of class checks. The overriding criteria were five or less checks, which would result in receiving the Mystery Motivator. If the class average was selected as the particular criterion, then five or less class average checks would have to be met for reinforcement. If the greatest number of checks was chosen, then the student with the highest number of checks would have to have had no more than five of them to meet the criterion. Besides randomizing the specific criterion for the reinforcers, the reinforcers were also randomized to ensure they were unknown. At the end of each class period, the teacher drew the criterion for reinforcement from the first jar. If the criterion of five or
less checks was met then the entire class would receive the particular Mystery Motivator selected from a second jar. The intervention was found to be very effective in reducing disruptive behavior.

**Interdependent Mystery Motivator Studies**

Studies have also assessed the effects of various interdependent group contingency Mystery Motivator interventions, on both academic and behavioral student performance. In a dissertation by Bennett (2010), the effectiveness this type of intervention on spelling performance in a general education second grade classroom was assessed. The entire class was rewarded for averaging fewer than four incorrectly spelled words on their weekly spelling test. Even though the entire class participated in the study, only six students were targeted for baseline data collection. Results indicated the intervention was successful in increasing spelling performance with the entire class receiving reinforcement 100% of the time during intervention phases. In addition, both students and teachers rated the intervention favorably.

In 2007, Murphy et al. assessed the effectiveness of the interdependent group contingency, combined with a Mystery Motivator, in reducing disruptive behavior in a preschool class. They employed an ABAB reversal design across nine preschoolers enrolled in a Head Start classroom. The students were told they would each need to receive five or fewer checks to earn an individual reward as well as all a potential mystery class prize. For the entire class to earn the opportunity to select a Mystery Motivator, each student had to receive five or fewer checks. If the class met the criterion, one reward was randomly selected from the Mystery Motivator box. Results revealed
remarkable reductions of disruptive behavior across all nine participants, with some reversal effects, although not to the same degree as during baseline, indicating lasting treatment effects. It appeared the students had learned appropriate classroom behavior. The classroom teacher also indicated satisfaction with the intervention.

In 2010, Hoag conducted another preschool study utilizing an interdependent group contingency with a Mystery Motivator. She compared the effectiveness of the aforementioned intervention to one with a known reinforcer in four preschool classrooms. In this study, every student in the class had to meet the criteria for five or less checks to receive reinforcement. An ABAC design was employed and four classrooms were randomly assigned to an ABAC or an ACAB condition. Disruptive behavior of four identified children in each classroom was videotaped and recorded to determine the effectiveness of the interventions. Results indicated Mystery Motivators produced a moderately larger and more consistent change in disruptive behavior than known reinforcers.

In 2010 Schanding and Sterling-Turner conducted a study assessing the use of the Mystery Motivator to reduce disruptive behavior in a high school biology class. To date, it is the only study conducted in a general education high school classroom utilizing an interdependent group contingency in combination with a Mystery Motivator. In this study, the researchers evaluated the effectiveness of the Mystery Motivator intervention on the disruptive classroom behavior in a biology class. An ABAB single-case design was used to evaluate the effects of this intervention on the behavior of three students identified as exhibiting problem behaviors. The effects of the intervention on the
behavior of non-identified students were also assessed through a sampling method. The teacher posted two envelopes on the board instead of displaying a weekly chart. One envelope, the reward envelope, contained the cards with names of the rewards the class could earn. The other envelope contained paper slips with either the letter M or X written on each slip. Results indicated a decrease in problem behaviors for the three identified students. Also, general decreases in problem classroom behavior were noted through a time sampling methodology.

To date, only one study evaluated the Mystery Motivator as a classroom behavior intervention in a general education elementary school classroom (Kraemer et al., 2012). Two fifth-grade classes, one math and one English, were selected from a suburban school district in Ohio. The student population at the school selected for the study was 92.6% White, non-Hispanic; 3.2% Asian/Pacific Islander; and 2.8% multi-racial. It was reported that 8.9% of the students came from economically disadvantaged homes and 14.4% received special education services. An ABCACBA design was utilized to compare the efficacy of the Mystery Motivator and the Get ‘Em on Task intervention on disruptive classroom behavior. The Get ‘Em on Task intervention is a computer-signaling program that helps teachers reward individual students based on individualized auditory signals monitoring student behavior. If a student is on-task when their individualized signal sounds, they earn points, which are recorded on individualized point sheets. Points earned during the Get ‘Em on Task intervention were used to make purchases from a reinforcement menu at the end of a week.
Data recording procedures during the Mystery Motivator intervention are not as clearly described. Kraemer et al. does not explain how data were collected to assess on-task and off-task behavior by the teacher, when the outside observer was not present and how eligibility for prizes was determined. It appears that quantitative, objective data were gathered only twice a week when the outside observer was present. A time sampling methodology utilizing the Behavioral Observation of Students in Schools (BOSS) measured academic engagement as well as off-task behavior. The results of the BOSS were used to determine the effectiveness of the interventions, but it does not appear the same BOSS data were used to determine eligibility for prizes. The study was implemented for 14-weeks, with the Mystery Motivator in place for six of those 14-weeks. Results indicated both interventions were very successful in reducing off-task behavior and both interventions were deemed acceptable by the teacher and “somewhat liked” by the students. Overall, the Get ‘Em on Task intervention was rated slightly higher by both the teacher and the students and was found to be about 16.75% more effective at reducing disruptive student behavior.

There were, however, serious limitations in the research design that could undermine the validity of these results. First of all, the study’s ABCACBA design does not account for multiple treatment inference or carry over effects. If multiple treatments (Mystery Motivator and the Get ‘Em on Task interventions) are given to the same subjects, it is difficult to control for the effects of the prior treatment (Kazdin, 2011). Also, both classes followed the same ABCACBA research design, thus there was a lack
of counterbalancing. As a result, the findings can only be generalized to those exposed to the same treatments in the same order of presentation (Kazdin, 2011).

Another limitation of this study is the collection of data through momentary time sampling method. In momentary time sampling, the observer looks up and records whether a behavior occurs or does not occur during a time interval. The BOSS was used to record the behavior of individual students in 15-second intervals. Once an individual student was observed for one 15-second interval, the researcher moved on to the next consecutive student. A weakness of momentary time sampling is it only provides an estimate of behavior as opposed to documenting every occurrence (Snowden, 2004). While the observer focused on one student the rest of the class might have been engaged in numerous disruptive behaviors that were not accounted for. Thus, this recording method may not be the best choice for gathering data on the behavior of the whole classroom.

Furthermore, Kraemer et al. (2012) did not ensure the reinforcers remained unknown when using the Mystery Motivator intervention. The reinforcement days were unknown to the students and variable, however, when the teacher judged that criteria for reinforcement were met, students selected prizes from a known reinforcement menu. Thus, an essential component of the original Mystery Motivator design was not implemented and the element of surprise was weakened. Also, the reinforcement menu was not the same for both interventions. As a result, it is possible the reinforcers on one menu could have been more or less powerful than those on the other. Additionally, in the *Get 'Em on Task* intervention, students were able to select an individual prize from the
classroom store whereas the entire class received the same prize during the Mystery Motivator intervention. Without controlling for the aforementioned variables it is difficult to truly compare the efficacy of the two interventions.

Finally, the Kraemer et al. (2012) study was conducted on a very homogeneous sample of students. It took place in only one school, with only one grade level and in only two classrooms. The sample consisted of primarily Caucasian (92.6%) students in a suburban district where only 8.9% of the students received a free or reduced priced lunch. The lack of diversity of the students participating in the study further limits the generalizability of the results.

**Intervention Acceptability**

Many educational interventions, such as the ones described above, have been proven to be successful in achieving desired goals. Despite the fact there are an abundant amount of research-based interventions; school personnel do not have a good record of implementing such interventions with fidelity (Schoenwald & Hoagwood, 2001; Walker, 2004). In fact, there is a very large discrepancy between the availability of evidence-based interventions and their effective use in the field of education (Walker, 2004).

However, schools should not take all the blame. The developers of interventions must share the responsibility by creating a context, which will allow for greater adaptation of research-based interventions (Schoenwald & Hoagwood, 2001; Walker, 2004).

In order for evidence-based interventions to be accepted and implemented by educators, they not only have to be efficacious in producing results, but must also be socially valid. Social validity is now viewed as a critical consideration in research
evaluating interventions in applied behavioral analysis, mental health and school psychology (Gresham & Lopez, 1996; Kazdin, 1977; Wolf, 1978). Social validity refers to judgments about the social importance of interventions on three levels: (1) treatment must be socially significant so the behaviors targeted are deemed socially important and relevant, (2) treatment procedures must be considered socially appropriate, and (3) the effects must be socially important or have meaningful significance (Wolfe, 1978). In order for teachers to implement interventions with integrity, they must find them acceptable for use in their classrooms and deem them to be appropriate, effective, and fair. In order for researchers to develop interventions that will be utilized on a wider scale, teachers’ overall assessment of these interventions, beyond their effectiveness, must be taken into consideration (Nastasi & Truscott, 2000). It is also helpful if students rate interventions they are a part of as satisfactory. Including students’ ratings of interventions increases the social validity of these interventions even further. In fact, in recent years, treatment acceptability has become increasingly important and viewed as a critical component of treatment effectiveness (Finn & Sladeczek, 2001).

**Summary**

The bulk of the existing research on the effectiveness of Mystery Motivators evaluates the effect of the intervention either on targeted individual students or targeted groups of students (DeMartini et al., 2000; Kehle & Bray, 2000; Kehle et al., 1998; LeBlanc, 1998; Maduas et al., 1994; Moore et al., 1994; Motram et al., 2000; Musser et al., 2001; Robinson & Sheridan, 2000; Theodore et al., 2000) and often these students have already been identified for special education services (Kehle & Bray, 2000;
Motram et al., 2000; Musser et al., 2001; Theodore et al., 2000) or as an at-risk group (Murphy et al., 2007). Only a few studies have evaluated the effectiveness of the Mystery Motivator in an interdependent context on the performance of all the students in a classroom (Bennett, 2010; Hoag, 2010; Kraemer et al., 2012; Murphy et al., 2007; Schanding & Sterling-Turner, 2010). Of these studies, only four examined the effects of the Mystery Motivator intervention on disruptive classroom behavior (Kraemer, 2012; Murphy et al., 2007; Schanding & Sterling-Turner, 2010). However, each of these studies had many important limitations.

Both the Schanding and Sterling-Turner (2010) and Kraemer et al. (2012) studies evaluated the effectiveness of the intervention on classroom behavior through a sampling method, which cannot be considered the most accurate means of data collection (Snowden, 2004). In the Schanding and Sterling-Turner study (2010), non-target students were randomly chosen and behavior was recorded to ascertain the effectiveness of the intervention on the class. In the Kraemer et al. (2012) study, data was gathered through biweekly observations using the BOSS on only a portion of the classroom. A sampling method provides only an estimate of behavior as opposed to documenting every occurrence (Snowden, 2004). It is possible that while the researcher focused on one student, the rest of the class might have been engaged in numerous disruptive behaviors that were not accounted for. Therefore use of a sampling method appears to be a weakness in the studies where the behavior of a whole classroom is the unit of analysis. The present study compensated for the aforementioned weakness by collecting disruptive classroom behavior data through a frequency count.
Another important limitation of many of the interdependent Mystery Motivator studies targeting disruptive classroom behavior involves burdensome data collection techniques. In the two studies (Hoag, 2010; Murphy et al., 2007) conducted in a preschool setting, teachers were required to keep track of the behavior of individual students by marking checks next to students’ names whenever they engaged in disruptive behavior. In each study teachers identified data collection techniques as cumbersome. Although, there were only nine students in the study conducted in the Head Start classroom, the teacher indicated she sometimes had difficulty marking checks when she was engaged in other activities (Murphy et al., 2007). When it was not possible for her to stop the activity to record disruptive behavior, she recorded the behavior at the first opportunity, which likely resulted in inconsistencies (Murphy et al., 2007). In the Kraemer et al. study, it appears that quantitative, objective data were gathered only when the outside observer was present during the Mystery Motivator. This was only for two out of the five days the intervention was in place. In the Schanding and Sterling-Turner (2010) an outside observer was also required to gather data to ascertain the effectiveness of the intervention on non-target students.

In the days of increasing resource and budget cuts, many teachers cannot expect to have an outside observer, such a school psychologist or intervention specialist, to come into their classroom to gather data. Most teachers are expected to gather data, particularly Tier 1 data, on their own. The expectation for teachers to collect their own data is especially likely to occur in low resource areas, which often have greater amounts of minority and low SES students, who have the greatest need for proactive behavior
management strategies (Kellam et al., 1998). Limited exposure of teachers to support staff in schools is illustrated by a study that found that 21.4% of the kindergarten to fifth grade teachers and 16.4% of teachers overall indicated they had no prior contact with a school psychologist (Erchul, Raven & Whichard, 2001).

To compensate for the aforementioned concerns regarding data collection procedures, the teachers in this study were provided with tally counters to record the frequency of disruptive behavior in their classrooms. Teachers did not have to stop instruction to record data. They simply added a tally to the counter, which they held in their hand by pressing on a small button. Furthermore, the tally counter provided a discriminative stimulus to notify students of disruptive behavior by making a clicking sound. This stimulus was much less intrusive to the flow of classroom activities than verbal redirection. Thus, the tally counter not only simplified data recording procedures, but also allowed teachers to provide behavioral feedback and consequences with consistency, immediacy and ease.

Another limitation of prior Mystery Motivator studies is that they were implemented over a short time period, usually around two to three weeks (DeMartini-Scully et al., 2000; Kehle et al., 2000; Kraemer, 2012; Madaus et al., 2003; Matovic, 2012; Musser et al., 2001). As a result, the studies did not ascertain whether the efficacy of the intervention did not decrease over time, as students become bored with the intervention due to the satiation effect (Hall & Hall, 1980). In present study, the intervention was implemented for eight weeks, which allowed for a better understanding of the Mystery Motivator’s ability to maintain its effectiveness over time.
Furthermore, only one study which examined the effect of the Mystery Motivator intervention on classroom behavior was conducted within an elementary, general-education setting (Kraemer et al., 2012). This study compared the effects of the Mystery Motivator and the Get ‘Em on Task interventions on disruptive classroom behavior using an alternating treatment design (Kraemer et al., 2012). Many limitations were identified in this study including: reinforcers did not remain a mystery; multiple treatment effects were not accounted for; the two interventions were presented in the same order in both classrooms, which led to a lack of counterbalancing; and the interventions were only implemented in only two classrooms, in one grade level with a very homogenous student population. In the present study, the Mystery Motivator Calendar intervention was implemented in seven different schools across eight different classrooms. The students in this study came from very diverse backgrounds. The heterogeneity of the sample in this study allowed for greater generalizability of the results. Also, because only one intervention was evaluated in the present study, the threats of multiple treatment effect and carry over effect, did not apply.

Beyond addressing the aforementioned limitations, this study also included individual handouts for students to serve as reminders of intervention goals, a component not typically used in Mystery Motivator studies. Also, a monthly Mystery Motivator Calendar was posted for the whole class to see and to monitor progress towards behavioral goals. Such visual cues are important given that Jones and Van Houte (1985) have shown public posting of an intervention can reduce disruptive behaviors and improve academic engagement.
Furthermore, continued research in the area of group contingency interdependent Mystery Motivator interventions targeting disruptive classroom behavior is still needed because the four existing studies utilized single case design. Often these studies were only conducted in one or two classrooms (Kraemer, 2004; Murphy et al., 2007; Schandling & Sterling-Turner, 2010) and at most in four classrooms (Hoag, 2010). Inherent in single case designs are limitations to external validity (Kazdin, 2011). Replications in multiple contexts, with populations composed of varying ages, grades, ethnic and socio-economic backgrounds are necessary to provide a stronger basis for observed relationship(s).
CHAPTER THREE

METHODOLOGY

Procedures

Recruitment

School administrators, school psychologists, and principals in urban and suburban areas of a large Midwestern city were contacted by email to recruit eight classroom participants (refer to Appendix A for recruitment email). A total of 11 classrooms responded with interest to participate in the study. Of these classrooms, one did not meet the inclusion criteria for disruptive classroom behavior. For the remaining ten classrooms, the intervention was implemented in the first eight that responded, while consultative services and intervention materials were provided for the remaining two classrooms.

Inclusion Criteria for Participant Selection

Geographical areas, which contained culturally, racially, socio-economically, and linguistically diverse student populations, were targeted. Kindergarten to fifth grade general education classrooms in which teachers were experiencing difficulties with disruptive student behavior was sought.

Consent

First, letters of cooperation were attained from the school principals (see Appendix B). Next, teacher consent was obtained (see Appendix B). Once, teacher
consent was obtained, parental consent was sought. The local university IRB granted a waiver of documented parental consent. Thus, parent signatures were not required for student study participation. Instead, parents were presented with the option to opt out of the study at any time. In accordance with the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99), teachers distributed parental consent forms, which the students were required to bring home. Additionally, when email information was available, teachers contacted parents by email with the consent documentation. Also, once the intervention was explained to students, verbal student consent was garnered.

**Participants and Setting**

The study was implemented in a total of eight classrooms, which included: three kindergarten classrooms, two first grade classrooms, one third grade classroom, and two fourth grade classrooms. Demographic data were gathered for each classroom including: (a) the number of males and females; (b) racial background of students; (c) the number of students receiving special education services; (d) the number of students undergoing a functional behavior assessment (FBA) or a case study for special education; (e) the number of students on individual behavior plans; (f) the number of students receiving English Language Learner services; (g) the number of students meeting local education standards, the number of students at-risk or below; (h) the number of students receiving a free or reduced price lunch (economic disadvantage status); (i) the number of school staff members in the classroom; (j) and current behavior management systems used. This data is presented in Table 1 below.
<table>
<thead>
<tr>
<th>Table 1. Classroom Participant Demographic Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
</tr>
<tr>
<td># Students</td>
</tr>
<tr>
<td>Grade</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Racial/Ethnic</td>
</tr>
<tr>
<td>African Amer.</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Biracial</td>
</tr>
<tr>
<td>Caucasian</td>
</tr>
<tr>
<td>Latino/Hispanic</td>
</tr>
<tr>
<td>Pakistani</td>
</tr>
<tr>
<td>Support services</td>
</tr>
<tr>
<td>Behavior Plan</td>
</tr>
<tr>
<td>Case Study</td>
</tr>
<tr>
<td>E.L.L.</td>
</tr>
<tr>
<td>I.E.P.</td>
</tr>
<tr>
<td>Standards</td>
</tr>
<tr>
<td>At &amp; above</td>
</tr>
<tr>
<td>At-risk</td>
</tr>
<tr>
<td>Below</td>
</tr>
<tr>
<td>Free/reduced</td>
</tr>
<tr>
<td># Staff</td>
</tr>
<tr>
<td>Current behavior systems</td>
</tr>
<tr>
<td>- PBS</td>
</tr>
<tr>
<td>- Olweus</td>
</tr>
<tr>
<td>- Colored cards</td>
</tr>
<tr>
<td>- Marble jar</td>
</tr>
</tbody>
</table>

Note: Behavior Plan indicates the number of students on individual behavior plans. Standards refers to educational tests used to assess student performance and categorize students at or above standards, at-risk or below standards. Free/reduced refers to the number of students receiving a free or reduced priced lunch. # Staff indicates the number of school personnel in each classroom. PBS stands for Positive Behavior Support. Olweus stands for the Olweus Bullying Prevention Program. Colored cards refers to a progressive behavioral system used in classrooms described in greater detail in the Participants, Classroom A section below. * In Classroom F three students were on individual behavior plans at the start of the intervention and one at completion of the intervention. Also, standards data in Classroom F is from the beginning of the school year when 26, not 24 students were enrolled. Standards data in Classroom H is reported for language arts, data was only available for 28 students. When the intervention began there were 24 students in Classroom F, when it ended there were only 23.
In addition, when available, school demographic data were obtained online from the Illinois Interactive Report Card (Illinois Interactive Report Card, 2011). Information gathered included the following: (a) the racial profile of the school, (b) teacher to student ratios, (c) average class size (d) the percentage of students meeting standards on the Illinois Standards Achievement Test, (f) adequate yearly progress (AYP) status, (g) and the percentage of students coming from homes described as economically disadvantaged. This data is presented in Table 2 below.

Table 2. School-Wide Demographic Data from Illinois Interactive Report Card

<table>
<thead>
<tr>
<th>Class:</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Racial profile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>14.9%</td>
<td>1.3%</td>
<td>17.6%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Black</td>
<td>2%</td>
<td>3.3%</td>
<td>0.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>67.6%</td>
<td>91.6%</td>
<td>76.6%</td>
<td>76%</td>
</tr>
<tr>
<td>Asian</td>
<td>14.6%</td>
<td>0%</td>
<td>4.7%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Multi-Racial</td>
<td>0.5%</td>
<td>3.8%</td>
<td>0.5%</td>
<td>6.1%</td>
</tr>
<tr>
<td>American</td>
<td>0.5%</td>
<td>0%</td>
<td>0.3%</td>
<td>0%</td>
</tr>
<tr>
<td>Indian</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISAT pass rate</td>
<td>82.7%</td>
<td>60.9%</td>
<td>75.8%</td>
<td>67.7%</td>
</tr>
<tr>
<td>AYP Status</td>
<td>Not Met</td>
<td>Not Met</td>
<td>Not Met</td>
<td>Not Met</td>
</tr>
<tr>
<td>Average Class Size</td>
<td>22</td>
<td>24</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>68%</td>
<td>83%</td>
<td>75%</td>
<td>68%</td>
</tr>
<tr>
<td>Student/Teacher Ratio*</td>
<td>15.5:1</td>
<td>20.5:1</td>
<td>15.5:1</td>
<td>20.5:1</td>
</tr>
</tbody>
</table>

Note: Data not available through Illinois Interactive Report Card for classrooms: A, B, C, & G.
* Student to teacher ratios reflect school district averages.
**Classroom A.** Classroom A was a kindergarten classroom in a parochial school in a suburb of a large Midwestern city. The classroom was comprised of 25 students, including 13 males and 12 females. One student was African American, six students were of Hispanic/Latino origin, one student was of Asian origin, and the remaining 17 students were Caucasian. In addition to the primary teacher, there was a full-time classroom aide. There were no students with Individualized Education Plans, no students who received English Language Learner Services, no students who received a free or reduced priced lunch, no students were undergoing a case study evaluation or FBA, and no students were on an individual behavior plan. Information regarding the number of students performing at or below standards was not available. Also, no data on school demographics were available on the Illinois Interactive Report Card.

The school had a Positive Behavior Support system implemented on a school-wide basis. Students were able to earn tickets for engaging in desired behavior. Each time a student received a ticket for desired behavior he/she would put his/her name on it and put it in a jar. At the end of each week, a raffle occurred in each classroom in which students with tickets had the opportunity to earn small age appropriate prizes. The Olweus Bullying Prevention Program had also been implemented school-wide. In addition, a progressive behavioral system was in place. A pocket chart containing a set of colored cards for each student was displayed. Cards were green, yellow, and red. The colors represented a progression of consequences from green to yellow to red for breaking rules. If the teacher deemed a student was not behaving appropriately, then the
student would have to change his/her color. Students had the opportunity to move backwards, from red to yellow, and yellow to green for good behavior.

**Classroom B.** Classroom B was also a kindergarten classroom located in a parochial school in a suburb of a large Midwestern city. The classroom was comprised of 24 students including 14 males and 10 females. One student was African American, two students were of mixed racial backgrounds, one student was of Asian origin, and the remaining 20 students were Caucasian. In addition to the primary teacher, there was a full-time classroom aide. There were no students with Individualized Education Plans, no students who received English Language Learner Services, no students who received a free or reduced priced lunch, no students undergoing a case study evaluation or FBA, and no students with an individual behavior plan. No information regarding the number of students who were performing at or below standards was available. Also, no data on school demographics were available on the Illinois Interactive Report Card. The behavioral supports in Classroom B were equivalent to the behavioral supports in Classroom A.

**Classroom C.** Classroom C was a kindergarten classroom located in a parochial school in a large Midwestern city. The classroom was comprised of 23 students, including 13 males and 10 females. One student was African American, three students were of mixed racial backgrounds, two students were of Asian origin, and the remaining 17 students were Caucasian. In addition to the primary teacher, there was a full-time classroom aide. There were no students with Individualized Education Plans, no students who received English Language Learner Services, no students who received a free or
reduced priced lunch; no students were undergoing a case study or an FBA, and no students with an individual behavior plan. Diagnostic Indicators of Basic Early Literacy Skills (DIBELS) were used as a screening tool to assess student academic performance.

According to DIBELS data, 20 students were performing at or above standards and three were found to be at-risk. This school was in its second year of implementing Schoolwide Positive Behavior Support. In addition, the same progressive behavioral system involving colored cards was in place as in the aforementioned classrooms. All students who stayed on green could earn a marble for the class; marbles were also used to reinforce desired behavior in the classroom. When the marble jar was filled, the students could earn a sticker, which they would place on their personal folder. No data on school demographics were available on the Illinois Interactive Report Card.

**Classroom D.** Classroom D was a first grade public school classroom located in a suburb of a large Midwestern city. The classroom was comprised of 17 students including nine males and eight females. One student was African American, three students were of mixed racial backgrounds, five students were of Hispanic/Latino origin, two students were Pakistani, and the remaining nine students were Caucasian. One student had an Individualized Education Plan, 11 students received English Language Learner Services, and 12 students received a free or reduced priced lunch. Two students were on individual behavior plans. No students were undergoing a case study or an FBA. According to AIMSweb Oral Reading Fluency data, 15 students were performing at or above standards, while two students were performing below standards. The behavioral supports in place at this school included the Olweus Bullying Prevention Program and the colored
card progressive system. Students also had the opportunity to earn classroom money based on behavior throughout the day. If they stayed on green the entire day they could earn classroom money. If they moved to yellow or red they would have to pay their teacher with the classroom money. When caught doing something helpful, they also had the opportunity to earn classroom money. The classroom money could be used at the end of the week to buy various prizes from the classroom store.

Schoolwide data available through the Illinois Interactive Report Card indicated the racial composition of this school was as follows: 14.9% White, 2% Black, 67.6% Hispanic/Latino, 14.6% Asian, 0.5% multi-racial and 0.5% American Indian. The percentage of students meeting or exceeding standards on the Illinois Standard Achievement Test (ISAT) was 82.7%. Overall, the school had not met Adequate Yearly Progress (AYP) goals within the last school year. The average student to teacher ratio was 15.5:1 in the district; individual school data were not available. The average class size was approximately 22 students. Finally, 68% of the students attending the school came from homes described as economically disadvantaged.

**Classroom E.** Classroom E was a first grade public school classroom located in a far suburb of a large Midwestern city. At the end of the intervention, the classroom was comprised of 31 students including 16 males and 15 females. Throughout the course of the intervention, two students left the classroom, one female and one male, and another three students, two males and one female joined the classroom. At the end of the intervention, nine students were African American, three were of mixed racial background, 19 were of Hispanic/Latino origin, and one student was Caucasian. Four
students in the classroom received special education services through an Individualized Education Program. One student was undergoing a case study evaluation for special education. None of the students received English Language Learner services. Five students in this classroom had behavior contracts that allowed them to earn daily points for positive behavior. Twenty-eight students qualified for a free or reduced priced lunch. Fountas and Pinnell assessments were used to measure language skills. According to this assessment, ten students were performing at or above standards, eight were at-risk, and 11 were performing below grade level.

Behavior management systems already in place included school-wide PBS. Students could earn tickets for displaying desired behaviors. These tickets could be used to purchase entry into school events or to purchase items in the school store each Friday. Finally, a progressive behavioral system, utilizing colored cards, was in place similar to the aforementioned schools.

Schoolwide data available through the Illinois Interactive Report Card indicated the racial composition of this school was as follows: 1.3% White, 3.3% Black, 91.6% Hispanic, and 3.8% multi-racial. The percentage of students meeting or exceeding standards on the ISAT was 60.9%. Overall, the school had not met AYP goals within the last school year. The average student to teacher ratio was 20.5:1 in the district; individual school data were not available. The average class size was approximately 24 students. Finally, 83% of the students attending the school come from homes described to be economically disadvantaged.
Classroom F. Classroom F was a third grade public school classroom located in a suburb of a large Midwestern city. The classroom was comprised of 24 students including 12 males and 12 females at the start of the intervention. All of the students in this classroom were of Hispanic origin. Midway through the intervention, one male student left the classroom. Four students in the classroom received special education services through an Individualized Education Program. One student in the class received English Language Learner services. At the start of the intervention three students had individual behavior plans, but at the end of the intervention there was only one student with an individual behavior plan. All of the students in this classroom qualified for a free or reduced priced lunch. No student was undergoing a case study or FBA. Markers of Academic Progress (MAP) data were used to assess student performance. According to Fall 2011 data, when 26 students were enrolled in the class, 15 students were performing at or above standards, four were at-risk, and another seven were performing below standards.

Behavior management systems already in place included Schoolwide PBS however, students could not earn tickets or tokens for appropriate behavior. Only schoolwide behavioral expectations had been developed. In addition, this classroom had the same progressive behavioral system, which utilized colored cards, as the aforementioned classrooms.

Schoolwide data available through the Illinois Interactive Report Card indicated the racial make up of this school was as follows: 17.6% White, 0.3% Black, 76.6% Hispanic, 4.7% Asian, 0.3% American Indian and 0.5% multi-racial. The percentage of
students meeting or exceeding standards on the ISAT was 75.8%. Overall, the school had not met AYP goals within the last school year. The average student to teacher ratio was 15.5:1 in the district; school data were not available. The average class size was approximately 25 students. Finally, 75% of the students attending the school came from homes described as economically disadvantaged.

**Classroom G.** Classroom G was a fourth grade parochial school classroom located in a large Midwestern city. The classroom was comprised of 24 students including 15 males and nine females. One student was African American, four students were of mixed racial background, four students were of Asian origin, seven students were of Hispanic/Latino background, and the remaining eight students were Caucasian. Two students in this classroom had an Individualized Education Program. An additional student had been recommended for evaluation at the local public school for significant symptoms associated with Attention Deficit Hyperactive Disorder. None of the students received English Language Learner services. Seven of the students in this classroom qualified for a free or reduced priced lunch. According to AIMSweb Oral Reading Fluency measures, 18 students were performing at or above standards, four were considered at-risk, and two were performing below standards. The same colored card progressive behavioral system was in place, as in the aforementioned classrooms, however, students could not move backwards for displaying appropriate behavior. No data on school demographics were available through the Illinois Interactive Report Card.

**Classroom H.** Classroom H was a fourth grade public school classroom located in a far suburb of a large Midwestern city. The classroom was comprised of 29 students,
including 18 males and 11 females. Four students were African American, 24 students were of Hispanic/Latino origin and one student was Caucasian. Over the course of the intervention, one Hispanic male student joined the classroom and one Hispanic male student left the classroom. Four students in this classroom had an Individualized Education Program. At the time the intervention was being implemented, one student was undergoing a case study evaluation for an emotional and behavioral disorder. The same student had an individual behavior plan. After, the intervention was completed, as the result of a full case study, this student was found to be eligible for an Individual Education Program and placement in a classroom for students with emotional and behavioral disorders was recommended. None of the students received English Language Learner services. Twenty-three students qualified for a free or reduced priced lunch. The most recent student performance data available was from the ThinkLink Discovery Education Test, which was completed in the winter of 2011. Based on this data, two students performed below standards and 26 met or exceeded standards in mathematics. In Reading/Language Arts, five students performed below standards while 23 met or exceeded standards. Data was only available for 28 students.

The school had implemented PBS. School-wide behavioral expectations had been developed, but the school no longer provided individual tickets to students for good behavior because teachers found this practice to be too cumbersome. Instead, entire classrooms could receive tickets when they were displaying good behavior. Classroom prizes were rewarded according to the number of tickets earned.
Schoolwide data available through the Illinois Interactive Report Card indicated the racial composition of this school was as follows: 3.4% White, 13.1% Black, 76% Hispanic, 1.5% Asian, and 6.1% multi-racial. The percentage of students meeting or exceeding standards on the ISAT was 67.6%. Overall, the school had not met AYP goals within the last school year. The average student to teacher ratio was 20.5:1 in the district; individual school data were not available. The average class size was approximately 24 students. Finally, 68% of the students attending the school came from homes described as economically disadvantaged.

Research Design

Single case designs allow for the collection of data repeatedly across time (Kazdin, 2011). These designs also allow researchers to evaluate the effects of the intervention condition compared to baseline data within and across participants (Kazdin, 2011). Such designs are well suited for the assessment of an intervention’s effectiveness within a classroom (Kazdin, 2011). In this research project, the classroom was the individual unit of analysis.

More specifically, this study employed a single-case, ABAB, changing-criterion design, and was replicated across eight classrooms. The first phase was baseline, the second was intervention, the third was withdrawal, the fourth was reinstatement and the fifth was a follow up. ABAB designs are well suited to assess functional relationships between performance and strategies for increasing or decreasing behavioral rates (Sulzer-Azaroff & Mayer, 1991). In an ABAB design the effects of the intervention are clear when performance improves during the first intervention phase, reverts when intervention
is withdrawn, and improves again when treatment is reinstated (Kazdin, 2011). In this study, complete reversals to baseline were not expected. It was hypothesized students would learn replacement behaviors during intervention, and they would not unlearn these behaviors during reversal as in the Murphy et al. 2007 study. However, it was expected that when potential reinforcement was removed, student motivation would decrease, and as a result increases in disruptive behavior would be seen.

Furthermore, this study utilized a changing criterion component between intervention phases. Changing criterion designs are appropriate for interventions in which there are specific criteria for earning rewarding consequences (Kazdin, 2011). If performance meets or surpasses the criterion, reinforcement/consequences are provided and a slightly more stringent criterion is implemented, this process is then repeated across sub-phases until the desired level of performance is met (Kazdin, 2011). Before changes in sub-phases are made, behavior should stabilize (Kazdin, 2011). In changing criterion designs, behavior changes occur in a stepwise fashion that directly corresponds to changes in criteria (Kazdin, 2011). In this study, significant reductions of 50% or more in disruptive behavior were required between sub-phases for students to earn reinforcement. Once students were able to achieve the set criterion for a period of approximately ten school days in total, then the criterion was reduced again by 50%. Changing criterion designs are better suited to situations in which behavior is altered gradually toward a terminal goal (Kazdin, 2011). However, because reductions in the criteria between phases for rewarding consequences were so significant in this study, it was not expected behavior would decrease in the quintessential stepwise process characteristic of changing
criterion designs. More frequent and less stringent changes in criteria were not made in this study in order to make the implementation of the intervention more teacher friendly.

Variables

**Independent variable.** The independent variable was the Mystery Motivator Calendar intervention (see Appendix E). On certain days, interspersed at irregular intervals, the letter “M” was written under a small square of paper. An “M” signified reinforcement was available for the designated time period. A random number generator was used to determine the calendar days on which “Ms” were available. For each classroom, reinforcement was available for 60% of the school days during the first month of intervention, and 50% of the school days during the second month on intervention (Sprick, 2011). Squares of paper were used to cover up the days on the calendar so the students could not see whether an “M” was available on that particular day. The days on which prizes were available were a mystery, and the prizes themselves were also a mystery. The available prizes were described on individual note cards placed in a manila envelope with a large question mark on it. The note cards were randomly drawn on the days students earned reinforcement. This process is described in greater detail in the intervention procedures section.

**Dependent variables.** Disruptive behavior was the primary dependent variable. Disruptive behavior was defined and operationalized on a classroom-by-classroom basis by the Problem Identification Interview (PII) (Kratochwill & Bergan, 1990). The time period of the day during which the intervention would be implemented was also identified during the PII. Teachers were asked to identify a time period during which they
believed students struggled most with behavior. (See Table 3 for the description of operationalized disruptive and replacement behavior per a classroom.)

Table 3. Targeted Disruptive and Replacement Behavior per a Classroom

<table>
<thead>
<tr>
<th>Class</th>
<th>Target Behavior</th>
<th>Replacement Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Calling out/Talking</td>
<td>Raising hand/No talking/Only answer in unison when cued</td>
</tr>
<tr>
<td></td>
<td>Sitting inappropriately</td>
<td>Sit cross-legged, hands to yourself</td>
</tr>
<tr>
<td>B</td>
<td>Calling out/Talking</td>
<td>Raising hand/No talking/Only answer in unison when cued</td>
</tr>
<tr>
<td></td>
<td>Sitting inappropriately</td>
<td>Sit cross-legged, hands to yourself</td>
</tr>
<tr>
<td>C</td>
<td>Calling out</td>
<td>Raising hand</td>
</tr>
<tr>
<td></td>
<td>Off-task during calendar</td>
<td>Eyes on teacher/speaker during calendar</td>
</tr>
<tr>
<td>D</td>
<td>Calling out/Talking</td>
<td>Raising hand/No talking</td>
</tr>
<tr>
<td></td>
<td>Sitting inappropriately</td>
<td>Sit cross-legged, hands to yourself</td>
</tr>
<tr>
<td>E</td>
<td>Calling out/Talking</td>
<td>Raising hand/No talking/Only answer in unison when cued</td>
</tr>
<tr>
<td></td>
<td>Desk open during instruction</td>
<td>Desks closed during instruction</td>
</tr>
<tr>
<td></td>
<td>Getting out of seat</td>
<td>Stay seated</td>
</tr>
<tr>
<td>F</td>
<td>Materials unprepared</td>
<td>Materials ready</td>
</tr>
<tr>
<td></td>
<td>Talking during instruction</td>
<td>When teacher speaks, stay quiet/talk only while proofreading</td>
</tr>
<tr>
<td></td>
<td>Off-task during writing</td>
<td>Writing, eyes on assignment or teacher</td>
</tr>
<tr>
<td>G</td>
<td>Calling out/Talking</td>
<td>Raising hand/No talking/Only answer in unison when cued</td>
</tr>
<tr>
<td></td>
<td>Getting out of seat</td>
<td>Stay seated</td>
</tr>
<tr>
<td></td>
<td>Off-task during reading/independent work</td>
<td>Eyes on book during reading, eyes on assignment during independent work</td>
</tr>
<tr>
<td>H</td>
<td>Calling out/Talking</td>
<td>Raising hand/No talking/Only answer in unison when cued</td>
</tr>
<tr>
<td></td>
<td>Getting out of seat</td>
<td>Stay seated</td>
</tr>
</tbody>
</table>

Teacher intervention acceptability, as determined by a modified version of the Intervention Rating Profile-20 (IRP-20) (Witt & Marten, 1983), was the second
dependent variable. The IRP-20 is discussed in greater detail in the instrumentation section that follows.

Student intervention satisfaction was the final dependent variable. Student ratings were assessed through two modified versions of the Children’s Intervention Rating Profile (CIRP) (Witt & Elliott, 1985). The Children’s Intervention Rating Profile discussed in greater detail in the section below.

**Instrumentation**

**Problem identification interview.** Disruptive classroom behavior was identified by the classroom teacher through a modified version of the problem identification interview (PII) by Kratochwill and Bergan (1990) (see Appendix C).

**Event recording.** During the PII, the teacher and researcher collaboratively identified and operationalized disruptive behaviors that typically occurred in the classroom. The frequency of disruptive behavior was tallied and recorded in each classroom from baseline through intervention to follow-up. During intervention, when a student engaged in disruptive behavior he/she would receive a tally on the tally counter. When a tally was added to the counter, it would make a clicking noise, which served as a discriminative stimulus. The teacher was advised to hold the tally counter in the direction of the student engaging in disruptive behavior. If the student appeared to unaware that he/she was engaging in the disruptive behavior, the teacher would redirect the student to the identified replacement behavior. If the student continued to engage in the disruptive behavior, four seconds after being redirected by the teacher, he/she would receive another
tally. Teachers collected data every school day the intervention was in place during the intervention period. The tally counter was not used during baseline and reversal phases.

**Intervention Rating profile.** A modified version of the Intervention Rating Profile-20 (IRP-20) developed by Witt and Marten (1983) was used to assess teachers’ satisfaction with the intervention. The IRP-20 was specifically designed to assess the acceptability of school-based interventions (Finn & Sladeczek, 2001; Witt & Martens, 1983). The items on the IRP-20 focused on assessing teacher perceptions regarding the practicality and appropriateness of the intervention for presenting problems, how time consuming the intervention was, and whether the teacher thought the intervention adversely affected the students involved (Witt & Marten, 1983). Teachers individually completed the IRP-20 upon completion of the follow-up phase.

The instrument included 20 items and a Likert type scale, ranging from 1 (Strongly Disagree) to 6 (Strongly Agree). The highest score that could be obtained on this scale was 120 (20 items * 6 points). A score of 100 or higher was considered acceptable (Witt & Marten, 1983) meaning the teacher found the intervention to be effective in decreasing disruptive behavior and he/she liked having the intervention in his/her classroom. The higher the teacher score, the greater the acceptability of the intervention (Finn & Sladeczek, 2001).

A study by Finn and Sladeczek (2001) determined the IRP-20 is a reliable and valid scale. Past studies in which teachers had completed the IRP-20 were noted, and it was found that the IRP-20 yielded internal consistency reliabilities ranging from .85 to .89 (Finn & Sladeczek, 2001). Criterion validity was also established with the Semantic
Differential (SD) scale, which measures concepts similar to those measured on the IRP-20. The IRP-20 produced a correlation of .86 with the Semantic Differential scale. The IRP-20 was also shown to discriminate between varying interventions (Finn & Sladeczek, 2001). Overall, the literature on the IRP-20 indicates it is a useful tool for assessing the acceptability of behavioral interventions in educational settings (Finn & Sladeczek, 2001). However, results also indicate the acceptability of an intervention can vary as a function of the severity of the problem behaviors, with higher ratings when behavior problems are more difficult (Finn & Sladeczek, 2001). Also, higher acceptance scores tend to be given when the time involvement in the intervention for teachers is low (Finn & Sladeczek, 2001). On this version of the IRP-20 internal consistency reliability was assessed through Cronbach’s resulting in excellent internal consistency (Alpha, á = 0.94).

**Children’s Intervention Rating Profile.** The Children’s Intervention Rating Profile (CIRP) was developed using items from the Intervention Rating Profile (Witt & Martens, 1983). It is the only rating scale specifically designed to assess treatment acceptability of children (Finn & Sladeczek, 2001). The CIRP consists of seven statements written at a fifth-grade reading level (Finn & Sladeczek, 2001). Items relate to the fairness, expected effectiveness, and the potential negative consequences associated with an intervention (Finn & Sladeczek, 2001). Students are asked to rate the intervention on a scale of 1 (I agree very much) to 5 (I disagree very much). A score of 35 is considered to be the most acceptable score. A modified version of the CIRP was individually administered to children in kindergarten and first grade. Third and fourth
grade participants completed another modified version of the CIRP as a class, with items read aloud by the researcher.

Studies utilizing the CIRP have found internal consistency coefficients range between .75 and .89 with elementary school students (Witt & Elliott, 1985), however this tool has been found to have internal consistency in studies with first grade to college aged students (Finn & Sladeczek, 2001). Studies have shown children are able to indicate their preferences on the CIRP, which lends to the validity of the measure with school-aged students (Elliott, 1988; Turco & Elliott, 1986).

Because the original CIRP is written at a fifth-grade reading level, it is possible its use may be limited to older children (Finn & Sladeczek, 2001). Nonetheless, studies have shown the CIRP is a reliable measure with younger children. In a study completed by Waas and Anderson (1991), the CIRP was given to a group of second graders. To ensure students understood the task, participants worked together in groups of five and were overseen by a research assistant (Finn & Sladeczek, 2001). In a study by Galloway and Sheridan (1994), children in first to third grades were asked to evaluate an intervention. To ensure students understood the items and the task, the CIRP was read to the students. Similar modifications were used with the students who participated in this study.

Two versions of the Children’s Intervention Rating Profile (CIRP) (Witt & Elliott, 1985) modified by the researcher were used to assess student intervention satisfaction. They were both modified to include language referring to the Mystery Motivator Calendar. One version with simplified language and a smiley face rating scale was used for kindergarten and first grade participants. Different smiley faces corresponded with
different number scores, with the face with the biggest smile being the equivalent of a 5 or “I agree very much,” the next face with the smaller smile was the equivalent of a 4 or “I sort of agree,” the neutral face was the equivalent of a 3 or “I neither agree nor disagree,” the face with a small frown corresponded to a 2 or “I sort of disagree,” and finally the face with the biggest frown corresponded to a 1 or “I disagree very much.” On this scale, items 2, 3 and 4 were reversed scored so a higher score would indicate a higher level of intervention satisfaction. (Please refer to Appendix I to see the kindergarten and first grade modified CIRP.) The scale was completed individually with kindergarten and first grade students to help ensure they understood the task. It was emphasized that honestly reporting how the students felt about the intervention was the most important factor to consider when completing the scale. On the kindergarten and first grade version of the CIRP, internal consistency reliability was assessed through Cronbach’s, resulting in questionable internal consistency (Alpha, \( \alpha = .60 \)).

For third and fourth grade participants, a numerical rating scale was used and language was not simplified and more closely mirrored the items on the original CIRP. Students were asked to rate the intervention on a scale of 1 (I agree very much) to 5 (I disagree very much). Please note the values corresponding to satisfaction level are in an inverse order compared to the scale used for kindergarten and first grade students. This design more closely reflects the original CIRP design. On this scale items 1, 5, 6 and 7 were reversed scored so a higher score would indicate a higher level of intervention satisfaction. Please refer to Appendix I to see the third and fourth grade participant modified CIRP form. For third and fourth grade participants, the items were read to the
whole class to provide students with the opportunity to ask for clarifications and to help ensure they understood the statements and the task. Once again, it was stressed that honestly reporting how the students felt about the intervention was the most important factor to consider when completing the scales. On this version of the CIRP, internal consistency reliability was assessed with Cronbach’s indicating poor internal consistency (Alpha, á = 0.55). When internal consistency reliability was assessed together for both modified versions of the CIRP, Cronbach’s indicated questionable consistency (Alpha, á = 0.65).

**Treatment Integrity**

To help ensure treatment integrity, an intervention protocol was developed. The protocol outlined all intervention components (see Appendix D). The protocol included: a sample lesson plan for teaching appropriate replacement behaviors, an announcement that the intervention was in place, a script describing the intervention, review of the behavioral goals, the procedure for recording tallies, the procedure for determining whether criterion for reinforcement was met, the procedure for choosing a reinforcer from the Mystery Motivator envelope, and the procedure to announce the criterion was not met. The researcher also observed 28% of the intervention sessions to help ensure the intervention was being implemented with fidelity. During these observations teachers implemented the intervention with 100% integrity.

**Intervention Procedures**

**Teacher interview.** Once consent was gathered from all participants, a preliminary teacher interview took place. Classroom demographic data was collected.
Disruptive classroom behaviors were identified through a modified version of the Problem Identification Interview (PII) (Kratochwill & Bergan, 1990). The collaborative problem solving process outlined in the PII aligns with PBS practices (Office of Special Education Programs Technical Assistance Center for Positive Behavioral Interventions & Supports, 2005). During this interview, the researcher and the teacher met to discuss the scope and the strength of the problem behavior. Next, target disruptive behaviors were chosen for intervention. The selected behaviors were defined and operationalized, and replacement behaviors were identified and operationalized. In addition, the time period during which the intervention would be implemented was chosen. Teachers were advised to pick a class/subject during which disruptive behavior was most problematic. Also, rewards used for mystery prizes, the time for gathering baseline data, and the intervention-training meetings, were determined in conjunction with the teacher during this initial meeting.

**Baseline.** The next step involved conducting observations in each classroom to tally the frequency of the targeted and operationalized disruptive behaviors to establish a baseline. Data were collected until the baseline stabilized, which involved between three and seven observation sessions depending on the variability of the behavior. Based upon the observed number of occurrences of the operationalized disruptive behavior, a criterion was set in collaboration with the teacher. The researcher advised the initial criterion be set at 50% of the number of disruptive behavioral events observed during baseline. In all but one classroom, Classroom H, teachers followed this advice. In
Classroom H, slightly more stringent criteria were established for phase one and phase two of the intervention.

**Teacher intervention training meeting.** After a baseline was established, a meeting was held with the teacher to train her on the intervention in accordance with Intervention Protocol. Teachers were provided with copies of the intervention protocol (see Appendix D). The importance of fidelity to the intervention procedures was discussed in this meeting. Each teacher was also provided with the opportunity to ask questions and voice concerns prior to implementation.

**Teaching behavioral goals.** During the initial phase of intervention implementation, students were taught behavioral goals in a lesson led by the researcher. The behavioral goals were the replacement behaviors identified in the problem identification interview. During this lesson, the class spent time discussing each goal. The rationale behind the goals and the importance of a classroom environment that encourages learning was also discussed. Students were asked to model and practice the behavioral goals. They were asked to provide positive and negative examples of the behavior. Feedback was to students on their demonstration and examples of replacement behavior.

**Intervention announcement and explanation.** On the first day of implementation, the intervention was explained in detail to the students. The explanation of the intervention followed the Intervention Protocol. The steps included: an announcement that the intervention was in place, a description of the intervention, a review of the behavioral goals discussed, an explanation of the procedure for counting
tallies on the tally counter and for making tally marks on the calendar, an explanation of the procedure for determining whether criterion for reinforcement was met, and an explanation of the procedure for drawing a reinforcer from the Mystery Motivator envelope. Verbal student assent was also obtained during this time. Students were also given the opportunity to seek clarification and ask questions.

**Visual reminders of behavioral expectations.** The behavioral goals were prominently displayed on the Mystery Motivator Calendars in each classroom to serve as a reminder for the students. In addition, each student received an individual handout so he/she would have a copy of the behavioral goals. These handouts were provided prior to intervention implementation (see Appendix H).

**Implementation.** The Mystery Motivator Calendar intervention was implemented in each classroom for approximately eight weeks. This roughly corresponds to the length of one marking period. The intervention took place during the time-period/class-period identified as most problematic during the PII. When a student engaged in an operationalized disruptive behavior during this class period, the teacher would count a tally on the tally counter, because the tally counter made a clicking sound, this created an environment in which immediate feedback was provided. The tally counter also provided students with a discriminative stimulus. At the earliest possible convenient time, the teacher would go up to the calendar and mark the number of tallies the students had received. Unlike, Murphy et al. (2007), this procedure greatly reduced the interference of data collection requirements on the instructional flow in classrooms. In Murphy et al., the teacher had to stop instruction in order to mark checks next to the names of the students.
engaged in disruptive behaviors. This procedure was identified as one of the study’s limitations. In this study, teachers did not have to stop instruction to make a tally, nor did they have to stop instruction to write down a tally mark on a sheet, instead they simply pressed the tally counter they held in their hand.

The Mystery Motivator Calendar was displayed for all students to see throughout the intervention period. Jones and Van Houte (1985) have shown public posting of an intervention can reduce disruptive behaviors and improve academic engagement. At the end of the period, the square covering that particular day on the calendar was removed, regardless of whether or not the students were able to achieve the set criterion for the day, so they could see whether an “M” was marked. If an “M” was marked, and the students met the established criterion, the prize was awarded at the earliest convenience to create an environment in which immediate reinforcement was provided. The processes for making tallies on the calendar, determining whether criterion for reinforcement was met, drawing a reinforcer from the Mystery Motivator envelope, and announcing when the criterion was not met, followed the Intervention Protocol.

Once students were able to meet the initial established criterion for a period of ten total school days, a new criterion was established. The researcher recommended that the new criterion be determined by once again reducing the number of allowed disruptive behavioral incidents by 50%. This procedure was followed until the teacher felt that the disruptive behavior had been reduced to a manageable number. If the number that was reduced by 50% was an odd number, then the new criterion would be rounded up to the nearest higher whole number.
There were a few departures from the intervention procedure that took place during implementation, however, they were minor and did not seriously undermine the fidelity of the intervention. First, in Classroom H, the initial criterion for being eligible for reinforcement was established as an approximate 67% reduction of disruptive behavioral events as opposed to an approximate 50% reduction as in all the other classes. Also, in Classroom H, the reduction from the first to second phase of intervention was greater than 50%. A 60% reduction in the allowed number of behavioral events from phase one to phase two of the intervention was made. For specific criteria established on a classroom-by-classroom basis please refer to the results section below.

Also, in six of the eight classrooms, teachers did not make timely reductions in criteria in accordance with the outlined procedures. Teachers should have decreased the allowed criteria by 50% after students were able to meet the prior criteria for 10 school days in total. Instead, one teacher changed the criterion a day too early. In this class students achieved the criterion for nine out of the first ten school days the intervention was in place. Thus, ten school days had passed, but not ten school days in total during which students were able to achieve the criterion. Also, another three teachers changed the criterion one, two or three days after students met the criterion for the required period of ten school days in total. The teachers reported to have forgotten to alter the criterion. Also, in one classroom, the teacher did not move on to the third criterion and third phase of the intervention, by reducing the allowed number disruptive behavioral events again, like she should have.
**Intervention materials.** To implement this intervention, the following materials were used: (a) Mystery Motivator Calendars; (b) a Mystery Motivator envelope which contained note cards with descriptions of individual mystery prizes; (c) a visual display of the expected behaviors stated in positive terms; (d) tangible rewards, such as candy, pencils, easers, etc., that corresponded to the reinforcers listed on the note cards; and e) tally counters to ensure accurate data collection by teachers. (Please see Appendix E for a visual of the Mystery Motivator Calendar. Please see Appendix F for a visual of the tally counter.)

**Reversal.** A two-day reversal was introduced in each classroom to further lend credibility to the intervention (Kazdin, 2011). This always took place after the students had been able to meet the established criterion for a period of at least nine school days in total. During this time period, the intervention was not implemented. The teacher no longer used the tally counter. The researcher came into the classrooms to tally the occurrence of disruptive behavior on her computer. A complete reversal to baseline was not expected because it was hypothesized that students would learn replacement behaviors similar to Murphy et al. (2007). Students were not expected to unlearn these behaviors during reversal; however, it was hypothesized that their motivation to meet behavioral expectations would decrease when reinforcement was removed.

**Reinstatement and follow-up.** A reinstatement of the intervention followed the reversal period. After the intervention had been implemented for a period of approximately eight school weeks in total, it was removed again for follow-up data. Once again, a return to baseline conditions was not expected during follow-up, but it was
hypothesized that problem behavior would slightly increase due to the removal of reinforcement.

**Social validity questionnaires.** At the end of the eight-week intervention period, the teachers were asked out to fill out a modified Intervention Rating Profile-20 to gauge their acceptance of the intervention (see Appendix J). Two modified versions of the Children’s Intervention Rating Profile were given to the students. One was modified for kindergarten and first grade participants. The other version was modified for third and fourth grade participants. Both modified versions asked the students to evaluate their satisfaction with the intervention. Individual student data was not gathered (see Appendix I).

**Inter-observer reliability.** Inter-observer reliability was calculated as a frequency ratio (Kazdin, 2011) between the teacher and the researcher for a total of 28% of the intervention sessions. Average inter-observer agreement for problem behavior across all classrooms was 92% (range of 70% to 100%). When observing in classrooms during Mystery Motivator Calendar intervention sessions, the researcher recorded incidents of operationalized disruptive behaviors on her computer. At the same time, the teacher recorded incidents of operationalized disruptive behaviors on the tally counter and on the calendar. The researcher then compared her recorded tallies to the teacher’s recorded tallies in each classroom. The percentage of agreement was used as indicator of reliability. This was calculated by dividing the lower number of tallied disruptive behavioral events by the higher number of tallied disruptive behavioral events and
multiplying by 100. For detailed information regarding inter-observer reliability data for individual classrooms, please refer to Table 4 below.

Table 4. Inter-Observer Agreement per Classroom

<table>
<thead>
<tr>
<th>Classroom</th>
<th>Percentage of Sessions Observed</th>
<th>Percentage Agreement</th>
<th>Range of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>25%</td>
<td>91%</td>
<td>74 to 100%</td>
</tr>
<tr>
<td>B</td>
<td>24%</td>
<td>92%</td>
<td>81 to 98%</td>
</tr>
<tr>
<td>C</td>
<td>25%</td>
<td>92%</td>
<td>84 to 100%</td>
</tr>
<tr>
<td>D</td>
<td>33%</td>
<td>90%</td>
<td>70 to 100%</td>
</tr>
<tr>
<td>E</td>
<td>25%</td>
<td>92%</td>
<td>85 to 100%</td>
</tr>
<tr>
<td>F</td>
<td>30%</td>
<td>90%</td>
<td>80 to 100%</td>
</tr>
<tr>
<td>G</td>
<td>31%</td>
<td>93%</td>
<td>87 to 100%</td>
</tr>
<tr>
<td>H</td>
<td>29%</td>
<td>91%</td>
<td>77 to 98%</td>
</tr>
<tr>
<td>Overall</td>
<td>28%</td>
<td>92%</td>
<td>70 to 100%</td>
</tr>
</tbody>
</table>

Data Analysis

Mystery Motivator Calendar Data

In order to ensure for comparability of data within classrooms, when intervention periods were of varying lengths, raw data of frequency tallies were converted to rates. The rate reported was the mode time the intervention was implemented in each given data set. The primary method of data evaluation in this study was visual inspection. Visual inspection of the data is the primary method of data-evaluations in single-case designs (Kazdin, 2011). The effects of the intervention at different points in time and between
subjects were analyzed according to criteria laid out by Kazdin (p. 288). First, characteristics related to the magnitude of the change were examined. This included investigating whether there had been a change in mean across phases and a change in level across phases. Characteristics related to the rate of change were also inspected. This included a change in trend or slope and the latency of the change. Finally, the overall pattern of the data within and between classrooms was examined, including the pattern of data during reversal and follow-up. This included any variability in the pattern of the data. Means of the number of occurrence of operationalized behavior were compared between baseline, intervention, reversal, reinstatement, and follow up. Mean standard deviations were also calculated. The number of days and percentage with which the students were able to meet the criterion in each intervention phase were also examined.

**Intervention Rating Profile**

The Intervention Rating Profile (IRP-20) was scored by adding the numbers corresponding to teachers’ responses on the IRP-20 form. Mean responses to individual items were also examined. Responses on the IRP-20 form were captured on a scale ranging from 1 to 6. The number 1 corresponded to ‘Strongly Disagree’, while 6 corresponded to ‘Strongly Agree’. A score of 100 or higher was considered to be acceptable and the higher the score, the greater the intervention acceptability (Finn & Sladeczek, 2001) (see Appendix J).

**Children’s Intervention Rating Profile**

Two modified versions of the Children’s Intervention Rating Profile (CIRP) (Witt & Elliott, 1985) were used to assess student intervention satisfaction. One version was
created for kindergarten and first grade participants. On this item, language was simplified and a smiley face rating scales was used. The biggest smile corresponded to “I agree very much” and the biggest frown corresponded to “I disagree very much”. Negatively stated items were reverse-scored.

Another version was created for third and fourth grade participants, which included the number scale used on the original CIRP. The CIRP consisted of seven statements that asked students to rate the intervention on a scale from 5 (I agree very much) to 1 (I disagree very much). Negatively stated items were reverse-scored. It is important to mention that scales ran in inverse order on the two modified versions developed.

The CIRP was scored by calculating the average number of points assigned to the intervention on a classroom basis. A classroom average score of 35 is the most acceptable score, indicating the classroom “agrees very much” the intervention was acceptable. A classroom average score of 28 indicates the students in the classroom, on average, “sort of agree” the intervention was acceptable. Mean score responses to individual items were also examined. (For copies of both versions of the CIRP used, please refer to Appendix I.)
CHAPTER FOUR

RESULTS

Presentation of Results

This study investigated whether the Mystery Motivator Calendar intervention, which utilized a group contingency, variable-ratio reinforcement schedule, would decrease disruptive classroom behavior. It was hypothesized that the results would replicate similar studies in which Mystery Motivator was found to be effective in modifying behavior. The study also examined whether the effectiveness of the Mystery Motivator Calendar intervention would decrease over the course of the study. It was hypothesized that intervention effects would decrease slightly as students became satiated with the intervention and it became less of a novelty. This study also examined teacher intervention acceptability ratings for the Mystery Motivator Calendar. It was hypothesized that teachers participating in the intervention would rate the intervention as acceptable given the results of previous studies. Finally, this study examined student satisfaction with the Mystery Motivator Calendar. It was hypothesized that students would positively rate the intervention, given the results of prior studies. The aforementioned hypotheses were tested. Results are based on a final sample of eight classroom participants. Results will be presented by research question followed by an overall summary of results (see Table 13).
Research Question One

How will the interdependent group contingency Mystery Motivator Calendar intervention affect the frequency of disruptive behavioral events in general education elementary school classrooms as identified by each teacher? Results for the first research question are presented below on a classroom-by-classroom basis. It was hypothesized that results would replicate similar studies in which Mystery Motivator was found to be effective in modifying behavior. Data collected in each classroom is presented in the form of a graph to allow for visual inspection of the effects of the intervention at different points in time. Also, the means of the number of operationalized disruptive behavioral events are compared between baseline, intervention, reversal, reinstatement, and follow up. Mean standard deviations are calculated and finally the number of days and percentage which with students were able to meet the criterion in each intervention phase are examined.

Classroom A

In Classroom A, which was a kindergarten classroom, it was determined the intervention would be implemented during the morning meeting. During this time, students engaged in activities related to the calendar, the weather, counting, prayer, the Pledge of Allegiance, songs, listening to the teacher read a book, and the introduction of the activities that would follow. Students were assigned jobs during the morning meeting so they could actively participate. The time period for the intervention was from 9:00am until 9:45am each day. The identified problem behaviors included calling out without raising one’s hand, talking, and not sitting
appropriately on the rug. The replacement behaviors identified included: always raising your hand, answering as a class when cued by the teacher, and sitting on the floor in “Criss Cross Applesauce” position, legs crossed and hands kept to yourself.

In Classroom A, data were collected for four points during baseline. The mean number of disruptive behavioral events during baseline was 119. Based upon this number, an initial criterion for mystery prize eligibility of 60 disruptive behavioral events or less was established in conjunction with the teacher for the first phase of the intervention, an approximate 50% reduction from baseline. The first phase of the intervention lasted 10 school days. Students were able to meet this criterion 100% of the time. During the second phase of the intervention, a new criterion of 30 or less disruptive behavioral events was established, again reducing the number of allowed disruptive behavioral events by 50%. This phase of the intervention lasted 20 school days in total. Students were able to achieve the criterion for 13 of the 20 school days, or 65% of the time. Finally, in the third phase of the intervention, a final criterion of 15 or less disruptive behavioral events was introduced. This phase lasted six school days in total. Students were able to achieve this criterion for one of the six school days, or 17% percent of the time.

Data for disruptive behavior are presented in Figure 1 and Table 5. A moderate amount of variability was noted during baseline (SD = 22.38). Upon introduction of the first phase of intervention (criterion is equal to or less than 60 tallies), an immediate change in level was observed with no latency noted. The mean of disruptive behavioral events decreased 65.21% from baseline (M = 119) to the first phase of intervention (M =
41.4). Additionally, the amount of variability in the data decreased from baseline (SD = 22.38) to the first phase of the intervention (SD = 12.15). The mean number of disruptive behavioral events decreased again in the second phase of the intervention (M = 28.65) by 40.64%.

Figure 1. Disruptive Behavior Classroom A

A reversal was introduced in the middle of the second intervention phase. An increase in the mean number of disruptive behavioral events was observed during reversal (M = 41.5), as compared to the mean of the second phase of the intervention (M = 28.65). This change did not reach baseline levels, although, the mean number of disruptive behavioral events during reversal (M = 41.5) was similar to the mean number of disruptive behavioral events observed during the first phase of the intervention (M = 41.4). Also, in the second phase of the intervention (criterion is equal to or less than 30
tallies) less variability (SD = 7.69) was observed as compared to the first phase of intervention (SD = 12.15), most notably after the introduction of the reversal. The highest number of disruptive behavioral events observed during reversal was 51, compared to the 99, the lowest number of disruptive events observed during baseline. Furthermore, the mean number of disruptive behavioral events observed during baseline was 119.

Table 5. Classroom A Intervention Phase Data

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Intervention Phase 2</th>
<th>Reversal</th>
<th>Intervention Phase 3</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Days</td>
<td>4</td>
<td>10</td>
<td>20</td>
<td>2</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Criterion Days Met</td>
<td>60</td>
<td>30</td>
<td>15</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Met</td>
<td>100%</td>
<td>65%</td>
<td>17%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>119</td>
<td>41.4</td>
<td>28.65</td>
<td>41.5</td>
<td>22.33</td>
<td>30</td>
</tr>
<tr>
<td>SD</td>
<td>22.38</td>
<td>12.15</td>
<td>7.69</td>
<td>13.44</td>
<td>7.633</td>
<td>9.9</td>
</tr>
</tbody>
</table>

Upon reinstatement, a decrease in disruptive behavioral events was again noted. During the third phase of the intervention (criterion is equal to or less than 15 tallies), a further decrease in the mean number of disruptive behavioral events (M = 22.33) as compared to the second phase of intervention (M = 28.65) was observed.

During follow-up a slight increase in the mean number disruptive behavioral events was noted (M = 30) as compared to the last phase of the intervention (M=22.33). Overall, an 81.24% decrease of disruptive behavioral events was noted from baseline to the third phase of intervention. A 74.79% decrease in disruptive behavioral events was noted from baseline to follow-up.
Classroom B

In Classroom B, which was also a kindergarten classroom, it was determined the intervention would also be implemented during the morning meeting. During this time, students engaged in activities related to the calendar, the weather, counting, prayer, the Pledge of Allegiance, songs, listening to the teacher read a book, and the introduction of the activities that would follow. Students were assigned jobs during the morning meeting so they could actively participate. The time period for the intervention was 40 minutes each day, from 8:45am to 9:25am. The identified problem behaviors included calling out without raising one’s hand, talking, and not sitting appropriately on the rug. The replacement behaviors identified included: always raising your hand, answering as a class when cued by the teacher, and sitting on the floor in “Criss Cross Applesauce” position, legs crossed and hands kept to yourself.

In Classroom B, data were collected for three points during baseline. Disruptive behavior across baseline was averaged at 133 per a 40-minute class period. Based upon this number, an initial criterion for mystery prize eligibility of 66 disruptive behavioral events or less was established, in conjunction with the teacher, an approximate 50% reduction from baseline. The first phase of the intervention lasted 12 school days. Students were able to meet this criterion for 10 of the 12 school days or 83% of the time. During the second phase of the intervention, a new criterion of 33 or less disruptive behavioral events was established, reducing the number of allowed disruptive behavioral events by 50%. This phase of the intervention lasted 13
school days in total. Students were able to achieve the criterion for 12 of the 13 school days, or 92% of the time. Finally, in the third phase of the intervention, a final criterion of 17 or less disruptive behavioral events was introduced. This phase lasted 12 school days in total. Students were able to achieve this criterion for 12 of the 12 school days, or 100% percent of the time.

Data for disruptive behavior are presented in Figure 2 and Table 6. A moderate amount of variability was noted during baseline (SD = 25.63). Upon introduction of the first phase of intervention (criterion is equal to or less than 66 tallies), an immediate change in level was observed with no latency noted. The mean of disruptive behavioral events decreased 65.94% from baseline (M = 133) to the first phase of intervention (M = 45.3). Additionally, the data became more stable in the first phase of the intervention (SD = 17.78) compared to baseline (SD 25.63).

The mean number of disruptive behavioral events decreased again by 52.12% in the second phase of the intervention (M = 21.69) as compared to the first phase of the intervention (M = 45.3). Also, in the second phase of the intervention (criterion is equal to or less than 33 tallies) less variability (SD = 6.54) in the data was observed in comparison to the first phase of intervention (SD = 17.78).

A reversal was introduced between the second and third intervention phases. An increase in the mean number of disruptive behavioral events (M=38.5) was observed, as compared to the second phase of the intervention (M= 21.69). However, the increase did not reach baseline conditions. The highest level of disruptive behavioral events observed during reversal was 42. This is significantly lower than 106, the lowest number of
disruptive events observed during baseline or 133 the mean number of disruptive
behavioral events observed during baseline.

Note: M stands for Mean. During the reversal there was a school day in between the two reversal days on which data was not collected.

Figure 2. Disruptive Behavior Classroom B

Table 6. Classroom B Intervention Phase Data

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Intervention Phase 2</th>
<th>Reversal</th>
<th>Intervention Phase 3</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Days</td>
<td>3</td>
<td>12</td>
<td>13</td>
<td>2</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Criterion Days Met</td>
<td>66</td>
<td>33</td>
<td></td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Met</td>
<td>83%</td>
<td>92%</td>
<td></td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>133</td>
<td>45.3</td>
<td>21.69</td>
<td>38.5</td>
<td>11.83</td>
<td>14</td>
</tr>
<tr>
<td>SD</td>
<td>25.63</td>
<td>17.78</td>
<td>6.54</td>
<td>4.95</td>
<td>2.89</td>
<td>0</td>
</tr>
</tbody>
</table>

Upon reinstatement an immediate decrease in disruptive behavior was noted again. During the third phase of the intervention (criterion is equal to or less than 17
tallies), a decrease in the mean number of disruptive behavioral events was observed (M = 11.83) as compared to the second phase of intervention (M = 38.5).

During follow-up a slight increase in the mean number of disruptive behavioral events was noted (M = 14) as compared to the third phase of the intervention (M = 11.83). Overall, a decrease of 91.11% disruptive behavioral events was noted from baseline to the third phase of intervention. A decrease of 89.47% disruptive behavioral events was noted from baseline to follow-up.

**Classroom C**

In Classroom C, which was a kindergarten classroom, it was determined the intervention would be implemented during the morning meeting. During this time, students engaged in activities related to the calendar, the weather, counting, prayer, the Pledge of Allegiance, songs, listening to the teacher read a book, and the introduction of the activities that would follow. Students were assigned jobs during the morning meeting so they could actively participate. The time period for the intervention was 30 minutes each day, from 8:15am to 8:45am. The identified problem behaviors included not sitting appropriately on the rug and looking at something other than the teacher during direct instruction. The replacement behaviors identified included: sitting on the floor in “Criss Cross Applesauce” position, legs crossed and hands kept to yourself, and keeping your eyes on your teacher during instruction.

In Classroom C, data were collected for three points during baseline. The mean number of disruptive behavioral events during baseline was 126. Based upon
this number, an initial criterion for mystery prize eligibility of 63 disruptive
behavioral events or less was established, in conjunction with the teacher, an
approximate 50% reduction from baseline. The first phase of the intervention lasted
15 school days, students were able to achieve the set criterion for 10 school days or
67% of the time. During the second phase of the intervention, a new criterion of 32 or
less disruptive behavioral events was established, reducing the number of allowed
disruptive behavioral events by approximately 50%. This phase of the intervention
lasted 16 school days in total. Students were able to achieve the criterion for 3 of the
16 school days, or 19% of the time.

Data for disruptive behavior are presented in Figure 3 and Table 7. A very small
amount of variability was noted during baseline (SD = 6). Upon introduction of the first
phase of intervention (criterion is equal to or less than 63 tallies), an immediate change in
level was observed with no latency noted. The mean of disruptive behavioral events
decreased 51.59% from baseline (M = 126) as compared to the first phase of intervention
(M = 61). Variability increased during the first phase of the intervention (SD = 18.64) as
compared to baseline (SD = 6). Also, an outlier was observed in the first phase of the
intervention.
The mean number of disruptive behavioral events decreased again from the first phase of the intervention (M = 61) to the second phase of the intervention (M = 45.87; criterion is equal to or less than 60 tallies) by 24.8%. An outlier was observed again during the second phase of the intervention. Also, in the second phase of the intervention (criterion is equal to or less than 32 tallies) a slight decrease in variability (SD = 15.85)
was observed.

A reversal was introduced in the middle of the second intervention phase. An increase in the mean number of disruptive behavioral events ($M = 73$) was observed, as compared to the first phase of the intervention ($M = 62$), but this change did not reach baseline levels. The highest level of disruptive behavioral events observed during reversal was 74 as compared to the 120, the lowest number of disruptive events observed during baseline, or 126 the mean number of disruptive behavioral events observed during baseline.

Upon reinstatement an immediate decrease in the number of disruptive behavioral events was again noted. During follow-up an even further decrease in disruptive behavioral events was noted ($M = 38.5$). Overall, a 63.6% decrease of disruptive behavioral events was noted from baseline to the second phase of intervention. A decrease of 69.44% disruptive behavioral events was noted from baseline to follow-up.

**Classroom D**

In Classroom D, which was a first grade classroom, it was determined the intervention would be implemented during mathematics. The time period for the intervention was an hour each day, from 9:10am until 10:10am. On Mondays, math was only 30 minutes, so the reported rate of tallies has been pro-rated to ensure comparability in the data. The first portion of math class involved sitting on the carpet for direct instruction, a smart board was used to guide instruction. Next, the class would complete problems as a group with the use of the smart board. This was followed by the completion of independent seatwork consisting of work sheets. Once
students finished independent work, they were allowed to play a math game in a corner of the room, away from students who were still working. The identified problem behaviors included calling out without raising one’s hand, talking, and not sitting appropriately on the rug. The replacement behaviors identified included: always raising your hand, answering as a class when cued by the teacher, and sitting on the floor in “Criss Cross Applesauce” position, legs crossed and hands kept to yourself.

In Classroom D, data were collected for four points during baseline. The mean number of disruptive behavioral events during baseline was 92.5. Based upon this number, an initial criterion for mystery prize eligibility of 46 disruptive behavioral events or less was established, an approximate 50% reduction from baseline. The first phase of the intervention lasted 10 school days. Students were able to meet this criterion 100% of the time. During the second phase of the intervention, a new criterion of 23 or less disruptive behavioral events was established, reducing the number of allowed disruptive behavioral events by 50%. This phase of the intervention lasted 10 school days in total, students were able to achieve the criterion for 10 of the 10 school days, or 100% of the time. Finally, in the third phase of the intervention, a final criterion of 13 or less disruptive behavioral events was introduced. This phase lasted 11 school days in total. Students were able to achieve this criterion for 10 of the 11 school days, or 91% percent of the time.

Data for disruptive behavior are presented in Figure 4 and Table 8. A moderately low amount of variability was noted during baseline (SD = 16.5). Upon introduction of
the first phase of intervention (criterion is equal to or less than 46 tallies), an immediate change in level was observed with no latency noted. The mean of disruptive behavioral events decreased 76.22% from baseline (M = 92.5) to the first phase of intervention (M = 22). Additionally, variability in the data decreased during the first phase of the intervention (SD = 10.26) as compared to baseline (SD = 16.5).

The mean number of disruptive behavioral events decreased again from the first phase of the intervention (M = 22) to the second phase of the intervention (M = 10.8) by 50.91%. Also, in the second phase of the intervention (criterion is equal to or less than 23 tallies) less variability (SD = 3.29) was observed as compared to the first phase of intervention (SD = 10.26).

Note: C3 stands for Criterion 3; the third intervention phase is interrupted by a reversal.

Figure 4. Disruptive Behavior Classroom D
Table 8. Classroom D Intervention Phase Data

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Intervention Phase 2</th>
<th>Reversal</th>
<th>Intervention Phase 3</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Days</td>
<td>4</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Criterion Days Met</td>
<td>46</td>
<td>23</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days Met</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Met</td>
<td>100%</td>
<td>100%</td>
<td>91%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>92.5</td>
<td>22</td>
<td>10.8</td>
<td>22.5</td>
<td>8.09</td>
<td>14</td>
</tr>
<tr>
<td>SD</td>
<td>16.5</td>
<td>10.26</td>
<td>3.29</td>
<td>3.54</td>
<td>4.25</td>
<td>2.83</td>
</tr>
</tbody>
</table>

Next, the third phase of the intervention was introduced (criterion is equal to or less than 12 tallies). A couple of days into the third phase of the intervention a reversal was introduced. An increase in the mean number of disruptive behavioral events (M = 22.5) was observed, as compared to the second phase of the intervention (M = 10.8), but this change did not reach baseline levels. The mean number of disruptive behavioral events during reversal (M = 22.5) was almost the same as the mean number of disruptive behavioral events observed during the first phase of the intervention (M = 22). The highest number of disruptive behavioral events observed during reversal was 25 as compared to the 73, the lowest number of disruptive events observed during baseline, or 92.5 the mean number of disruptive behavioral events observed during baseline.

Upon reinstatement, there was one day of latency before the number of disruptive behavioral events decreased again. During the third phase of the intervention (criterion is equal to or less than 12 tallies), a decrease in the mean number (M = 8.09) of disruptive behavioral events was observed as compared to the second phase of intervention (M = 10.08).

During follow-up a slight increase in the mean number of disruptive behavioral
events (M = 14) was noted as compared to the mean number of disruptive behavioral events observed during the third phase of the intervention (M = 8.09). Overall, a 92.5% decrease in the mean number of disruptive behavioral events was noted from baseline to the third phase of intervention. An 84.86% decrease in the mean number of disruptive behavioral events was noted from baseline to follow-up.

**Classroom E**

In Classroom E, which was a first grade classroom, it was determined the intervention would be implemented during mathematics class in the afternoon. The time period for the intervention was estimated to be 50 minutes during the Problem Identification Interview, although there were days on which the math lesson was a little shorter or little bit longer. The first part of the math lesson involved direct instruction with a smart board followed by individual or small group work on math worksheets at the student’s desks. The desks were arranged in clusters. Next, the class would come together and review the worksheets.

The identified problem behaviors included calling out without raising one’s hand or talking during instruction; getting out of one’s seat during instruction for any reason other than getting a pencil or a tissue; and opening and closing desks during direct instruction. The replacement behaviors included: always raising your hand to talk in class, staying quiet while the teacher is speaking, staying in your seat during instruction (unless the student needed a tissue or pencil) and finally keeping desks closed during instruction.

Because the duration of math class varied slightly from day to day the teacher was
asked to record the start and stop times of each class. After data were collected, it was found that the average length of math class was 48 minutes, with a range of 25 to 70 minutes. Despite this irregularity, the behavioral criteria for students to receive reinforcement, was not pro-rated, thus it did not adjust for the length of the class. To maintain consistency in reporting the dependent variable, the number of disruptive behavioral events occurring per class, were adjusted as a ratio to reflect a 50-minute time period.

In Classroom E, data were collected for five points during baseline. The mean number of disruptive behavioral events during baseline was 227. Based upon this number, an initial criterion for mystery prize eligibility of 114 disruptive behavioral events or less was established, in conjunction with the teacher, for the first phase of the intervention, an approximate 50% reduction from baseline. The first phase of the intervention lasted 11 school days. Students were able to meet the criterion for 10 out of the 11 school days or 91% of the time. During the second phase of the intervention, a new criterion of 57 or less disruptive behavioral events was established, reducing the number of allowed disruptive behavioral events by 50%. The second phase of the intervention lasted 11 school days. Students were able to meet the criterion for 10 out of the 11 school days or 91% of the time. Finally, in the third phase of the intervention, a final criterion of 29 or less disruptive behavioral events was introduced. This phase lasted 13 school days in total. Students were able to achieve the new criterion for 4 of the 13 school days, or 31% of the time.
Data for disruptive behavior are presented in Figure 5 and Table 9. A moderately large degree of variability was noted during baseline (SD = 42.3). Upon introduction of the first phase of intervention (criterion is equal to or less than 114 tallies), an immediate change in level was observed with no latency noted. The mean number of disruptive behavioral events decreased by 72.03% from baseline (M = 227.2) to phase one of the intervention (M = 63.54). A moderate amount of variability was observed during the first phase of the intervention (SD = 30.37).

Note: M refers to the mean of disruptive behavioral events during baseline and reversal.

Figure 5. Disruptive Behavior Classroom E
Table 9. Classroom E Intervention Phase Data

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Intervention Phase 2</th>
<th>Reversal</th>
<th>Intervention Phase 3</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Days</td>
<td>5</td>
<td>11</td>
<td>11</td>
<td>2</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>Criterion Days Met</td>
<td>114</td>
<td>10</td>
<td>57</td>
<td>29</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>% Met</td>
<td>91%</td>
<td>91%</td>
<td>91%</td>
<td>31%</td>
<td>31%</td>
<td>31%</td>
</tr>
<tr>
<td>Mean</td>
<td>227.2</td>
<td>63.54</td>
<td>37.72</td>
<td>83</td>
<td>41.54</td>
<td>81.5</td>
</tr>
<tr>
<td>SD</td>
<td>42.3</td>
<td>30.37</td>
<td>14.45</td>
<td>25.45</td>
<td>13.13</td>
<td>2.12</td>
</tr>
</tbody>
</table>

The mean number of disruptive behavioral events decreased again from the first phase of the intervention (M = 63.54) to the second phase of the intervention (M = 37.72) by 40.64%. Also, in the second phase of the intervention (criterion is equal to or less than 57 tallies) less variability (SD =14.45) was observed as compared to the first phase of intervention (SD = 30.37).

During the third phase of the intervention (criterion is equal to or less than 29 tallies), a slight increase of 9.2% in the mean number of disruptive behavioral events (M = 41.54) was observed as compared to the second phase of intervention (M = 37.72). Data stabilized further in the third phase of the intervention (SD = 13.13).

A reversal was introduced in the middle of the third intervention phase. An increase in the mean number of disruptive behavioral events (M = 83) was observed during reversal, but this change did not reach baseline levels. The highest level of disruptive behavioral events observed during reversal was 124 as compared to the 222, the lowest number of disruptive events observed during baseline, or 227.2 the mean number of disruptive behavioral events observed during baseline.
After the reversal, the third phase of intervention was implemented again. Upon reinstatement an immediate decrease in the number of disruptive behavioral events was noted again. During follow-up an increase in disruptive behavioral events was noted (M = 81.5), but once again it did not return to baseline conditions. Overall, an 81.72% decrease of disruptive behavioral events was noted from baseline to the third phase of intervention. A 64.11% decrease of disruptive behavioral events was noted from baseline to follow-up.

Classroom F

In Classroom F, which was a 3rd grade classroom, it was determined the intervention would be implemented during writing. The time period for the intervention was 25 minutes, from 9:45am until 10:10am each day. Writing class involved anywhere from two to ten minutes of direct teacher instruction followed by individual writing at the students’ desks. The identified problem behaviors included: not having materials ready for writing when asked, talking, and looking at something other than the teacher or the writing assignment. Identified replacement behaviors included: having materials ready, writing during independent work time, keeping your eyes on the teacher or the assignment, and talking to classmates only in order to proofread papers.

In Classroom F, data were collected for six points during baseline. The mean number of disruptive behavioral events during baseline was 39.66. Based upon this number, an initial criterion for mystery prize eligibility of 20 disruptive behavioral events or less was established, in conjunction with the teacher, an approximate 50% reduction from baseline. The first phase of the intervention lasted 10 school days. Students were
able to meet the criterion for 9 out of the 10 school days or 90% of the time. During the second phase of the intervention, a new criterion of 10 or less disruptive behavioral events was established, reducing the number of allowed disruptive behavioral events by 50%. The second phase of the intervention lasted 12 school days. Students were able to meet the criterion for 10 out of the 12 school days or 83% of the time. Finally, in the third phase of the intervention, a final criterion of 5 or less disruptive behavioral events was introduced. This phase lasted 11 school days in total. Students were able to achieve the new criterion for 11 of the 11 school days, or 100% of the time.

Data for disruptive behavior are presented in Figure 6 and Table 10. Data were not collected on a day-to-day basis as evidenced by the breaks in the line in the graph found in Figure 6 above. Variability was low during baseline (SD = 5.72). Upon introduction of the intervention in the first phase (criterion is equal to or less than 20 tallies), an immediate change in level was noted without latency.

The mean number of disruptive events decreased by 81.59% in the first phase of the intervention (M = 7.3) as compared to baseline (M = 39.66). Low variability (SD = 5.71) in the number of disruptive behavioral events was noted in the first phase of intervention.

A reversal was introduced between the first phase of the intervention (criterion is equal to or less than 20 tallies) and the second phase of the intervention (criterion is equal to or less than 10 tallies). During the reversal, an increase in the mean number of disruptive behavioral events (M = 24.5) was noted, but this change did not reach baseline levels (M = 39.66). The highest number of disruptive behavioral events during reversal
was 28, as compared to 32, the lowest number reached during baseline, or the mean of 39.66 observed during baseline.

Note: M refers to the mean of disruptive behavioral events during baseline and reversal.

Figure 6. Disruptive Behavior Classroom F

Table 10. Classroom F Intervention Phase Data

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Reversal</th>
<th>Intervention Phase 2</th>
<th>Intervention Phase 3</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Days</td>
<td>6</td>
<td>10</td>
<td>2</td>
<td>12</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Criterion Days Met</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>% Met</td>
<td>90%</td>
<td>90%</td>
<td>90%</td>
<td>90%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>39.66</td>
<td>7.3</td>
<td>24.5</td>
<td>7.67</td>
<td>3.63</td>
<td>15.5</td>
</tr>
<tr>
<td>SD</td>
<td>5.72</td>
<td>5.71</td>
<td>4.95</td>
<td>4.44</td>
<td>1.29</td>
<td>6.36</td>
</tr>
</tbody>
</table>

Upon reinstatement, the second intervention phase (criterion is equal to or less than 10 tallies) was introduced and an immediate decrease in the number of disruptive
behavioral events was again noted. Little variability was observed during the second phase of the intervention (SD = 4.44) and data became more stable than in the first phase of intervention. During the second phase of the intervention a very slight increase of 4.82% was noted in the mean number of disruptive behavioral events (M = 7.67) as compared to the first phase of intervention (M = 7.3).

During the third intervention phase (criterion is equal to or less than 5 tallies), a further decrease of 52.67% was noted in the mean number of disruptive behavioral events (M= 3.63) as compared to the second phase of the intervention (M = 7.67). There was very little variability in the data in this phase (SD = 1.29). During follow up, a small increase in disruptive behavioral events was noted again, but once again levels did not return to baseline conditions. Overall, a 90.85% decrease of disruptive behavioral events was noted from baseline to the third phase of intervention. A decrease of 60.92% disruptive behavioral events was noted from baseline to follow-up.

**Classroom G**

In Classroom G, which was a fourth grade classroom, it was determined the intervention would take place in the afternoon during social studies and science. Some days of the week the students would have science during this time period and on other days of the week they would have social studies. The schedule was not regular. The time period during which the intervention was implemented was from 1:30pm to 2:30pm on Mondays, Tuesdays, Thursdays, and Fridays and from 1:30pm to 2:00pm on Wednesdays. To maintain consistency in reporting the dependent variable, Wednesday data were adjusted as the rate of disruptive behavioral events occurring per a 60-minute
time period.

The identified problem behaviors included: shouting out without raising one’s hand or talking, getting out of one’s seat, and looking at something other than the book during independent reading. The replacement behaviors included: always raising your hand to talk in class, only answering as a class when cued by the teacher, staying in your seat during instruction and independent work, and keeping your eyes on your book during silent reading.

In Classroom G, data were collected for three points during baseline. The mean number of disruptive behavioral events during baseline was 246 per a one-hour class period. Based upon this number, an initial criterion of 123 disruptive behavioral events or less on Mondays, Tuesdays, Thursdays and Fridays and of 62 or less on Wednesdays was established, in conjunction with the teacher, a 50% reduction from baseline. The first phase of the intervention lasted 13 school days. Students were able to meet the criterion for 10 out of the 13 school days or 77% of the time. During the second phase of the intervention, a new criterion of 62 or less disruptive behavioral events on Mondays, Tuesdays, Thursdays and Fridays, and 31 or less on Wednesdays was established, reducing the number of allowed disruptive behavioral events again by approximately 50%. The second phase of the intervention lasted 19 school days. Students were able to meet the criterion for 11 out of the 19 school days or 58% of the time. In this classroom, a third phase was not implemented.

Data for disruptive behavior are presented in Figure 7 and Table 11. A moderate amount of variability was noted during baseline (SD = 44.5). Upon introduction of the
first phase of intervention (criterion is equal to or less than 123 tallies), an immediate change in level was observed with no latency noted. The mean of disruptive behavioral events decreased 59.65% from baseline ($M = 246.33$) to the first phase of intervention ($M = 109.15$). An outlier and moderate variability ($SD = 29.72$) were observed during phase one of the intervention, however, data stabilized in comparison to variability observed during baseline ($SD = 44.5$).

Figure 7. Disruptive Behavior Classroom G
The mean number of disruptive behavioral events decreased again from the first phase of the intervention (M = 109.15) to the second phase of the intervention (M = 66.57) by 39.01%. Data was not collected upon a day-to-day basis during the second phase of the intervention as evidenced by the breaks in the line in the graph. Also, in the second phase of the intervention (criterion is equal to or less than 62 tallies) less variability (SD = 18.26) was observed as compared to the first phase of intervention (SD = 29.72).

A reversal was introduced in the middle of the second intervention phase. A very significant amount of variability was observed during the reversal (SD = 119.5). An increase in the mean number of disruptive behavioral events (M = 163) was observed, as compared to the first (M = 109.5) and second (M = 66.57) phases of the intervention. On one day of the reversal, the frequency of disruptive behavioral events returned to same level observed during baseline. The highest level of disruptive behavioral events observed during reversal was 247. While 202 was the lowest number of disruptive events observed during baseline and 246.33 was the mean number of disruptive behavioral events observed during baseline. Upon reinstatement, an immediate decrease in the

<table>
<thead>
<tr>
<th>School Days</th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Reversal</th>
<th>Intervention Phase 2</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days Met</td>
<td>3</td>
<td>13</td>
<td>2</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>% Met</td>
<td>123</td>
<td>10</td>
<td>62</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Mean</td>
<td>246.33</td>
<td>109.15</td>
<td>163</td>
<td>66.57</td>
<td>168</td>
</tr>
<tr>
<td>SD</td>
<td>44.5</td>
<td>29.72</td>
<td>119.5</td>
<td>18.26</td>
<td>79.20</td>
</tr>
</tbody>
</table>

Table 11. Classroom G Intervention Phase Data
number of disruptive behavioral events was again noted. A third phase of intervention was never introduced in this classroom.

During follow-up, an immediate increase in the mean number of behavioral events was noted once again. On one follow-up day, the frequency of disruptive behavioral events returned to same levels observed during baseline conditions. The highest level of disruptive behavioral events observed during follow-up was 224. This number was more than 202; the lowest number of disruptive events observed during baseline and approached 246.33, the mean number of disruptive behavioral events observed during baseline.

Overall, a 72.98% decrease in the mean number of disruptive behavioral events was noted from baseline to the third phase of intervention, and a 31.8% decrease in the mean number of disruptive behavioral events was noted from baseline to follow-up.

Classroom H

In Classroom H, which was a fourth grade classroom, it was determined the intervention would take place in the afternoon during mathematics at approximately 1pm. On Mondays, Tuesdays and Thursday, mathematics class lasted 50 minutes, on Wednesdays it lasted 40 minutes and on Fridays it lasted 20 minutes. The variety in the schedule was due to students having special classes on certain days of the week. To maintain consistency in reporting the dependent variable, data were adjusted as the rate of disruptive behavioral events occurring per a 50-minute time period.

The first part of the math lesson involved the students sitting at their desks for direct instruction. A smart board was used to guide instruction. Next, the class would
complete problems at their desks either in small groups or individually. Afterwards, the class would come together and review the worksheets with the use of the smart board. The identified problem behaviors included calling out without raising one’s hand or talking during instruction or independent work time, and getting out of one’s seat during instruction and independent work. The replacement behaviors included: always raising your hand to talk in class, answering as a class when cued by the teacher, and staying in one’s seat during instruction and independent work.

In Classroom H, data were collected for seven points during baseline. The mean number of disruptive behavioral events during baseline was 463.43 per a 50-minute period. Based upon this number an initial criterion for mystery prize eligibility, of 150 disruptive behavioral events or less, per a 50-minute time period, was established, in conjunction with the teacher, a 67.62% reduction in disruptive behavior. Also, the criterion for reinforcement was pro-rated on the two days of the week during which math was not 50-minutes long. The first phase of the intervention lasted 10 school days. Students were able to meet the criterion for 10 out of the 10 school days or 100% of the time. During the second phase of the intervention, a new criterion of 60 or less disruptive behavioral events per a 50-minute time period was established, in conjunction with the teacher, reducing the number of allowed disruptive behavioral events by approximately 60%. The second phase of the intervention lasted 14 school days. Students were able to meet the criterion for 11 out of the 14 school days or 79% of the time. Finally, in the third phase of the intervention, a final criterion of 30 or less disruptive behavioral events per a 50-minute time period was introduced. This phase lasted 10 school days in total. Students
were able to achieve the new criterion for 4 of the 10 school days, or 40% of the time.

Data for disruptive behavior are presented in Figure 8 and Table 10. A very large amount of variability was noted during baseline (SD = 128.89). Upon introduction of the first phase of intervention (criterion is equal to or less than 150 tallies), an immediate change in level was observed with no latency noted. The mean number of disruptive behavioral events decreased by 82.31% from baseline (M = 463.43) to phase one of the intervention (M = 82). Moderate variability in the data was observed during the first phase of the intervention (SD = 31.11), but the data stabilized significantly as compared to baseline (SD = 128.89). The mean number of disruptive behavioral events decreased again from the first phase of the intervention (M = 82) to the second phase of the intervention (M = 45) by 45.12%. Also, in the second phase of the intervention (criterion is equal to or less than 60 tallies) even less variability (SD = 15.48) was observed as compared to the first phase of intervention (SD = 30.37) and baseline (SD = 128.89).

A reversal was introduced in the middle of the second intervention phase. A very significant amount of variability was observed during the reversal (SD = 180.31) and an increase in the mean number of disruptive behavioral events (M = 318.5) was observed during reversal. On one of the reversal days, the number of disruptive behavioral events returned to baseline conditions. The highest level of disruptive behavioral events observed during reversal was 446. Much higher than 280, the lowest number of disruptive events observed during baseline, and approaching 463.43, the mean number of disruptive behavioral events observed during baseline. Upon reinstatement an immediate decrease in the number of disruptive behavioral events was noted again. During the third
phase of the intervention (criterion is equal to or less than 30 tallies), a slight decrease in the mean number of disruptive behavioral events ($M = 42.4$) was observed as compared to the second phase of intervention ($M = 45$). Variability increased during the third phase of the intervention ($SD = 19.06$) in comparison to the second phase of intervention ($SD = 15.48$).

Note: C2 refers to the continuation of Criterion 2, which was interrupted by a reversal.

Figure 8. Disruptive Behavior Classroom H
Table 12. Classroom H Intervention Phase Data

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Intervention Phase 1</th>
<th>Reversal</th>
<th>Intervention Phase 2</th>
<th>Intervention Phase 3</th>
<th>Follow Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Days</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>14</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Criterion 50/min Days Met</td>
<td>150</td>
<td>60</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Met</td>
<td>100%</td>
<td>11%</td>
<td>79%</td>
<td>40%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>463.43</td>
<td>82</td>
<td>318.5</td>
<td>45</td>
<td>42.4</td>
<td>147</td>
</tr>
<tr>
<td>SD</td>
<td>128.89</td>
<td>31.11</td>
<td>180.31</td>
<td>15.48</td>
<td>19.06</td>
<td>105</td>
</tr>
</tbody>
</table>

During follow-up an increase in disruptive behavioral events was noted (M= 147), but once again the increase did not return to baseline conditions. Overall, a decrease of 90.85% disruptive behavioral events was noted from baseline to the third phase of intervention. A decrease of 68.28% disruptive behavioral events was noted from baseline to follow-up.

**Research Question Two**

How will the effectiveness of the Mystery Motivator Calendar intervention change over time? Results for the second research question are presented in this section. It was hypothesized intervention effects would decrease, as the intervention became less of a novelty to students. The mean numbers of disruptive behavioral events across various stages of the intervention were analyzed to investigate the results of this research question.

In Classroom A, the mean number of disruptive behavioral events continued to decrease as each new phase of the intervention was introduced. The mean number of disruptive behavioral events was lower during phase two (M = 28.65) of the intervention...
as compared to phase one (M = 41.4) of the intervention and the mean number of disruptive behavioral events was lower during phase three (M = 22.33) of the intervention than in phase two (M = 28.65). Furthermore, significant changes were maintained during follow-up (M = 30) as compared to baseline (M = 119). (Please refer to data presented in Table 5 and Figure 1 for more information.)

In Classroom B, the mean number of disruptive behavioral events also continued to decrease as each new phase of the intervention was introduced. The mean number of disruptive behavioral events was lower during phase two (M = 21.69) of the intervention as compared to phase one (M = 45.3) of the intervention and the mean number of disruptive behavioral events was lower during phase three (M = 11.83) of the intervention than in phase two (M = 21.69). Furthermore, significant changes were maintained during follow-up (M = 14) as compared to baseline (M = 133). (Please refer to data presented in Table 6 and Figure 2 for more information.)

In Classroom C, there were only two phases of intervention because it took more time for behavior to stabilize in this classroom and for students to be able to meet the criterion for 10 total school days. Nonetheless, the mean number of disruptive behavioral events was lower during phase two (M = 45.87) of the intervention as compared to phase one (M = 61). Furthermore, significant changes were maintained during follow-up (M = 38.5) as compared to baseline (M = 126), phase one and phase two of the intervention. (Please refer to data presented in Table 7 and Figure 3 for more information.)

In Classroom D, the mean number of disruptive behavioral events continued to decrease as each new phase of the intervention was introduced. The mean number of
disruptive behavioral events was lower during phase two (M = 10.8) of the intervention as compared to phase one (M = 22). The mean number of disruptive behavioral events was lower during phase three (M = 8.09) of the intervention than in phase two (M = 10.8). Furthermore, significant changes were maintained during follow-up (M = 14) as compared to baseline (M = 92.5). (Please refer to data presented in Table 8 and Figure 4 for more information.)

In Classroom E, the mean of disruptive behavioral events was lower in phase two (M = 37.72) of the intervention as compared to phase one of the intervention (M = 63.54). However, a very slight increase in the mean number of disruptive behavioral events was seen during phase three (M = 41.54) of the intervention as compared to phase two (M = 37.72) of the intervention. This could partially be attributed to the outlier observed during phase three of the intervention. On one day of the intervention during phase three, a peak was observed in disruptive behavioral events higher than any other day during phases two and three of the intervention. Nonetheless, the mean number of disruptive behavioral events during the third phase of the intervention (M = 41.54) was much lower than baseline (M = 227.2). Furthermore, significant changes were maintained through follow-up (M = 81.5) as compared to the baseline (M = 227.2). (Please refer to data presented in Table 9 and Figure 5 for more information.)

In Classroom F, the mean number of disruptive behavioral events was slightly higher in phase two (M = 7.67) of the intervention as compared to phase one (M = 7.3) of the intervention. However, the mean number of disruptive behavioral events was lower during phase three (M = 3.63) of the intervention as compared to both phase one (M =
7.3) and phase two (M = 7.67). Furthermore, significant changes were maintained during follow-up (M = 15.5) as compared to baseline (M = 39.66). (Please refer to data presented in Table 10 and Figure 6 for more information.)

In Classroom G, there were only two phases of intervention because the teacher forgot to introduce the third phase. Nonetheless, the mean number of disruptive behavioral events was lower during phase two (M = 66.57) of the intervention as compared to phase one (M = 109.15) of the intervention. Furthermore, significant changes were maintained during follow-up (M = 168) as compared to baseline (M = 246.33). (Please refer to data presented in Table 11 and Figure 7 for more information.)

Finally, in Classroom H, the mean number of disruptive behavioral events continued to decrease as each new phase of the intervention was introduced. The mean number of disruptive behavioral events was lower during phase two (M = 45) of the intervention as compared to phase one (M = 82) of the intervention. Also, the mean number of disruptive behavioral events was lower during phase three (M = 42.4) of the intervention than in phase two (M = 45). Furthermore, significant changes were maintained during follow-up (M = 147) as compared to baseline (M = 463.43). (Please refer to data presented in Table 12 and Figure 8 for more information.)

**Research Question Three**

How will teachers rate the acceptability of the Mystery Motivator Calendar intervention as measured by a modified version of the Intervention Rating Profile - 20 (Witt & Martens, 1983)? Results for the third research question are presented in this section. It was hypothesized teachers participating in the intervention would rate the
intervention as acceptable given the results of previous studies. Teacher intervention acceptability was assessed using a modified version of the Intervention Rating Profile 20 (IRP-20). Acceptability scores were established by adding the numbers corresponding to teachers’ responses. Responses were captured on a scale ranging from 1 to 6. The number 1 corresponded to ‘Strongly Disagree’ and 6 corresponded to ‘Strongly Agree’. A score of 100 or higher was considered to be acceptable and the higher the score, the greater the intervention acceptability, with a maximum score of 120. Data capturing teacher responses per a classroom are presented in Tables 13 and 14 below.

Overall, seven of the eight teachers rated the intervention as acceptable. An average teacher acceptability score of 111.36 out of 120 was found (range of 92 to 119), with a mean item score of 5.56. Only the Classroom G teacher did not assign a score of 100 or more to the intervention.

Individual item scores were also analyzed. Items with the highest mean, indicating the strongest agreement included: behavioral problems were significant enough to warrant the use of this intervention (M = 5.75); the intervention was practical in the amount of staff contact time required (M = 5.88); the intervention was practical in terms of recording keeping time (M=5.75); teachers are likely to use this intervention again (M = 5.75); and teachers are likely to use this intervention because it requires little training (M = 5.88). The items with the lowest mean scores included: teachers would find this intervention appropriate for other problem behaviors (M = 5.13); the intervention was effective in changing problem behavior (M = 5.25); and that the intervention would
be appropriate before making a referral (M =5.13). All mean item scores on the IRP-20, including the lowest continue to indicate strong teacher agreement.

Table 13. IRP-20 Question Mean Scores and Questions Scores per Teacher

<table>
<thead>
<tr>
<th>Question</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Teachers would find the intervention suitable for problem behavior.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>2. Teachers would find the intervention appropriate for other problem behaviors.</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5.13</td>
</tr>
<tr>
<td>3. Behavior problems were severe enough for this intervention.</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5.75</td>
<td></td>
</tr>
<tr>
<td>4. This intervention was effective in changing problem behavior.</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5.25</td>
<td></td>
</tr>
<tr>
<td>5. This intervention was acceptable for problem classroom behavior.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5.63</td>
<td></td>
</tr>
<tr>
<td>6. Overall, this intervention would be beneficial for the classroom.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>5.38</td>
<td></td>
</tr>
<tr>
<td>7. I would use this intervention again in the classroom setting.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>5.75</td>
<td></td>
</tr>
<tr>
<td>8. This intervention would be appropriate before making a referral.</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5.13</td>
</tr>
<tr>
<td>9. This intervention would not result in negative side effects for children in the classroom.</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>10. This intervention would not result in risk to children.</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>11. This intervention would not be considered a &quot;last resort&quot;.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5.5</td>
</tr>
<tr>
<td>12. This intervention is practical for parent contact time required.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5.63</td>
</tr>
<tr>
<td>13. This intervention is practical for staff contact time required.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5.88</td>
<td></td>
</tr>
<tr>
<td>14. This intervention is practical in terms of record keeping time.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5.75</td>
<td></td>
</tr>
<tr>
<td>15. This intervention is practical for out of school work required.</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5.63</td>
<td></td>
</tr>
<tr>
<td>16. It would not be difficult to implement in a classroom of 30.</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>17. This intervention was not disruptive to classroom functioning.</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>18. It was not difficult to use this intervention and meet student needs.</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>19. Teachers are likely to use this intervention because it requires little technical skill.</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5.63</td>
<td></td>
</tr>
<tr>
<td>20. Teachers are likely to use this intervention because it requires little training.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5.88</td>
<td></td>
</tr>
</tbody>
</table>

Note. Wording of scale items has been summarized to fit table
Table 14. Overall Teacher Acceptability Score and Mean Item Score per Classroom

<table>
<thead>
<tr>
<th>Classroom</th>
<th>Overall Acceptability Score</th>
<th>Mean Item Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>119</td>
<td>5.95</td>
</tr>
<tr>
<td>B</td>
<td>114</td>
<td>5.7</td>
</tr>
<tr>
<td>C</td>
<td>115</td>
<td>5.75</td>
</tr>
<tr>
<td>D</td>
<td>115</td>
<td>5.75</td>
</tr>
<tr>
<td>E</td>
<td>117</td>
<td>5.85</td>
</tr>
<tr>
<td>F</td>
<td>115</td>
<td>5.75</td>
</tr>
<tr>
<td>G</td>
<td>92</td>
<td>4.55</td>
</tr>
<tr>
<td>H</td>
<td>104</td>
<td>5.2</td>
</tr>
<tr>
<td>Mean</td>
<td>111.36</td>
<td>5.56</td>
</tr>
</tbody>
</table>

**Research Question Four**

On average, how will classroom students rate their satisfaction with the Mystery Motivator Calendar intervention, as measured by two modified versions the Children’s Intervention Rating Profile (Witt & Elliott, 1985)? Results for the final research question are presented below. It was hypothesized that overall classroom students would rate the intervention as acceptable given the results of prior studies (Bennett, 2010; Deutscher, 2004; Madaus et al., 2003; Musser et al., 2001).

Two modified versions of the Children’s Intervention Rating Profile (CIRP) (Witt & Elliott, 1985) were used to assess student intervention satisfaction. One version of the scale, with simplified language and smiley faces instead of numbers, was used for kindergarten and first grade participants. On this scale different smiley faces corresponded with different number scores, with the face with the biggest smile being the equivalent of a 5 or “I agree very much.” The next face with the smaller smile was the
equivalent of a 4 or “I sort of agree,” the neutral face was the equivalent of a 3 or “I
neither agree nor disagree,” the face with a small frown corresponded to a 2 or “I sort of
disagree,” and finally the face with the biggest frown corresponded to a 1 or “I disagree
very much.” On this scale, items 2, 3 and 4 were reversed scored so a higher score would
indicate a higher level of intervention satisfaction. For third and fourth grade participants,
a numerical rating scale was used. Students were asked to rate the intervention on a scale
of 1 (I agree very much) to 5 (I disagree very much). On this scale items 1, 5, 6 and 7
were reversed scored so a higher score would indicate a higher level of intervention
satisfaction. Data capturing student responses per a classroom are presented in Tables 15
and 16 below.

Overall, students in kindergarten and first grade gave an average consumer
satisfaction score of 31.57 with 35 being the highest possible score. They gave a mean
item score of 4.5. Those in third and fourth grades assigned an average consumer
satisfaction score of 26.41 out of 35 with a 3.78 mean item score. Combined the average
intervention satisfaction score was 28.99 out of 35, with a mean item score of 4.14. Some
classrooms rated the intervention much higher than others. An analysis of the mean
scores of items reveals that for the most part students liked the Mystery Motivator
Calendar, they thought it was fair, and they thought it would be good for other students in
other classes. However, older students did give some indication that the intervention
caued problems in their class. Results also indicated that kindergarten and first grade
participants gave higher intervention ratings. However, these results must be interpreted
with significant caution given that the ratings of internal consistency on the modified
versions of the CIRP were questionable at best.

Table 15. Mean CIRP Item Scores and Mean Classroom Scores for Kindergarten and
First Grade

<table>
<thead>
<tr>
<th>Items</th>
<th>Classrooms</th>
<th>Mean Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I thought the Mystery Motivator Calendar was fair.</td>
<td>4.32</td>
<td>4.63</td>
</tr>
<tr>
<td>2. I thought my teacher was NOT fair when assigning tally marks on the calendar.</td>
<td>4.16</td>
<td>4.49</td>
</tr>
<tr>
<td>3. I thought the Mystery Motivator Calendar caused problems in my class.</td>
<td>3.48</td>
<td>3.83</td>
</tr>
<tr>
<td>4. There are better ways, different from the Mystery Motivator Calendar, to help children behave in class.</td>
<td>3.8</td>
<td>4.11</td>
</tr>
<tr>
<td>5. The Mystery Motivator Calendar will be good for other children in other classrooms.</td>
<td>4.64</td>
<td>4.75</td>
</tr>
<tr>
<td>6. I liked having the Mystery Motivator Calendar in my class.</td>
<td>5</td>
<td>4.85</td>
</tr>
<tr>
<td>7. The Mystery Motivator Calendar can help students in other classes behave.</td>
<td>4.56</td>
<td>4.79</td>
</tr>
<tr>
<td>Mean Score per an Item on Scale</td>
<td>4.29</td>
<td>4.5</td>
</tr>
<tr>
<td>Total Score Classroom Score from 35</td>
<td>30</td>
<td>31.57</td>
</tr>
</tbody>
</table>

*Note.* Items number 2, 4, 6 were reverse scored. A higher score connotes a higher level of intervention acceptability. Items were rated on a 5-point smiley face scale.
Table 16. Mean CIRP Item Scores and Mean Classroom Scores for Third and Fourth Grades

<table>
<thead>
<tr>
<th>Items</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Mean Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I thought the Mystery Motivator Calendar was fair.</td>
<td>4.96</td>
<td>5</td>
<td>4.04</td>
<td>4.66</td>
</tr>
<tr>
<td>2. I thought my teacher was too harsh in assigning tallies marks on the calendar.</td>
<td>2.57</td>
<td>3.26</td>
<td>2.92</td>
<td>2.92</td>
</tr>
<tr>
<td>3. I thought the Mystery Motivator Calendar might cause problems with my friends.</td>
<td>3.39</td>
<td>3.35</td>
<td>3.64</td>
<td>3.46</td>
</tr>
<tr>
<td>4. There are better ways to handle classroom behavior than to use the Mystery Motivator Calendar.</td>
<td>4.30</td>
<td>2.91</td>
<td>2.92</td>
<td>3.37</td>
</tr>
<tr>
<td>5. The Mystery Motivator Calendar will be good for other children in other classrooms.</td>
<td>3.61</td>
<td>4.04</td>
<td>3.8</td>
<td>3.82</td>
</tr>
<tr>
<td>6. I liked using the Mystery Motivator Calendar to manage the classroom behavior.</td>
<td>4.91</td>
<td>3.78</td>
<td>3.92</td>
<td>4.20</td>
</tr>
<tr>
<td>7. The Mystery Motivator Calendar would help other students manage their behavior in other classrooms.</td>
<td>4.09</td>
<td>4.22</td>
<td>3.6</td>
<td>3.97</td>
</tr>
<tr>
<td>Mean Score per an Item</td>
<td>3.98</td>
<td>3.80</td>
<td>3.55</td>
<td>3.78</td>
</tr>
<tr>
<td>Total Classroom Score from 35</td>
<td>27.83</td>
<td>26.57</td>
<td>24.83</td>
<td>26.41</td>
</tr>
</tbody>
</table>

Note. Items number 1, 5, 6 & 7 were reverse scored. A higher score connotes a higher level of intervention acceptability. Items were rated on a 5-point numerical scale.

Summary of Results

The present study investigated whether the Mystery Motivator Calendar intervention, which utilized a group contingency, variable-ratio of reinforcement schedule, would decrease disruptive classroom behavior. An ABAB, changing criterion design was utilized. Results were consistent with the hypothesis and significant reductions in disruptive behavior were observed in all eight participating classrooms. A summary of this data is presented in Table 17 below.
Table 17. Summary of Disruptive Behavior Data

<table>
<thead>
<tr>
<th>Class:</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Mean</td>
<td>119</td>
<td>133</td>
<td>126</td>
<td>92.5</td>
<td>227.2</td>
<td>39.66</td>
<td>246.33</td>
<td>463.43</td>
</tr>
<tr>
<td>Criterion 1</td>
<td>60</td>
<td>66</td>
<td>63</td>
<td>46</td>
<td>114</td>
<td>20</td>
<td>123</td>
<td>150</td>
</tr>
<tr>
<td>Phase 1 Mean</td>
<td>41.4</td>
<td>45.3</td>
<td>61</td>
<td>10.26</td>
<td>63.54</td>
<td>7.3</td>
<td>109.15</td>
<td>82</td>
</tr>
<tr>
<td>Criterion 1 Met %</td>
<td>100%</td>
<td>83%</td>
<td>67%</td>
<td>100%</td>
<td>91%</td>
<td>90%</td>
<td>77%</td>
<td>100%</td>
</tr>
<tr>
<td>Reversal Mean</td>
<td>41.5</td>
<td>38.5</td>
<td>73</td>
<td>22.5</td>
<td>83</td>
<td>24.5</td>
<td>163</td>
<td>318.5</td>
</tr>
<tr>
<td>Criterion 2</td>
<td>30</td>
<td>33</td>
<td>32</td>
<td>23</td>
<td>57</td>
<td>10</td>
<td>62</td>
<td>60</td>
</tr>
<tr>
<td>Phase 2 Mean</td>
<td>28.65</td>
<td>21.69</td>
<td>45.87</td>
<td>10.8</td>
<td>37.72</td>
<td>7.67</td>
<td>66.57</td>
<td>45</td>
</tr>
<tr>
<td>Criterion 2 Met %</td>
<td>65%</td>
<td>92%</td>
<td>19%</td>
<td>100%</td>
<td>91%</td>
<td>83%</td>
<td>58%</td>
<td>79%</td>
</tr>
<tr>
<td>Criterion 3</td>
<td>15</td>
<td>17</td>
<td>N/A</td>
<td>12</td>
<td>29</td>
<td>5</td>
<td>N/A</td>
<td>30</td>
</tr>
<tr>
<td>Phase 3 Mean</td>
<td>22.33</td>
<td>11.83</td>
<td>N/A</td>
<td>8.09</td>
<td>41.54</td>
<td>3.63</td>
<td>N/A</td>
<td>42.4</td>
</tr>
<tr>
<td>Criterion 3 Met %</td>
<td>17%</td>
<td>100%</td>
<td>N/A</td>
<td>91%</td>
<td>31%</td>
<td>100%</td>
<td>N/A</td>
<td>40%</td>
</tr>
<tr>
<td>Follow-Up Mean</td>
<td>30</td>
<td>14</td>
<td>38.5</td>
<td>14</td>
<td>81.5</td>
<td>15.5</td>
<td>168</td>
<td>147</td>
</tr>
</tbody>
</table>

Note: Please note the reversal phase occurred at various points in the above-mentioned classrooms. N/A stands for non-applicable. Classrooms C & G did not reach a third intervention phase. * In Classroom H, the initial criterion was a 67.62% reduction from baseline and the second criterion was a 60% reduction from Phase 1; this was the only class in which the initial criterion was more than a 50% reduction from baseline.

Data indicated the Mystery Motivator Calendar intervention was very effective at reducing the frequency with which disruptive behavioral events occurred in all eight classrooms. In all classrooms, immediate decreases in disruptive behavior were observed with no latency period from baseline to the first phase of intervention. According to both visual inspection of the data and comparisons of means across various intervention phases, data indicate the Mystery Motivator Calendar produced significant decreases in the frequency of disruptive behavioral events in each classroom. While complete reversals to baseline were not observed during withdrawal phases as expected in an
ABAB design, in almost all the classrooms the frequency of disruptive behavioral events increased. Additionally, in Classrooms B, D, and H, the pattern of data closely resembled a typical changing criterion design. In Classrooms A, E, F and G, the pattern of data somewhat resembled a changing criterion design. In Classroom C, the pattern of data would also have somewhat resembled a changing criterion design if not for an outlier in both phase one and phase two of the intervention. Furthermore, intervention effects were maintained in each classroom through follow-up.

The study also examined whether the effectiveness of the Mystery Motivator Calendar intervention would decrease over the course of the study as students became satiated by the intervention and the novelty wore off. Data were not consistent with the hypothesis that intervention effects would decrease over time. In six of the eight classrooms, the mean number of the frequency of disruptive behavioral events continued to decrease across intervention phases. In Classroom E, although the mean number of disruptive behavioral events increased in phase three of the intervention as compared to phase two, the increase was very small and could potentially be explained by the outlier that occurred during the third phase of the intervention. In Classroom F, a very slight increase in the mean number of behavioral events, of less than 1, was observed from phase one of the intervention to phase two of the intervention. However, the mean number of disruptive behavioral events was lower in phase three of the intervention than in both phase one and phase two of the intervention, indicating the effects of the intervention continued. Furthermore, in all eight classrooms, the effects of the intervention were maintained through follow-up when compared to the mean number of
disruptive behavioral events during baseline. Nonetheless, slight increases in disruptive behavior were observed when reinforcement was removed in most classrooms. Results were not consistent with the hypothesis because the effects of the intervention continued to have an impact of decreasing disruptive behavior throughout the intervention and continued through follow-up.

This study also examined teacher intervention acceptability ratings for the Mystery Motivator Calendar. It was hypothesized teachers participating in the intervention would rate the intervention as acceptable given the results of previous studies. For the most part, the results were consistent with the hypothesis. Teachers in seven of the eight classrooms rated the intervention as acceptable.

Finally, this study examined classroom satisfaction with the Mystery Motivator Calendar. It was hypothesized that students would be satisfied with the intervention, given the results of prior studies. In regards to this research question results were mixed. Five classrooms rated the intervention as acceptable, one classroom rated the intervention as somewhat acceptable, and while the remaining two classrooms rated the intervention more favorably than unfavorably, scores were not high enough to indicate satisfaction. Results also indicated kindergarten and first grade participants gave much higher intervention ratings than third and fourth grade classroom participants. However, these results must be interpreted with significant caution given that ratings of internal consistency on the modified versions of the CIRP were questionable at best.
CHAPTER FIVE
DISCUSSION

Introduction

The discussion is organized into six sections. First, the theoretical orientation of the study is explored. Next, a brief summary of the purpose is outlined. Then the results are discussed in the context of both the literature and the study’s theoretical orientation coupled with potential implications for the field of education. Next, the study’s limitations are discussed followed by suggestions for future research. Finally, overall conclusions are made.

Theoretical Orientation

Scientists, philosophers, and educators have long sought to understand why individuals act the way that they do. Many theoretical orientations seeking to explain human behavior have arisen and evolved over the years. Psychodynamic theorists believe that human behavior is motivated by an individual’s life history and internal forces (Schloss & Smith, 1998). These internal forces can include aggression, concern, sex drive and the need for love. Proponents of the medical model/biophysical theory believe that human behavior is driven by physical and biological factors coupled with environmental conditions (Alberto & Troutman, 2009; Schloss & Smith, 1998). Ecological theorists believe that both internal and external forces drive human behavior and that an
individual’s behavior interacts with environmental conditions (Schloss & Smith, 1998). Cognitive theorists believe that human behavior is influenced by an individual’s perception of reality (Alberto & Troutman, 2009).

This research project was based upon the principals of behavior theory and applied behavior analysis (ABA). Behavior theory states that all human behavior is learned and most behaviors are responses to environmental stimuli (Sprick, 2009). This means that every behavior can also be unlearned and shaped (Sprick, 2009). Behavior theory is based on the concept of conditioning. Conditioning includes three important principles: reinforcement, punishment and extinction (Skinner, 1953). Punishment decreases the occurrence of a behavior while reinforcement increases the occurrence of a behavior (Skinner, 1953). Extinction refers to the elimination of a previously learned response to a cue (Skinner, 1953).

An approach commonly utilized to both understand and modify the behavior of students, which grew from behavior theory is known as applied behavior analysis (ABA). Baer, Wolf and Risley first defined the phrase, “Applied Behavior Analysis” in the 1968 inaugural edition of the Applied Behavior Analysis Journal (Schloss & Smith, 1998). They defined ABA as the “process of applying sometimes tentative principles of behavior to the improvement of specific behaviors and simultaneously evaluating whether or not changes noted are indeed attributable to the process of application” (Baer, Wolf & Risley, 1968). In ABA the term “applied” refers to targeted behaviors, which are socially relevant or important. The term “behavior” refers to events that are observable and
measurable (Schloss & Smith, 1998). The term “analysis” refers to the process by which the effects of behavior change are measured (Schloss & Smith, 1998).

There are many advantages in applying the principles of ABA in schools as compared to the other theories discussed. ABA is practical, falls into the realm of expertise of educators, can be used to address a wide variety of behaviors with students from diverse backgrounds and with very diverse needs (Schloss & Smith, 1998), and ABA allows for evaluation of outcomes through precise, objective and verifiable data collection procedures. Hundreds of school-based studies utilizing ABA principles have demonstrated success (Schloss & Smith, 1998). This study falls into the same category.

**Purpose**

This study was designed to extend the research on variable ratio, interdependent group contingency classroom behavioral interventions, specifically those that assess the effectiveness of the Mystery Motivator. This study examined whether the Mystery Motivator Calendar would reduce disruptive classroom behavior in general education elementary school classrooms with diverse populations. A single-case ABAB, changing-criterion design was used to assess rates of disruptive classroom behavior. Secondly, this study also sought to examine whether the effectiveness of the Mystery Motivator Calendar would change over time as the novelty of the intervention wore off. Third, the study examined whether teachers would find the Mystery Motivator Calendar intervention acceptable for reducing disruptive classroom behavior. A modified version of the Intervention Rating Profile-20 was used to assess teacher intervention
acceptability. Finally, the study assessed classroom intervention satisfaction using two modified versions of the Children’s Intervention Rating Profile.

**Research Question One**

First, this study examined how the interdependent group contingency Mystery Motivator Calendar intervention would affect the frequency of disruptive behavioral events as identified by general education elementary school teachers. It was hypothesized this study would replicate the results of prior studies, which documented the effectiveness of the Mystery Motivator as an interdependent group contingency intervention in decreasing disruptive behavior in classrooms (Hoag, 2010; Kraemer, 2012; Murphy et al., 2007; Schanding & Sterling-Turner, 2010). Overall, results supported the original hypothesis. The Mystery Motivator Calendar was found to be very successful in decreasing disruptive classroom behavior. All eight classroom-participants showed immediate and significant reductions in disruptive behavior with intervention effects being maintained through follow-up. These findings are also consistent with previous research that has utilized the Mystery Motivator intervention to reduce disruptive behavior in individual students (DeMartini-Scully, Bray, & Kehle, 2000; Kehle et al., 2000; Matovic, 2010, Mottram, Bray, Kehle, Broudy, & Jenson, 2002).

As expected, in most cases, a complete return to baseline was not observed during withdrawal or follow-up. However, increases in disruptive behavior were observed when the intervention was removed. In ABAB designs, researchers expect behavior to revert to baseline or close to baseline conditions during intervention withdrawal phases (Kazdin, 2011). In this study, it was hoped the Mystery Motivator Calendar would teach students
to use replacement behaviors and these competencies would remain during withdrawals. However, it was expected that when potential for reinforcement was removed, student motivation to meet behavioral expectations would decrease and increases in disruptive behavior would be seen. This is consistent with behavioral principles, which ascertain that behavior is repeated when it is reinforced (Skinner, 1953).

Expectations were confirmed in all eight classrooms to various degrees. All classroom participants demonstrated continuation of the effect of the intervention during reversals and demonstrated continuation of the intervention effects through follow-up. Even though increases in disruptive behavior were seen during follow up, the frequency of disruptive behavioral events did not return to baseline conditions in any of the classrooms. These results were similar to Teta’s (2008) study, which examined the use of the Mystery Motivator on homework completion in children with ADHD. Even after the intervention was removed, homework continued to be completed at a much higher rate than during baseline. Results were also similar to Murphy et al.’s (2007) study, which utilized Mystery Motivators to reduce disruptive pre-school behavior. Improvement in classroom behavior continued long after the intervention was removed. Finally, Kehle et al. (1998), who utilized a Mystery Motivator package to address selective mutism in grade school children, also found that children continued to speak in school long after they were no longer rewarded for this behavior.

In addition to an ABAB design, this study utilized a changing criterion design. In a changing criterion design, researchers expect behavior to decrease or increase in a step-by-step process (Kazdin, 2011). In three of the eight classrooms, data patterns strongly
resembled that of a changing criterion design, and in another four classrooms, data patterns somewhat resembled a changing criterion design. In this study, it was not expected data would decrease in the quintessential stepwise fashion required of changing criterion designs (Kazdin, 2011). This was because the changes in criteria were significant, whereas typical changing criterion designs utilize more gradual reductions. The pattern of change in this study, likely reflects the rigorous and significant changes in criteria, which resulted in significant decreases in disruptive behavior between intervention phases.

The variable ratio component of the study design may also have contributed to the intervention’s dramatic results. Prior studies have shown variable ratio reinforcement schedules are known to strengthen reinforcers to the greatest degree (Alberto & Troutman, 2009; Beaman, Stoffer, Woods, & Stoffer, 1983; Cooper et al., 2007; Davis & Blankenship, 1996; Ferster & Skinner, 1957; Popkin & Skinner, 2003; Skinner, 1953). In Teta’s (2008) study, dramatic behavioral changes were also observed. Student homework completion significantly increased as a result of the Mystery Motivator intervention and the author partially attributed strengths of these changes to the known powerful effects of variable ratio reinforcement pattern. Many other Mystery Motivator researchers observed powerful intervention effects that could be attributed to the variable ratio reinforcement pattern (Bennett, 2010; Deutscher, 2004; Matovic, 2010; Schanding & Sterling-Turner, 2010). Likewise, in this study the dramatic degree of change may have been attributed to the unique power of variable ratio (Cooper et al., 2007; Davis & Blankenship, 1996) in comparison to other reinforcement schedules.
It is possible that the potent results of the intervention could also be attributed to more consistent behavioral feedback from teachers and clarification of behavioral expectations. Prior to the intervention, teachers had been inconsistently addressing the disruptive behavior. Their responses ranged from ignoring it, to correcting it, to sometimes even reinforcing it. For example, teachers would sometimes respond to questions of students who called out without raising their hands or those who got out of their seats at inappropriate times to answer questions. The Mystery Motivator Calendar coupled with the tally counters created an environment much more consistent in terms of behavioral goals, feedback and consequences. Also, due to the design of the Mystery Motivator Calendar intervention, students were reminded each day of the behavioral expectations for the intervention period. These reminders likely strengthened students’ awareness for expected behavior. These findings were similar to Teta (2008), who noted that prior to the implementation of the Mystery Motivator intervention, teachers provided students with inconsistent feedback regarding homework completion.

In this study, the tally counters provided additional benefits beyond teachers providing feedback with consistency. They also helped ensure feedback was immediate and consequences for disruptive behavior were provided in an efficient manner that did not interfere with instruction. Teachers simply held tally counters in the direction of the disruptive student and the counter made a clicking noise to signify that an additional tally had been added. In prior Mystery Motivator studies, teachers identified data collection procedures, which required them to stop teaching as burdensome and disruptive to the flow of classroom activities (Hoag, 2010; Matovic, 2010; Murphy et al., 2007). In this
study, the teachers not only praised the efficiency data recording procedures, but each teacher requested to keep the tally counter because it had been so easy to use.

The strength of the intervention may have also resulted from the type of contingency contract used. Studies that have utilized interdependent group contingency contracts, have found them to be very effective (Gresham & Gresham, 1982; Skinner et al., 1996; Skinner, Cashwell & Dunn, 1996). In fact, some studies have found interdependent group contingency interventions to be more effective than independent contingencies in decreasing disruptive classroom behavior (Gresham & Gresham, 1982). These types of contingency contracts can harness the power of peer pressure to promote behavioral changes. This phenomenon can be attributed to the fact that interdependent group contingencies have been found to promote increased social interactions and group cooperation as a result of the class working toward mutual goals (Skinner, Cashwell & Dunn, 1996). During the Mystery Motivator Calendar intervention, students were observed redirecting each other towards the behavioral goals. For example, if a student shouted out, another student would turn and put a finger to his/her lips to signify the other student to be quiet. This was not observed prior to intervention implementation. Prior to the Mystery Motivator Calendar, the inappropriate behavior of classmates was ignored or even sometimes reinforced. For example, students would laugh at the inappropriate comments or shout-outs of other students, thus reinforcing inappropriate behavior. Once the intervention was implemented students were observed encouraging each other to follow the behavioral goals. They refrained from reinforcing disruptive behavior more often. Thus, the interdependent group contingency design seemed to, at least partially, be
responsible for removing reinforcement for disruptive behavior.

The intervention also created an environment in which potential reinforcement was immediately available for appropriate behavior on a daily basis. The immediacy and frequency of reinforcers also may have contributed to the dramatic decreases in disruptive behavior. If students were able to meet behavioral expectations and a mystery prize was available, reinforcement was provided right away or the very next day. Research has indicated that providing reinforcement immediately after a behavior occurs strengthens the behavior (Skinner, 1953). Although, many other positive behavior support strategies were in place in most of the classrooms participating in this study prior to the Mystery Motivator Calendar, they provided reinforcement much less frequently. For example, some of the classrooms utilized tokens for positive behavior that could be traded in for prizes or entered in a raffle at the end of the week. For younger children, it could be especially difficult to control immediate behavior for the promise of far off rewards (Schloss & Smith, 1998; Sprick, 2009). The immediate availability of potential reinforcement in this study as compared to previously implemented strategies could be another reason why the Mystery Motivator Calendar produced such dramatic decreases in disruptive behavior.

The length of the intervention may have also contributed to the dramatic decreases in disruptive behavior observed. In this study, the Mystery Motivator Calendar was in place in for eight weeks. This time period allowed the researcher to use more than one criterion, with each consecutive criterion decreasing disruptive behavior to a greater extent. The requirement that students had to meet the target criterion for a period of ten
school days in total before the next criterion was implemented allowed for behavior to stabilize. It is likely that the cumulative effect of multiple criteria and stabilizing periods contributed to the large decreases in disruptive behavior that were observed.

Another reason the Mystery Motivator Calendar may have been so effective at eliciting behavioral change is that it increased students’ extrinsic motivation. Because this project started more than four months into the school year in each of the classrooms, it could be presumed that at this point, students were aware of behavioral expectations, but they were not displaying these behaviors with regularity. Thus, it could be assumed the disruptive classroom behavior was more likely the result of a performance deficit as opposed to a skill deficit (Sprick, 2009). The performance deficit was likely the result of insufficient motivation. This notion seems to be supported by the immediate and drastic reductions in disruptive behavior observed in all the classrooms on the very first day of the intervention. There was no latency, indicating the students already had the appropriate skills, but were not using them with regularity.

These immediate results further demonstrate the importance of using reinforcement in classrooms where students may lack the motivation to display appropriate behavior. The return to baseline during one day of reversal in two of the classrooms also supports the aforementioned notion. When the potential for reinforcement was removed, disruptive behavior increased dramatically. Once again, these results are consistent with the principles of behavior theory (Skinner, 1953). In Classroom H this increase in disruptive behavior was purposeful. Students actually voiced that if they could not have the calendar then they would talk during instruction
and shout out more often. Also, in Classroom D, a student who was redirected by another student to sit appropriately during the reversal replied that since his teacher was not using the tally counter, he did not have to sit down.

**Research Question Two**

The second research question was “how will the effectiveness of the Mystery Motivator Calendar intervention change over time?” It was hypothesized that a slight decrease in efficacy would be observed as the novelty of the intervention faded and students became satiated with the intervention. Study results did not support this hypothesis. Overall, the intervention’s effectiveness did not appear to decrease over time. In seven of the eight classrooms, the lowest mean number of disruptive behavioral events occurred during the last phase of the intervention. In six of the eight classrooms, disruptive behavior continued to decrease across all phases of the intervention. In the remaining classrooms, very slight increases in disruptive behavior were seen between phase one and phase two in one classroom, and between phase two and phase three in the other classroom. Furthermore, intervention effects were maintained through follow-up in all eight classrooms, even though some increases in disruptive behavior were observed once reinforcement was removed.

In the two classrooms in which disruptive behavior did not continue to decrease over the course of the intervention, it is difficult to ascertain whether this can be attributed to the novelty of the intervention wearing off or the fact that the criterion to receive reinforcement became more and more difficult to attain. Because the criteria for reinforcement became most difficult in the last phase, in many classrooms the number of
days the final criterion was met was the lowest. As a result, students received reinforcement less frequently. Thus, their motivation to meet behavioral goals may have slightly decreased. In order for interventions to be successful and for students to be motivated, attainable criteria should be developed (Schloss & Smith, 1998; Sprick, 2009). According to the Expectancy x Value = Motivation model, students’ expectancy to be successful is an integral part of motivation (Feather, 1982). As expectancy to succeed decreases so does motivation. In the last phase of the intervention, students’ expectancy to meet behavioral goals might have dropped too much to sustain their motivation.

Several students even made comments such as “we’ll never make it” after seeing the final criterion. Perhaps, the continuous reduction of disruptive behavior by 50% between phases was too high and went too far. Possibly smaller, more gradual changes should have been made, as is usually required of a changing criterion design (Kazdin, 2011), allowing for more realistic final criteria to be set.

Prior studies had not examined the effect of the Mystery Motivator intervention over longer periods of time. Therefore, the hypothesis that the effects of the intervention would decrease over time was based on the findings of Hall and Hall (1980), who suggest that variety of reinforcement is more effective, while repetition of the same reinforcer can lead to satiation. Thus, one would assume as students had more and more opportunity to earn prizes through the Mystery Motivator Calendar, the reinforcing effect of the calendar would not remain as strong. Because many of the studies that found the Mystery Motivator intervention to be very successful at eliciting behavioral change were only implemented for a period of two to three weeks (DeMartini-Scully et al., 2000;
Kehle et al., 2000; Matovic, 2010; Madaus et al. 2003; Musser et al., 2001), it was difficult to predict what would happen in an 8-week intervention period. The findings that the interventions’ effects did not decrease over an eight-week time period provide further evidence of the strength and sustainability of the intervention. This is congruent with research findings indicating variable ratio schedules are less prone to extinction than other reinforcement schedules (Alberto & Troutman, 2009; Beaman, Stoffer, Woods, & Stoffe, 1983; Cooper et al., 2007; Davis & Blankenship, 1996; Ferster & Skinner, 1957; Popkin & Skinner, 2003; Skinner, 1953). The Mystery Motivator Calendar likely provided enough variety for satiation not to occur. Students continued to be motivated to make behavioral changes over the course of the intervention.

Other potential reasons that the effects of the intervention did not decrease over time are examined in the prior section analyzing the impact of the Mystery Motivator Calendar on disruptive classroom behavior. These include: the opportunity for immediate reinforcement, an environment in which feedback and consequences were provided with greater regularity, use of contingency contracting, an increase in student motivation, and continued changes in criteria which kept on challenging the students.

Furthermore, results indicated that not only did the intervention effects not decrease over time, but continuation effects were also observed through follow-up. The Mystery Motivator literature is inconsistent in terms of whether intervention effects remain through follow up. Some studies showed a continuation of treatment effects upon completion (Kehle et al., 1998; Motram et al., 2002; Teta, 2008). Others, noted increases in disruptive behavior (DeMartini-Scully et al., 2000; Kehle et al., 2000; Matovic, 2010;
Musser et al., 2001), with one study seeing a complete return to baseline after the
intervention was withdrawn (Madaus et al., 2003). In this study, while increases in
disruptive behavior were almost always seen when reinforcement was removed, in all
cases behavior changes were maintained through follow-up. The continuation of
intervention effects through follow-up was most evident in kindergarten and first grade
classrooms and, to a lesser extent, in the third and fourth grade classrooms.

The extended, eight-week time period of the intervention may also have
contributed to continuation of intervention effects through follow-up. In other studies, the
lack of intervention effects during follow-up were attributed to the short duration of the
intervention, which was only two weeks in many cases (DeMartini-Scully et al., 2000;
Kehle et al., 2000; Matovic, 2010; Madaus et al., 2003; Musser et al., 2001). Kehle et al.
(2000) stated if more sufficient time been given to thin the intervention, behavior change
could have been more significant and long lasting.

**Research Question Three**

The third research question was “how will teachers rate the acceptability of the
Mystery Motivator intervention as measured by a modified version of the Intervention
Rating Profile – 20” (Witt & Martens, 1983)? It was hypothesized this study would
replicate the results of prior studies, and teachers would rate the intervention as
acceptable (Bennet, 2007, DeMartini-Scully et al., 2000, Kehl et al., 2000; Madaus et al.,
2003; Moore et al., 1994; Madaus, Kehle, Madaus, & Bray, 2003; Mottram et al., 2002).
For the most part, the results supported the hypothesis. Seven of the eight classroom
teachers rated the intervention as acceptable. According to the IRP-20, teachers suggested
they found the Mystery Motivator Calendar to be practical in the amount of staff contact time required, they were likely to use it because it required little training, the intervention was practical in terms of recording keeping time, and that they are likely to use it again in the future. Anecdotally, some of the teachers expressed that Mystery Motivator Calendar would be a good tool to teach behavioral expectations at the beginning of the year. Also, most teachers, in whose classrooms the intervention did not run until the end of the school year, chose to reinstate the intervention on their own after follow-up.

It was interesting to note that both classroom teachers who rated the Mystery Motivator Calendar intervention lowest contained one student with clinically significant behavior problems. These two students were often responsible for the majority of the tally marks in these two classrooms. In Classroom H, this student was found eligible for placement in a separate classroom for students with severe emotional and behavioral problems. However, due to logistical reasons, he remained in Classroom H until the completion of the intervention and school year. In Classroom G, which was a parochial school with no special education services, the student’s behavior problems were significant enough that a university team came to observe the student. Teachers completed behavior-rating scales and results were given to the student’s parents in the hopes they would take the student for a medical evaluation. In both classrooms, students would have been eligible for reinforcement on a much more frequent basis if not for the behavior of the one student. Also, it is important to note these two classrooms showed the highest reversals to baseline conditions during intervention withdrawal and follow-up. It
is very likely that all the aforementioned reasons had an impact on the teachers’ ratings of the intervention.

Nonetheless, in Classroom H the teacher still rated the intervention as acceptable, whereas in Classroom G, the teacher did not. Interestingly, this would seem to contradict, Finn and Sladeczek’s (2001) findings that teachers tend to rate interventions as more acceptable when behavior problems are more difficult. However, this may simply indicate there may be better interventions than the Mystery Motivator Calendar when the behavior of concern involves primarily one student. In such cases, it may be more appropriate to implement individual behavior interventions, or individual behavior interventions in conjunction with classroom behavior interventions. Or independent contingency group interventions would be more appropriate than interdependent contingency interventions. In such cases, caution must also be taken in order to prevent scapegoating of students (Maag, 1999), especially when they have been identified as having an emotional or behavioral disorder or when suspected of such a disorder as in the case of Classroom G.

Furthermore, the only classroom in which the teacher did not rate the intervention as acceptable, Classroom G, not only had one student with clinically significant behavioral problems, but was also the classroom in which the original teacher who agreed to take part in the study, abruptly quit his job six days into the intervention. The new teacher kindly kept the project going, but it is important to note that she did not originally volunteer to participate in the study. Thus, her buy in may not have been at the same level as that of the other teachers who originally volunteered. This is consistent with studies
that show that teachers are more inclined to buy in to interventions that were developed or approved with their participation (Datnow & Stringfield, 2011). It is possible the demands of taking on a new position, coupled with the demands of the Mystery Motivator study, challenges related to having a student with significant behavior problems and lack of personal buy-in all had a cumulative effect in her rating of the intervention.

**Research Question Four**

The fourth research question was “on average, how will classroom students rate their satisfaction with the Mystery Motivator Calendar intervention, as measured by two modified versions the Children’s Intervention Rating Profile” (Witt & Elliott, 1985)? It was hypothesized this study would replicate the results of prior studies and students would positively rate the intervention (Bennett, 2010; Deutscher, 2004; Madaus et al., 2003; Musser et al., 2001). For the most part, classroom students indicated that they were satisfied with the intervention. Five classrooms rated the intervention as “liked”, one classroom rated the intervention as “somewhat liked” and the remaining two classrooms rated the intervention more favorably than unfavorably, however ratings were not high enough to indicate intervention satisfaction.

However, these results must be interpreted with significant caution given the limitations of the instrument used to assess student intervention acceptability. It is important to note that an analysis of the internal consistency of the scale indicated questionable validity at best. These results must also be interpreted with caution given that anecdotal evidence is not congruent with the results. The lukewarm ratings of the
intervention in many classes were contradictory to the evidence of the popularity of the intervention. For example, in both Classroom D and in Classroom G students stated that they purposefully misbehaved during the reversals because they wanted to have the calendar back. In Classroom F students admitted that they had agreed to rate the intervention as poor for other classrooms because they were worried that they would lose the Mystery Motivator Calendar in their class if they indicated it was good for other classes. In Classroom C, where the intervention ran up to 2 days until the end of the year, when students saw the calendar being removed they shouted “no” in unison. The teacher promised to count tallies for the remaining two days. Also, in Classrooms A, E, F and H, students requested the teachers to keep the intervention until the end of the school year. In the remainder of the classrooms the intervention was delivered until the end of the school year.

Prior literature, which found students like the Mystery Motivator further supports the questionable student intervention satisfaction results in the present study (Bennett, 2010; Deutscher, 2004; Madaus et al., 2003; Musser et al., 2001). Once again, it is important to mention this contradiction may be attributed to the limitations of the instrument used, which will be discussed in greater detail in the limitations section.

Interestingly, the two classrooms in which teachers rated the intervention the lowest were the same two classrooms in which students rated the intervention the lowest. Many of the same reasons for low ratings by teachers in Classrooms G and H could be attributed to low student ratings. Namely, in both of these classrooms, one student was primarily responsible for most of the tallies. As a result, the classroom students were not eligible for reinforcement on a more frequent basis. This situation likely impacted
classroom ratings of the intervention. In Classroom G, on the day the class received the lowest number of tallies, students remarked they were able to accomplish this because the student with behavioral difficulties was absent. Thus, it is likely that students were unhappy that the whole class was being held accountable for the behavior of one student.

Results, of the Kraemer (2012) lend support to the aforementioned explanation. In this study, students gave more favorable ratings to the *Get ‘Em on Task* intervention than to the Mystery Motivator intervention even though they indicated satisfaction with both. It is important to mention that in the *Get ‘Em on Task* intervention students were rewarded for individual behavior while in the Mystery Motivator intervention, the group was rewarded for group behavior. It is likely students prefer interventions in which rewards are available for individual behavior as opposed to awards being available for the whole group. In a group contingency one student can “ruin it for everybody”. This idea is further supported by anecdotal evidence. In Classroom C, when the researcher was administering the CIRP item that assess the fairness of the Mystery Motivator Calendar, one of the kindergarten students said, “the Mystery Calendar can be fair ‘cause we get prizes, but it can also be not fair because when people are still being bad, they sometimes make us not get our prizes.”

Nonetheless, it is likely more important to consider teacher intervention ratings than student intervention ratings. Teachers are primarily responsible for implementing and facilitating behavioral interventions in classrooms coupled with a plethora of other obligations and responsibilities. Thus, teacher time is critical to consider. While students may prefer independent to interdependent contingency interventions, when the choice lies
between a time saving intervention and no intervention, it is preferable to implement the 
time saving one. Particularly, the intervention has been shown to be effective in 
managing classroom behavior and does put students at risk.

The student ratings of the intervention in Classroom G and Classroom H may 
further confirm the previously mentioned concept that the Mystery Motivator Calendar 
may not be an appropriate intervention for classrooms in which one student displays 
clinically significant disruptive behavior, which continues despite the intervention. In 
other classrooms, there were one or a few students, whose behavior was most 
problematic, however, these students responded to the intervention. Also, the behavior of 
these students did not approach clinical significance. As previously mentioned, in 
contexts in which one student has chronic and clinically significant behavior problems an 
independent group contingency intervention, or an individual behavior intervention 
coupled with an interdependent group contingency, may be more appropriate. While 
interdependent group contingency interventions have been shown to promote class 
cooperation and to save teacher time (Skinner et al., 2002; Skinner, Cashwell & Dunn, 
1996), these factors must be balanced with the needs of individual students in the 
classroom. Also, as previously mentioned, the potential for the scapegoating (Maag, 
1999) of students with disabilities must be seriously considered. The notion that the 
Mystery Motivator Calendar may not be the most appropriate intervention, given the 
circumstances discussed above, is also supported by an individual item analysis in 
Classroom G and in Classroom H. In both of these classrooms, student responses 
indicated that many agreed that there are better ways to handle classroom behavior.
An individual item analysis of participating classrooms also revealed students often thought their teacher was too strict/not fair in assigning tallies. It is possible that prior to the intervention students did not believe teachers were as strict because feedback and consequences were not as consistent. As noted in the discussion for Research Question One, prior to the intervention teachers had been inconsistently addressing the disruptive behavior from ignoring it, to correcting it, to sometimes even reinforcing it. Once the intervention was in place this was no longer the case. This may have impacted student’s perceptions that their teacher had suddenly become strict.

It is also interesting to note, while most students seemed to indicate they liked the calendar and thought it was fair, this did not necessarily mean they felt the intervention would be good for other students in other classes, their teacher was fair and that the Mystery Motivator Calendar was the best tool for handling classroom behavior. This is interesting, given that all the items on the CIRP are meant to measure the same construct, student intervention acceptability (Witt & Elliot, 1985). Perhaps, a reason for this discrepancy is a large majority of the CIRP’s validity data cited in the literature concerns its utility for discriminating between interventions (Finn & Sladeczek, 2001), as opposed to rating individual interventions, despite the fact that researchers have used the CIRP in this way (Robinson & Sheridan, 2000; Schnee, 2010). This will be discussed in greater detail in the limitations section below.

**Limitations**

Several threats to internal and external validity as well as methodological concerns may have impacted the results of this investigation. Many of the threats to
validity discussed below stem from the realities of conducting research in an applied setting like a school. To add clarification to the results, first the limitations, which impacted all or many classroom participants will be discussed, followed by limitations applicable to individual classrooms.

**Research Design Issues/Methodological Considerations**

A primary limitation of this study, which is a limitation of all single-case designs, is the generalizability is restricted to classrooms similar the sample (Theodore et al., 2004). While, students in participating classrooms came from quite diverse backgrounds, there were no classrooms in this study, which contained a primarily African American student population. Also, the focus of this study was on general education grade school classrooms, thus generalizability may be limited to this population only. The small sample size of the study (n=8), also limits the level of generalizability.

Another methodological concern that should be discussed as a limitation is the brief amount of time given to reversal and follow-up. These phases were implemented for two days to control for threats to validity. However, this two day time period may have been too brief. In many of the classrooms, especially Classroom G and Classroom H, significant variability was observed during the reversals. Specific school events could explain some of this variability, however, the research design could have been stronger, and threats to validity could have been minimized further if data were given the opportunity to stabilize during reversal and follow-up.

Also, a methodological concern that came into play across classrooms was teachers’ problems remembering they should change the criterion by lowering it by 50%
after students were able to achieve the criterion for 10 school days in total. As a result, the criteria did not get adjusted in a timely manner across classrooms.

Some unexpected consequences resulted from the intervention’s popularity that must be mentioned as limitations. For example, in some classes students verbalized that they would purposefully talk out and engage in disruptive behavior during reversals because they wanted to have the calendar back. In Classroom F, as previously mentioned, students admitted they did not agree the intervention would be good for children in other classrooms because they were afraid this meant they would no longer have the Mystery Motivator Calendar in their class. In all of the classrooms, in which the intervention did not run until the end of the year, students asked to have it reinstated until the end of the year. This might indicate the intervention was almost too reinforcing. Teachers must take into consideration that it might be very difficult to fade or remove this intervention and maintain behavior change in their classroom. Prior studies have also noted it can be difficult to fade reliance on behavioral systems, which provide students with unnatural reinforcement, such as this one (Schloss & Smith, 1998). Taking this into consideration, instead of removing the intervention when desired behavior changes are observed, teachers could instead continue to add additional behavioral goals or they could extend the duration of the intervention.

Another limitation of this study, from the perspective of the PBS framework, is the intervention focused on disruptive behavior. These meant teachers might have paid more attention to negative behavior than positive behavior while the intervention was in place. One of the primary principles of PBS is the recognition of appropriate or positive
student behavior (Sprick, 2009). In fact, PBS researchers often recommend a 3:1 ratio of recognition of positive behavior (Knoster, 2008; Sprick 2009) to correction of inappropriate behavior. However, in many of the classrooms in the study, the disruptive behavior was occurring with such frequency that it would have been difficult for teachers counterbalance their corrections with the appropriate amount of attention directed towards recognizing positive behavior. Conversely, the intervention strongly aligns with other principles of PBS such as clear expectations, consistent feedback, rewards for appropriate behavior, use of the problem-solving model, and objective data gathering to monitor the intervention.

Issues Pertaining to History

Also, another threat to validity is history, or the events that occurred between measurements. In many of the classrooms, data was not collected with complete regularity, especially in Classroom F and Classroom G. In many other classrooms, school events made it impossible to implement the calendar every school day. Also, in each classroom, with the exception of Classroom E, a Spring Break or Easter Break occurred sometime in the midst of the intervention. While this interruption may have had an effect on the data, the time of the year the study was conducted was beneficial. By January, February, or March, when the intervention began, students should have learned and been aware of classroom expectations. Thus, the threat of maturation was minimized. Also, research indicates behavior problems tend to increase in the last month of school (Sprick, 2009). The Mystery Motivator Calendar was implemented into the last month of school in seven of the eight classrooms that participated, furthermore lending support to the
notion that the threats of maturation or history were unlikely to come into play in the success of the intervention.

Additionally, many others events occurred between measurements, in all of the classrooms that participated in the study. These events included: fire drills; assemblies; play practice; schedule changes; shortened periods; intervention periods being switched to another time of the day to accommodate special events; changes in regular classroom activities; students who received additional social emotional or academic support, coming in and out of the classroom during intervention periods; changes in recess; and substitutes in classrooms. Also, the reversal phase occurred at varying points into the intervention, which may have impacted the data. However, in every classroom the reversal was implemented after the students successfully completed the first intervention phase with the first criterion. Thus, it was hoped enough time had passed for the students to learn to use the replacement behaviors when the reversal was implemented. Also, the various classroom behavioral interventions and schoolwide programs already in place may have impacted the manner in which classroom participants responded to the intervention.

Some classrooms experienced experimental mortality, most specifically Classroom E, Classroom F and Classroom H. However, even though students left all of these classrooms, in Classroom E and Classroom H, new students were also added to the classroom rosters after moving to the district. Thus, both the addition and the loss of students may have impacted the data. Also, certain students being absent, or being sent to the office, during intervention time, resulted in changes to the sample, and thus may have had an impact on the data.
**Instrumentation/Measurement Issues**

Despite the modifications made to the Children’s Intervention Rating Profile, students in many of the classrooms had difficulty understanding the instrument and completing the rating scales. For example, even though the CIRP was completed one on one with the researcher and the task was explained to kindergarten and first grade participant, students were observed rating items addressing a similar idea very differently, such as, “The Mystery Motivator Calendar will be good for other children in other classrooms” and “The Mystery Motivator Calendar would help other students manage their behavior in other classrooms.” For many of these students, the task may have simply been too difficult. For example, when asked whether they agree or disagree about whether there are better ways to handle classroom behavior, some students just shrugged and said, “I don’t know.” Other times, when asked to rate individual items, the students pointed to a smiley face and said, “I want this one.” When asked if this meant they agreed or disagreed, many students simply shrugged. When questioned whether the Mystery Motivator Calendar caused problems in his class, one kindergarten student replied, “Yes, once it fell off the wall.” Also, the probability of student’s giving socially desired responses must be considered with kindergarten and first grade participants. Since students completed the rating scale one on one with the researcher, it is possible that they felt more uncomfortable rating the intervention unfavorably, despite the fact they were told that telling the truth was most important.

Similar limitations to the ones noted above were observed in Classroom F, a third grade classroom of bilingual Latino students. Many of the students had difficulty
understanding the language on the CIRP, most specifically, the negatively worded items. Also, a rumor was started that if the students would rate the intervention as “good for other classrooms” the Mystery Motivator Calendar would be moved to another class. Thus students agreed not to give favorable ratings to this item. Because the students had been discussing this in Spanish and the primary classroom teacher was not in the room when the students were filling out the survey, the researcher was unable to address the rumor at the time data was being collected.

The aforementioned information indicates that despite the modifications made to the CIRP, it may not have been an appropriate tool to assess intervention acceptability in the present sample. The validity of this instrument as a tool to assess student intervention satisfaction must also be considered. As previously mentioned, the literature has primarily indicated that the CIRP is a valid instrument for discriminating between interventions (Finn & Sladeczek, 2001) as opposed to rating individual interventions. It was also designed to be used with older students. Few studies used this tool with students younger than third grade and no studies assessed the use of this tool with students as young as kindergarten (Finn & Sladeczek, 2001).

It is also important to consider students completed the Children’s Intervention Rating Profile forms upon completion of the most demanding intervention phase. Timing may have impacted how students responded to the survey. In many of the classrooms, students met the final criterion least frequently. This was the case in both classrooms, which rated the intervention the lowest. It is possible that the less frequently the students received reinforcement, the lower they would rate the intervention and the intervention
might have been rated more favorably during the first intervention phase when the criteria were much easier to achieve.

Another limitation of the study design involves the use of the tally counter. The tally counter made a clicking sound when the teacher recorded disruptive behavior. This sound likely alerted the researcher to the behavior noticed by the teacher, thus potentially inflating the rate of inter-observer agreement. This phenomenon may partially have contributed to the high 92% rate of inter-observer agreement. However, the benefits of using the tally counter to simplify data recording outweighed this limitation.

**Issues Specific to Individual Classrooms**

Now that limitations that impacted participants across multiple classrooms have been discussed, limitations applicable to individual classrooms will be examined. First of all, in Classroom G, six days into the intervention, the primary classroom teacher quit his job. An aide that had been working in the classroom replaced this teacher. This event was unexpected for both kids and staff members and spikes in disruptive behavior were seen during the transition period. The replacement teacher did not volunteer to participate in the project, but kept it going in the classroom both as a favor to the kids, who wanted to keep it and to the researcher. This may have impacted her ratings of the intervention. Also, in Classroom G, on the first day of the reversal the students took a test and as a result their behavior was different than on the days of instruction during which the intervention was usually implemented. There was a decrease in the frequency of disruptive behavioral events in comparison to instructional days. On the second day of reversal, a student’s parent was in the room for a special project, and the students were
very excited. As a result, there was an increase in disruptive behavior. Furthermore, on the last day of the follow-up, the students had just returned from an ice cream social in the gym. They had consumed a large amount of sugar and a hard time transitioning back to classroom activities. These events may partially explain the significant variance observed during reversal and follow-up in Classroom G.

In Classroom H, a student with significant behavior and emotional problems, responsible for a large amount of the disruptive behavior in the classroom, was out of school for two five-day periods over the course of the intervention. Another event occurred in Classroom H, during the reversal that is important to mention. An incident occurred with the classroom students during lunch that prompted the principal to severely scold the class. He threatened if anything else happened in the classroom for the remainder of the day, students would not be allowed to go on a special fieldtrip the next day. He also took three of the most disruptive students out of the class and to his office for the remainder of the intervention period to discuss the aforementioned incident. Conversely, on the other day of the reversal, when the students heard the Mystery Motivator Calendar would be removed for two days, they said they would purposefully talk and shout out during class. These events may explain the significant variability observed during reversal in Classroom H. Finally, on the last day of follow-up in Classroom H the students took a math test. This change in typical activities may have been responsible for the significant variance in the occurrence of disruptive behavioral events during follow-up.

Also, in Classroom C, on the first day of follow-up, two of the most disruptive
students, who were responsible for receiving the greatest number of tallies, were absent. This had not occurred at any other time during the intervention. This likely had an impact on the data on the first day of follow-up. Students received a very low number of tallies on this day in comparison to other days during the intervention. In fact, it was the second lowest number of tallies recorded. On the second day of follow-up, both of these students had returned, which may account for the variability observed.

**Intervention Transportability**

Many of the limitations discussed above result from the difficulties of conducting research in applied settings, such as schools. While, on one hand, many of these uncontrolled for events could be seen as a weakness of research design, on the other hand, they reflect the reality of life in schools. The success of the intervention in each classroom, despite many unexpected events, speaks strongly to the transportability of this intervention. Transportability refers to the notion of bridging the gap between research and applied settings (Schoenwald & Hoagwood, 2001). Research has shown that many interventions that demonstrate efficacy in clinical settings fail to have appropriate transportability to applied settings (Dobson & Hamilton, 2002; Schoenwald & Hoagwood, 2001). In the field of education, in particular, there is a very large discrepancy between the number of existing evidence-based interventions and the number of interventions implemented (Walker, 2004). This can be partly explained by the lack of attention paid by researchers to the issue of transportability.
Directions for Future Research

There are many ways in which future studies could address the aforementioned limitations. One way to control for the various events that occurred across classrooms and schools could be to conduct a future study in multiple classrooms in the same school so special events would occur with more regularity across classrooms. Future researchers could also consider utilizing the Mystery Motivator Calendar intervention with students of different age groups and in different educational settings, such as special education classrooms. Examining the effectiveness of this intervention in special education or resource classrooms and with both older and younger students could contribute to further generalization. Also, future researchers should consider longer reversal and follow-up periods than those in the present study to allow for data to stabilize.

Future studies may also improve the Mystery Motivator Calendar’s alignment with the principals of PBS, most importantly by paying more attention to the recognition of positive behavior. Research has shown that students engage in appropriate behavior more frequently when that behavior is recognized (Sprick, 2009). Thus, it would be interesting to see if the impact of the Mystery Motivator Calendar would be even stronger if teachers focused on positive behavior. For example, instead of counting the frequency of disruptive behavioral events, future studies could, for example, assign classroom points for time intervals in which all students in the class are displaying appropriate behavior. By earning a certain number of points per a day, students would become eligible for reinforcement. In this way, both the effects of an interdependent group contingency and variable ratio reinforcement schedule would remain, but the focus would
shift from disruptive to positive behavior. If the intervention would be restructured in such a way, then it would be much easier for teachers to adhere to the 3:1 ratio of recognizing positive behavior to providing corrections recommended by PBS researchers (Sprick, 2009). A potential drawback for structuring the intervention in such a way is that it may be more time and work intensive for teachers. Also, teachers may forget to give points to students for appropriate behavior without additional support in the form of sound signals. Such signals would be important because researchers have noted teachers have more difficulty remembering to provide reinforcement when a class is behaving well (Sprick, 2009). Another potential solution would be the use of the tally counter to record positive behavior as opposed to disruptive behavior. However, this could also be difficult, given that students are expected to engage in appropriate behavior most if not all of the time. Despite, this study’s focus on disruptive behavior, it is important to consider that the design allowed for simple data recording procedures, which increased the intervention’s social validity. Also, the use of the tally counter to provide redirection for inappropriate behavior is much less disruptive and provides students with much less attention than the use of verbal redirection.

Future researchers should also consider, translating instruments, such as the CIRP, into the students’ native language when possible. Because the students were instructed primarily in English in Classroom F and because this tool had been used with students in the third grade in prior studies (Finn & Sladeczek, 2001), the researcher did not consider translating the instrument. Furthermore, given the plethora of limitations discovered with the use of the CIRP, future studies should consider developing different
ways of assessing student intervention acceptability. The CIRP appears to be a better tool for comparing interventions than for evaluating individual ones (Finn & Sladeczek, 2001). In addition to its limitations for evaluating individual studies, the CIRP does not seem to be a valid tool for evaluating intervention satisfaction in younger students, for whom it is too difficult even with modifications and individual assistance.

Future studies might also examine how the Mystery Motivator Calendar intervention impacts disruptive behavior across the school day. This study did not take into consideration the impact of the Mystery Motivator Calendar on student behavior during times the intervention was not utilized. It would be interesting to measure generalization of behaviors learned through the Mystery Motivator Calendar. Kindergarten and first grade teachers did anecdotally indicate that behaviors had generalized to the rest of the day, but no formal measurements were taken.

Finally, it would be interesting to examine whether the Mystery Motivator Calendar intervention focused on addressing disruptive behavior could influence academic work production and academic outcomes. In Classroom F, the intervention was implemented during individual writing time. When students were observed to be off-task, they would receive tally marks. In this classroom, the teacher anecdotally indicated students were producing more written work as a result of the intervention. With skills like writing, practice should improve ability. Thus, it would have been interesting to assess academic performance pre and post intervention to examine whether the Mystery Motivator Calendar, while addressing disruptive behavior, could improve academic outcomes.
Conclusions

Despite the limitations of this research project, results are consistent with previous studies that have found the Mystery Motivator to be a powerful tool for eliciting behavioral change in student populations. This investigation contributed to the literature in the field by applying the Mystery Motivator Calendar intervention to eight classrooms with diverse student populations across seven different elementary schools. In this study significant reductions in disruptive behavior were seen in all eight classrooms from baseline to the last phase of intervention. Furthermore, intervention effects continued through to follow-up even though some increases in disruptive behavior were seen when reinforcement was removed.

Secondly, results indicated the effectiveness of the intervention did not appear to decrease over time. In the majority of classrooms, disruptive behavior continued to decrease across phases, even while the criteria for intervention became more difficult to attain. Prior studies had not assessed the effects of the Mystery Motivator intervention over the time. Furthermore, results were consistent with prior studies that found that teachers rated the intervention as acceptable. Seven of the eight teachers who participated in this study rated the Mystery Motivator Calendar as an acceptable intervention. Finally, results were somewhat mixed in terms of student ratings of intervention satisfaction. Five classrooms indicated they “liked” the intervention, one classroom indicated the intervention was “somewhat liked,” and the remaining two classrooms rated the intervention more favorably than unfavorably, however, ratings were not high enough to indicate satisfaction. Due to significant limitations inherent in the tool and manner in
which student intervention acceptability was assessed coupled with contradictory anecdotal evidence, student intervention satisfaction data must be interpreted with caution. Overall, results indicated that the Mystery Motivator Calendar was a very powerful tool for reducing disruptive classroom behavior, the effects of the intervention did not decrease over time, and for the most part this intervention was accepted by teachers and liked by students.
APPENDIX A

RECRUITMENT EMAIL
Hello ____________,

My name is Eva Kowalewicz. I am a doctoral student at Loyola University Chicago. I’m looking for 8 general education grade school classrooms in which to implement a classroom behavioral intervention. This research will be conducted for my doctoral dissertation. The intervention can be considered as typical educational practice and does not involve the collection of any personally identifiable data. The intervention is known as the Mystery Motivator Calendar. It is very easy to implement and requires much less time on the teacher’s part than, for example, a token economy (i.e. students are able to earn tickets for good behavior which can later be turned in for prizes).

The intervention involves the following steps:

1. Determine 3 classroom behaviors that are the greatest impediment to classroom functioning.
2. Gather a baseline measure of the problem behavior.
3. State the expected behavior in positive terms.
4. Based on the baseline data set a behavioral goal.
5. Actively teach and reinforce the rules with behavior-specific praise and feedback. Display the rules.
6. Implement the intervention. This intervention involves creating a poster calendar. On certain days, interspersed at irregular intervals, there will be a letter “M” written under a small square of construction paper. If students are able to achieve set criteria for the day and there is an “M” on that particular calendar day, then they will receive a mystery prize. Squares of construction paper are used to cover up each day on the calendar so the students do not know whether they do or do not receive a prize on that particular day. Not only are the days on which prizes are available a mystery, but the prizes themselves are also a mystery.
7. If the students meet the criterion for the day and there is a letter “M” on the calendar, they will draw a card from the manila envelope and receive the designated prize.
8. At the end of the period, the square will be removed, regardless of whether or not the students were able to achieve the set criterion so they can see whether or not an “M” was available for that day.
9. The number of inappropriate behavioral incidents will be recorded directly on the calendar. This calendar will serve as a classroom progress-monitoring tool.
10. After the intervention the teacher and students will evaluate their satisfaction with the intervention with brief surveys that should each take less than 5 minutes to complete.

I will provide all the intervention materials (including calendars and rewards), coaching and consultation. I’m looking for elementary school classrooms willing to commit to implementing the intervention for a total of 8 weeks. An additional period in which baseline will be gathered will also be required prior to intervention implementation. If you are interested, or would like more information please email me or call me at: redacted, eva.kowalewicz@gmail.com or ekowalewicz@luc.edu.

Kind Regards,

Eva Kowalewicz
Doctoral Candidate in School Psychology
Loyola University Chicago
ISBE Certified School Psychologist, Type 73 Certificate # 2489564
APPENDIX B

CONSENT DOCUMENTATION
Dear Parent,

Your child’s classroom has been selected as potential classroom for participation in a research study. The purpose of this letter is to inform you of your rights and of your child’s rights as a research study participant. You have the right to decline your child’s participation in the research study at any time.

Project Title: The Mystery Motivator Calendar


Introduction: This research is investigating the reduction of disruptive classroom behavior using the Mystery Motivator Calendar. The study is being conducted to meet dissertation requirements for a PhD in School Psychology at Loyola University Chicago. The following two university faculty co-sponsors are supervising the study: Dr. Gina Coffee (redacted) and Dr. Pamela Fenning (redacted). You are being asked to participate because your child is a student in a general education classroom with a diverse student population.

Expected Duration of the Study: The study will be conducted over an approximately 12-week period. First baseline data will be collected. Eight weeks will be dedicated to the Mystery Motivator calendar intervention.

Risk: There are no foreseeable risks involved in your child’s participation in the research beyond those experienced in everyday life.

Procedure/Intervention: The intervention is based on the principles of positive behavior support, in which the students will be rewarded for appropriate behavior. Research shows interventions based on positive behavior support are the very effective for efficient classroom management. Classroom management is critical because behavior problems in the classroom can lead to decreased academic learning time, decreased academic performance and lower standardized test scores.

The intervention utilized will be known as the “Mystery Motivator Calendar”. A poster will be created. On certain days, interspersed at irregular intervals, there will be a letter “M” written under a small square of construction paper. If students are able to achieve set criteria for the day and there is an “M” on the particular calendar day, then they will receive a mystery prize. Squares of construction paper are used to cover up each day on the calendar so the students do not know whether they do or do not receive a prize on that particular day. Not only are the days on which prizes are available a mystery, but the prizes themselves are also a mystery. The prizes are written on note cards placed in a manila envelope with a large question mark on it. If the students meet the criterion for the day and there is a letter “M” on the calendar, they will draw a card from the manila envelope and receive the designated prize. When a student engages in inappropriate behavior, the teacher will count the behavior on a tally counter. If the class gets less than a certain number of tallies they will be eligible for a prize. At the end of the period, the square will be removed, regardless of whether or not the class was able to achieve the set criterion so they can see whether or not an “M” was available for the day. At the conclusion of the intervention, both
students and the teacher will be asked to complete rating scales to ascertain their level of satisfaction with the classroom intervention. Data will be gathered to ascertain the effectiveness of the classroom intervention for a doctoral dissertation. No identifiable information will be gathered about your child at any time during the study.

**Benefits to the Participant:** Participants will not benefit directly from participating in the research. A contribution to the existing literature on classroom behavioral management methods can be obtained. Such information can then be applied to other schools and other classrooms.

**Confidentiality:** No identifiable records of participants will be collected. The data gathered will be examined only on a whole class basis. No data will be gathered on the behavior of individual students.

**Voluntary Participation:**
Your child’s participation in this study is voluntary. If you do not want your child to take part in this study, you have the right to withdraw him/her. If you decide to allow your child to participate in the study, you are free to withdraw his/her participation at any time without penalty and your child is free not to have to answer any question at any time.

**Statement of Consent:**
You do not need to sign a form to give consent for your child to participate in the research. However, if you do not want your child to participate in the study, please contact the researcher. You have the right to decline your child’s participation AT ANY TIME.

**Researcher/Contact Person for Questions:**
If you have any questions about this intervention or would like to decline the participation of your child, please contact Eva Kowalewicz, at eva.kowalewicz@gmail.com or ekowalewicz@luc.edu, or redacted. If you have questions about your rights as a research participant, you may contact the Loyola University Office of Research Services at (773) 508-2689.

Kind Regards,

Eva Kowalewicz, M.Ed.
Doctoral Candidate in School Psychology
Loyola University Chicago
ISBE Certified School Psychologist
Type 73, Certificate # 2489564
ekowalewicz@luc.edu
redacted
Carta de Consentimiento Para Padres

Queridos Padres,

La clase de su hijo/a ha sido seleccionada para participar en un estudio de investigación. El propósito de esta carta es para informarle de los derechos que tienen usted y su hijo/a como participantes en el estudio de investigación. Usted tiene el derecho a rechazar la participación de su hijo/a en el estudio de investigación en cualquier momento.

Título del proyecto: The Mystery Motivator Calendar (El Calendario Misterioso y Motivador)

Investigador(a): Eva Kowalewicz, M.Ed., Candidata al Doctorado en Psicología Escolar.

Introducción: Este estudio busca reducir la conducta perturbadora en el salón de clases con El Calendario Misterioso y Motivador (The Mystery Motivator Calendar). El estudio se llevará a cabo para cumplir con los requisitos de disertación para el doctorado en la Escuela de Psicología de la Universidad Loyola de Chicago. Los siguientes dos profesores universitarios y co-patrocinadores estarán supervisando el estudio: Dra. Gina Coffee y la Dra. Pamela Fenning. Usted está siendo invitado a participar porque su hijo/a es un estudiante en un salón de clases de educación general con una diversa población estudiantil.

Duración Esperada del Estudio: El estudio se llevará a cabo durante un período de aproximadamente 12 semanas. Los primeros datos de referencia serán recogidos. Ocho semanas se dedicarán a la intervención de El Calendario Misterioso y Motivador (Misterio motivador Calendar).

Riesgo: No hay riesgos previsibles involucrados en la participación de su hijo/a en la investigación más allá de los experimentados en la vida cotidiana.

Procedimiento / Intervención: La intervención se basa en los principios de apoyo del comportamiento positivo, en el que los alumnos serán recompensados por el comportamiento apropiado. La investigación muestra las intervenciones basado en el apoyo de comportamiento positivo son muy eficaces para el manejo del comportamiento del salón de clases. El manejo del salón de clases es fundamental, porque los problemas de conducta en el salón de clases puede llevar a la disminución del tiempo de aprendizaje académico, disminución del rendimiento académico y bajos puntajes en las pruebas estandarizadas.

La intervención utilizada será conocido como el "Calendario Misterioso y motivador". Un cartel será creado. En ciertos días, intercalados por intervalos irregulares, habrá una letra "M", escrito en un pequeño cuadrado de papel de construcción. Si los estudiantes son capaces de alcanzar los criterios establecidos para el día y hay una "M" en el día en particular, entonces, ellos recibirán un premio sorpresa. Cuadros de papel de construcción se utilizan para cubrir cada día en el calendario para que los estudiantes no sepan si recibirán o no en ese día un premio. No sólo son los días en que se dan los premios un misterio, pero los mismos premios también son un misterio. Los premios se escribe en tarjetas colocadas en un sobre con un gran signo de interrogación sobre ella. Si los estudiantes cumplen con el criterio para el día y hay una letra "M" en el calendario, ellos sacarán una tarjeta del sobre de papel y recibirán el premio designado.
Cuando un estudiante se involucra en un comportamiento inapropiado, el profesor hará que el estudiante escriba una señal encima o por debajo del papel de construcción para el día. Al final del período, el cuadrado será eliminado, independientemente de si los estudiantes fueron capaces de alcanzar los criterios establecidos o puedan ver o no una "M" disponible para el día. Los estudiantes también serán elegibles para recibir beneficios adicionales de comportamiento en dos semanas, un mes y dos meses en la intervención. Al término de la intervención, tanto a los estudiantes y el profesor se le pedirá que complete las escalas de calificación para determinar su nivel de satisfacción con la intervención en el salón de clases. Los datos se reunieron para determinar la eficacia de la intervención en el salón de clases para una tesis doctoral. No se reunieron información o identificación alrededor de su hijo/a en ningún momento durante el estudio.

**Beneficios para el participante:** Los participantes no se beneficiarán directamente de la participación en la investigación. Una contribución a la literatura existente sobre los métodos de manejo de la clase de comportamiento se pueden obtener. Dicha información puede ser aplicado a otras escuelas y otros salones de clase.

**Confidencialidad:** Ningún registro de identificación de los participantes serán archivados. Los datos recogidos serán examinados sólo con toda la clase. Ningúnos de los datos sobre el comportamiento de los estudiantes serán archivados.

**Participación voluntaria:** La participación de su hijo/a en este estudio es voluntaria. Si usted no desea que su hijo/a participe en este estudio, usted tiene el derecho de retirarlo/a del estudio. Si usted decide permitir que su hijo/a participe en el estudio, usted es libre de retirar su participación en cualquier momento sin penalidad y el niño/a estará libre para no responder a preguntas en cualquier momento.

**Declaración de Consentimiento:** No es necesario que firme un formulario para dar su consentimiento para que su hijo/a participe en la investigación. Sin embargo, si usted no quiere que su hijo/a participe en el estudio, por favor póngase en contacto con el investigador. Usted tiene el derecho a rechazar la participación de su hijo/a EN CUALQUIER MOMENTO.

**Investigador / Persona de contacto para preguntas:**
Si usted tiene alguna pregunta acerca de esta intervención o si desea rechazar la participación de su hijo/a, por favor póngase en contacto con Eva Kowalewicz, en eva.kowalewicz@gmail.com o ekowalewicz@luc.edu, o al redacted. Si usted tiene preguntas sobre sus derechos como participante de una investigación, puede comunicarse con la Oficina de la Universidad de Loyola de Servicios de Investigación al (773) 508-2689.

Mis cordiales saludos,

Eva Kowalewicz, M.Ed.
Candidata al Doctorado en Psicología Escolar
Universidad Loyola de Chicago
ISBE Psicóloga certificada de la escuela
Tipo 73, N ° de certificado 2489564
ekowalewicz@luc.edu
redacted
Teacher Consent Form

Project Title: The Mystery Motivator Calendar


Introduction: This research is investigating the reduction of disruptive classroom behavior using the Mystery Motivator Calendar. The study is being conducted to meet dissertation requirements for a PhD in School Psychology at Loyola University Chicago. The following two university faculty co-sponsors are supervising the study: Dr. Gina Coffee (redacted) and Dr. Pamela Fenning (redacted). You are being asked to participate because you are a teacher in a general education classroom with a diverse student population. Please read this form carefully and ask any questions you may have before deciding whether to participate in the study.

Expected Duration of the Study: The study will be conducted over an approximately 12-week period. First baseline data will be collected. Eight weeks will be dedicated to the Mystery Motivator Calendar intervention.

Risk: There are no foreseeable risks involved in participating in this research beyond those experienced in everyday life.

Procedure: The intervention utilized will be known as the Mystery Motivator Calendar. A poster calendar will be created. On certain days, interspersed at irregular intervals, there will be a letter “M” written under a small square of construction paper. If students are able to achieve set criteria for the day and there is an “M” on that particular calendar day, then they will receive a mystery prize. Squares of construction paper are used to cover up each day on the calendar so the students do not know whether they do or do not receive a prize on that particular day. Not only are the days on which prizes are available a mystery, but the prizes themselves are also a mystery. The prizes are written on note cards placed in a manila envelope with a large question mark on it. If the students meet the criterion for the day and there is a letter “M” on the calendar, they will draw a card from the manila envelope and receive the designated prize. When a student engages in inappropriate behavior, the teacher will count the behavior on a tally counter. At the end of the period, the square will be removed, regardless of whether or not the students were able to achieve the set criterion so they can see whether or not an “M” was available for that day.

The first activity the teacher will be involved in includes the Problem Identification Interview. The purpose of this interview will be to identify the three most disruptive classroom behaviors, to identify replacement behaviors and to gather basic demographic data on the class in general. The next activity the teacher will be involved in will be training on the implementation of the intervention according to the intervention protocol. Next, the teacher will implement the intervention according to the procedure outlined above. During 25% of the time period that the intervention is implemented, the researcher will be observing in the classroom for inter-rater reliability purposes. Inter-rater reliability means that the researcher will be checking whether she would assign check marks for the same behaviors that the teacher is assigning check marks. This is done to strengthen the research design not to evaluate the teacher in any way. Upon completion of the Mystery Motivator Calendar study, the teacher will be asked to fill out a short, 20-question survey, which should not take more than 5 minutes to complete. The survey evaluates the teacher’s satisfaction with the intervention.

Benefits to the Participant: By participating in the study, the classroom environment may be improved. Disruptive behavior of students may be reduced, which could lead to improved learning opportunities for all students. In addition, a contribution to the existing literature on classroom behavioral management methods can be obtained. Such information can then be applied to other schools and other classrooms. Participants will not benefit directly from participating in the research.
Confidentiality: No records of the students participating will be revealed to anyone else. The data will be examined for the class as a whole and there will be no data gathered on the behavior of individual students. Signed teacher consent forms will be stored at the researcher’s home. No one else will have access to these consent forms. No information linking the teacher the classroom will be published. Consent forms including teacher and school names will be destroyed upon dissertation defense.

Voluntary Participation: Participation in this study is voluntary. If you do not want to be in this study, you do not have to participate. Even if you decide to participate, you are free not to answer any question or to withdraw from participation at any time without penalty.

Statement of Consent: Your signature below indicates that you have read the information provided above, have had an opportunity to ask questions, and agree to participate in this research study. You will be given a copy of this form to keep for your records.

Researcher/Contact Person for Questions: If you have questions about this research study, please feel free to contact the researcher. If you have questions about your rights as a research participant, you may contact the Loyola University Office of Research Services at (773) 508-2689.

Eva Kowalewicz, M.Ed.
Doctoral Candidate in School Psychology
Loyola University Chicago
ISBE Certified School Psychologist
Type 73, Certificate # 2489564
ekowalewicz@luc.edu
redacted

___________________________________________  ____________________
Teacher Name  School

___________________________________________  ____________________
Teacher Signature  Date

___________________________________________  ____________________
Researcher Signature  Date
Letter of Cooperation

Project Title: The Mystery Motivator Calendar

Purpose of the Study: This research is investigating the reduction disruptive classroom behavior using the Mystery Motivator Calendar. The study is being conducted to meet dissertation requirements for a PhD in School Psychology at Loyola University Chicago. The following two university faculty co-sponsors are supervising the study: Dr. Gina Coffee (redacted) and Dr. Pamela Fenning (redacted).

Expected Duration of the Study: The study will be conducted over an approximately 12-week period. First baseline data will be collected. Eight weeks will be dedicated to the Mystery Motivator Calendar intervention.

Risk: Risks to participation are minimal and no more than what would be encountered in day-to-day events.

Procedure: The intervention utilized will be known as the Mystery Motivator Calendar. A poster will be created. On certain days, interspersed at irregular intervals, there will be a letter “M” written under a small square of construction paper. If students are able to achieve set behavioral criteria for the day and there is an “M” on that particular calendar day, then they will receive a mystery prize. Squares of construction paper are used to cover up each day on the calendar so the students do not know whether they do or do not receive a prize on that particular day. Not only are the days on which prizes are available a mystery, but the prizes themselves are also a mystery. The prizes are written on note cards placed in a manila envelope with a large question mark on it. If the students meet the criterion for the day and there is a letter “M” on the calendar, they will draw a card from the manila envelope and receive the designated prize. When a student engages in inappropriate behavior, the teacher will count a tally on a tally counter. If students get less than a certain number of tallies they will be eligible for a prize. At the end of the period, the square will be removed, regardless of whether or not the students were able to achieve the set criterion so they can see whether or not an “M” was available for that day. Students and the teacher will be surveyed regarding their satisfaction with the intervention at the completion of the study. Also, classroom demographic data will be collected.

Confidentiality: No records of the students participating will be revealed to anyone else. The data will be examined for the class as a whole and there will be no data gathered on the behavior of individual students. Signed teacher consent forms will be stored at the researcher’s home. No one else will have access to these documents. No information linking the data to specific schools, classrooms or teachers will be published. Letters of Cooperation will be sent directly to Loyola’s Institutional Review Board.

Researcher/Contact Person for Questions:
Eva Kowalewicz, M.Ed.
Doctoral Candidate in School Psychology, Loyola University Chicago
ISBE Certified School Psychologist, Type 73, Certificate # 2489564
ekowalewicz@luc.edu
redacted
This document certifies that _______________________________ the signatory official has a clear understanding and approves of the research procedures outlined in the research protocol, (e.g., recruitment, consent, and data collection) and is a willing participant in the research project. This project will evaluate the effectiveness of the Mystery Motivator classroom behavioral intervention and will be completed under the supervision of faculty members at Loyola University Chicago, School of Education, School Psychology Department.

________________________________ _________________________
Administrator Signature  Date

* Please note letter of cooperation should be printed on institutional/organizational letterhead.
Hello! My name is Eva Kowalewicz and I am researcher with Loyola University Chicago. I would like to ask you to participate in a study in your classroom. You don’t have to participate in this study if you don’t want to. This study is about classroom behavior. You will be rewarded as a class for good classroom behavior. The study involves a Mystery Motivator Calendar (show the students the calendar). On some days there is a ‘M’ under the paper, on some days there isn’t. The days on which there are ‘Ms’ under the pieces of paper are a mystery. In this study we will use the Mystery Motivator Calendar for 2 months. So if we start at the beginning of ______, we would finish at the end of ______.

We will have class behavior goals. Your teacher will let you know what these goals are before we start to use the Mystery Motivator Calendar. An example of a goal you might have is to stay quiet while working at your seats or to always raise your hand if you want to speak in class.

If you are able to achieve your class goals then you will be eligible for a mystery prize as a class. If you meet the class goal and there is a ‘M’ under the piece of paper on the calendar for the day you will be able to choose a mystery prize from the envelope. To be eligible for a mystery prize, you will have to get less than a certain number of tallies for the day as a class.

You get tallies when you fail to meet behavior goals. For example, if your classroom goal is to always raise your hand before talking in class, if you shout out an answer without raising your hand, you will earn a tally. Your class will have to get less than a certain number of tallies each day to be eligible for a class prize.

At the end of the study, you will be asked to fill out a survey about the Mystery Motivator Calendar. You will tell me whether your agree or disagree with statements such as:

- The Mystery Motivator Calendar was fair
- I liked having the Mystery Motivator Calendar in my class
- I think that other children would like to have a Mystery Motivator Calendar in their classrooms.

This is a research project and not a test. You won’t be graded on anything you do and the results will not affect your school grade. All you have to do is try as hard as you can to meet the behavioral goals for the class. No personal information about you will be gathered. There are no dangers involved in participating in this study. Do you have any questions? If you have any questions at any time you can contact me (I will give them my phone number and email address) or you can ask your teacher.

Do you agree to participate in this study?
APPENDIX C

PROBLEM IDENTIFICATION INTERVIEW
Date:
Classroom:
Grade:
School:
Interviewee’s Name:

1. Gather background demographic data about the classroom.

2. General statement to introduce discussion/identify problem: (e.g., Please describe the behaviors in your classroom you find disruptive in your teaching).

3. Behavior specification/operationalize: (e.g., Can you provide me with examples of problem behaviors).

4. What time of the day/class period/subject do you find these behaviors to be the most problematic or disruptive?

5. Specify behavior priorities (Prioritize the problems, “On a scale of 0 to 10, where 0 = no problem and 10 = severe problem, would you rate the problem behaviors you identified).

6. Identify the perceived behavior strength (e.g. How frequently does the behavior occur?).

7. Tentative definition of goal (i.e. How frequently could the behavior occur without causing problems?)

8. Question interviewee about past approaches interviewee has tried to solve the problem

9. Develop a summary of the three of the most problematic behaviors.

10. For each behavior identified, state the expected behavior in positive terms.

11. Provide rationale for data recording and for gathering baseline data.

12. Discuss data collection procedures.

13. Establish date and time to begin data collection

14. Establish data and time for next meeting
APPENDIX D

INTERVENTION PROTOCOL
1. Procedure for Teaching Behavioral Goals

The objective of this activity is for students to be able to identify and demonstrate their understanding of the behavioral goals that will be targeted through the Mystery Motivator Calendar intervention.

List the expectations to be taught on the board:

- Goal 1: Stated in positive terms.
- Goal 2: Stated in positive terms.
- Goal 3: Stated in positive terms.

Ask students why each behavioral goal is important.
“Class, today we will be talking about Behavioral Goal 1, Behavioral Goal 2 and Behavioral Goal 3. These are important skills we need to all learn and demonstrate so that our classroom runs smoothly. Can you tell me why you think Behavioral Goal 1 is important? Can you tell me why you think Behavioral Goal 2 is important? Can you tell me why you think Behavioral Goal 3 is important?”

Remember to praise all correct replies.

Have the students model the appropriate behavior.
“Great now that we have discussed why it is important to Behavioral Goal 1, Behavioral Goal 2 and Behavioral Goal 3, I’d like for you to practice what our goals look like. Can you show me how you would Behavioral Goal 1? Can you show me how you would Behavioral Goal 2? Can you show me how you would Behavioral Goal 3?

Remember to praise appropriate behavior and provide feedback.

Ask the students to describe negative examples of the behavior.
“Wonderful, now I know that you know what Behavioral Goal 1, Behavioral Goal 2 and Behavioral Goal 3 looks like, I’d like for you to give me some examples of what each one of our goals doesn’t look like. Can you tell me what it would not look like to Behavioral Goal 1? Give me a negative or opposite example of Behavioral Goal 1. Can you tell me what it would not look like to Behavioral Goal 2? Give me a negative or opposite example of Behavioral Goal 2. Can you tell me what it would not look like to Behavioral Goal 3? Give me a negative or opposite example of Behavioral Goal 3.

Remember to praise appropriate behavior and provide feedback.

Closing
“Great job. You have really shown me that you understand why our behavioral goals are important. You also did a good job showing me what our behavioral goals look like and a great job telling me what our behavioral goals do not look like. Before we finish up, can I have a volunteer tell me again what our behavioral goals are? (Allow student response). Great, now we are going to be talking about a classroom plan that we will use to make sure that you are meeting these behavioral goals. This plan will include a reward system.”
2. Announcement the Intervention is in Place

“Class, today we will be starting a classroom behavior intervention known as the Mystery Motivator Calendar. Today is DATE and we are going to start recording behavior on the calendar. As you can see there is a place for tally marks on the calendar. During the day we will record behavior directly on the calendar with tallies. Now we are going to talk about our classroom behavioral goals again.”

3. Review of the Behavioral Goals

“Let me review them again. Our goals as a classroom are the following:”

- Goal 1: Stated in positive terms.
- Goal 2: Stated in positive terms.
- Goal 3: Stated in positive terms.

“In order to help you remember what the classroom behavioral goals are each of you will receive a handout on which they are written (pass out the handouts). There is also a poster attached to the Mystery Motivator Calendar on which the goals will always be written so you do not forget. Do you have any questions about our goals?”

*This will vary depending on the target behaviors identified by the classroom teacher in the problem identification interview. The rules will be stated in positive terms. (e.g. instead of do not shout out answers, the rule will be always raise your hand.)

4. Script Describing Intervention

“Now, I will describe how the classroom goals we talked about work with the calendar. I want to reward you as a class for being able to meet the goals we just talked about. So today we are going try to get less than X tally marks on the calendar. You get a tally mark on the calendar each time you do not meet a classroom goal. We will talk about this more in a few minutes. At the end of the day we will remove the piece of paper to see whether there is an “M” behind the piece of paper. If there is an “M” and you have met your goal, then you will be eligible for a mystery prize. If there is no “M” then you will not be eligible for a prize.”

5. Procedure for Making Tallies on the Calendar

“Tally marks get counted and then written down when behavioral goals are not met. Remember when we talked about negative or opposite examples? Well, each time I see you do the opposite of what is expected your teacher will count a tally. For example, one of our goals is X (i.e. always raise your hand) that means that whenever you X (i.e. shout out an answer without raising your hand), your teacher will count a tally on the tally counter. (Teacher will hold the tally counter in the direction of the student who broke the
goal and count a tally). She will write the tallies you have received at her earliest 
convenience.”

*Please note teachers were instructed to redirect the student to the replacement behavior if the student appeared to be unaware that he/she was engaging in the disruptive 
behavior. If the student continued to engage in the disruptive behavior, four seconds after 
being redirected by the teacher, he/she would receive another tally.

6. Procedure for Determining Whether Criterion for Reinforcement is Met

* The criterion will be determined based on baseline data so the number of tally marks 
may vary. “At the end of the day the we will add up the number of tallies counted, if the class 
receives less than X number of tally marks then you will be eligible for prize. Even if the 
goal is not met for the day, the square piece of paper under which there may be an “M” 
will be removed so the class can see whether there was an “M” available for the day.”

7. Procedure for Choosing a Reinforcer from the Mystery Motivator Envelope

“If the class meets the behavioral goals for the day, and there is an “M” under the sheet of 
paper, then the class will be eligible for a prize. The teacher will pick one student to come 
up and pull out a sheet of paper on which the mystery prize is written from the manila 
envelope.”

Teachers should strive to provide students with the earned reinforcement as soon as 
possible so the reinforcement remains strongly tied to the behavior.

8. Procedure for Announcing the Criterion was Not Met

“So it looks like we weren’t able to meet the classroom behavioral goals for the day 
today. Tomorrow is another day. What can you do differently tomorrow so you are able 
to meet the goal?” Elicit responses from students; if they are unable to come up with 
answers provide feedback. Remind the students of the classroom behavioral goals.

9. Procedure for Announcing the Reversal

“Class, I have an announcement to make. Tomorrow, we will be suspending the Mystery 
Motivator Calendar for 2 days. The calendar will not be available and you will not be 
eligible for mystery prizes. Don’t worry; we will start up the calendar again in two days. 
However, for the next we are simply putting the calendar on hold.”

10. Procedure for Announcing Reinstatement

“Class today, we will be starting the Mystery Motivator Calendar. That means that I will 
be counting every time you are not able to meet the behavior goals we set on the tally
counter. Remember our goals are Goal 1: Stated in positive terms, Goal 2: Stated in positive terms and Goal 3: Stated in positive terms.”

11. Procedure for Announcing New Criterion

“Class, today we will be cutting the number of tallies that you are allowed to receive from X to X. You have been doing so well and you have been able to meet our last goal for 10 school days already. Congratulations! Today, we will have a new challenge. Remember our goals are Goal 1: Stated in positive terms, Goal 2: Stated in positive terms and Goal 3: Stated in positive terms. This means that today in order to be eligible for a mystery prize the class must get less than X tallies.”

12. Daily Procedure For Intervention Period

“Class, we are going to start to take tallies for our Mystery Calendar. As a class, let’s quickly review our goals. Remember our goals are Goal 1: Stated in positive terms, Goal 2: Stated in positive terms and Goal 3: Stated in positive terms. Great, now also remember that today we need to get X tallies in order to be eligible for our mystery prize. All right, we are starting tallies now!
APPENDIX E

MYSTERY MOTIVATOR CALENDAR IMAGES
Classroom A: April

APRIL

MONDAY TUESDAY WEDNESDAY THURSDAY FRIDAY
M M M M

16 17 18 19 20

23 24 25 26 27

30

1. Always Raise Your Hand to Talk!
2. No Talking or Shout Outs
3. Sit Criss Cross Apple Sauce

Your Teacher Will Signal for You to Answer as a Class

MONDAY TUESDAY WEDNESDAY THURSDAY FRIDAY
4 19 3 28 17

Spring Break

16 17 18 19 20

23 24 25 26 27

30
Classroom B: April

<table>
<thead>
<tr>
<th>MONDAY</th>
<th>TUESDAY</th>
<th>WEDNESDAY</th>
<th>THURSDAY</th>
<th>FRIDAY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td>🎉 School</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>🎉 Spring</td>
</tr>
</tbody>
</table>

Always Raise Your Hand to Talk!
* No Talking or Shout Outs

Your Teacher Will Signal for You to Answer as a Class

**Easter Break**

<table>
<thead>
<tr>
<th>MONDAY</th>
<th>TUESDAY</th>
<th>WEDNESDAY</th>
<th>THURSDAY</th>
<th>FRIDAY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Classroom F: April

<table>
<thead>
<tr>
<th>MONDAY</th>
<th>TUESDAY</th>
<th>WEDNESDAY</th>
<th>THURSDAY</th>
<th>FRIDAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>19</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>26</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Laid-In Feet Still When Mrs. Santiago Is Teaching
- Materials Out When First Asked
- Writing During Independent Work Time
- Eyes On Mrs. Santiago
- No School
- Instructional Time
- Instructional Time
APPENDIX F

TALLY COUNTER IMAGE
APPENDIX G

MYSTERY PRIZES PER CLASSROOM
Classroom A
- The teacher reads a story to class
- Candy
- Fruit snack
- Goldfish crackers
- Juice box
- Extra science experiment
- Eraser
- Pencil
- Music of choice during independent work
- 10 minutes outside recess
- Out of uniform day
- The teacher has to wear a silly hat
- Video after lunch instead of rest and read
- Sticker for the class
- Pajama day
- Lunch in the class with a movie
- Charger chip for the class

Classroom B
- The teacher reads a story to class
- Candy
- Fruit snack
- Goldfish crackers
- Juice box
- Extra science experiment
- Eraser
- Pencil
- Music of choice during independent work
- 10 minutes outside recess
- Out of uniform day
- The teacher has to wear a silly hat
- Video after lunch instead of rest and read
- Sticker for the class
- Pajama day
- Lunch in the class with a movie
- Charger chip for the class
- Show and share

Classroom C
- The teacher reads a story
- Juice box after recess
- Extra science experiment
- Eraser
- Pencil
- 5 minutes extra outside recess
- Teachers have to wear a silly hat
- Watch video after lunch instead of rest
- Sticker
• Bonus marble for everyone in the class
• One extra computer class

**Classroom D**
• Candy
• Fruit snack
• Goldfish crackers
• Eraser
• Pencil
• Music of choice during independent work
• 10 minutes free time
• 5 minutes free time
• The teacher has to wear a silly hat
• Video during lunch
• Pajama day
• 10 minutes time to play a game in class
• Extra $0.50 to spend in the class store

**Classroom E**
• Homework pass
• 10 minutes game time
• 5 minutes extra recess
• 5 minutes in the computer lab to play a game
• Goldfish crackers
• Rice crispy treat
• Candy
• Paw for the class
• Hat day
• Pencil
• Sticker
• Eraser
• Extra art time
• Teacher reads a story

**Classroom F**
• Teacher reads a story
• Students can bring a stuffed animal to class
• Candy
• Rice crispy treat
• Fruit snack
• Eraser
• Pencil
• Pencil sharpener
• Homework pass
• Music during work time
• 10 minutes free time
• 5 minutes free time
• Hat day
• Watch a video
• Pajama Day
• Movie during recess/lunch
• Smartboard game during recess

**Classroom G**
• Homework pass
• Candy
• Eraser
• Pencil
• Mini notepad
• Hat day
• Watch a video for 20 minutes in class
• Music of choice during independent work
• 5 minutes extra recess
• 10 minutes free time

**Classroom H**
• Homework pass
• 10 minutes free time
• 5 minutes extra recess
• Lunch in the classroom with a video
• Pajama day in the class
• Music of choice during independent work
• Juice box
• Candy
• Hat day
• Pencil
• Sticker
• Eraser
APPENDIX H

CLASSROOM BEHAVIORAL GOALS HANDOUTS
Classroom A Handout

1. Raise Your Hand.
   • No Talking
   • No Shout Outs

2. Signal to Answer as a Class

3. Sit Criss Cross Apple Sauce
1. Raise Your Hand.
   - No Talking
   - No Shout Outs

2. Signal to Answer as a Class

3. Sit Criss Cross Apple Sauce
1. Sit Criss Cross Apple Sauce

2. Eyes on Teacher
1. Raise Your Hand.
   • No Talking
   • No Shout Outs

2. Signal to Answer as a Class

3. Sit Criss Cross Apple Sauce
1. Raise Your Hand.
   - No Talking
   - No Shout Outs

2. Stay in Your Seats During Class

3. Keep Your Desk Closed During Class
## Classroom F Handout

<table>
<thead>
<tr>
<th>+ What this Looks Like</th>
<th>- What this Doesn’t Look Like</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Materials Out When First Asked</td>
<td>- Materials Not Ready</td>
</tr>
<tr>
<td>- Writing During Independent Work Time</td>
<td>- Not Working On Writing</td>
</tr>
<tr>
<td>- Only Talking About Writing</td>
<td>- Talking About Anything Besides The Assignment</td>
</tr>
<tr>
<td>- Eyes On Teacher When She Is Teaching</td>
<td>- Looking Around the Room When Teacher Teaching</td>
</tr>
<tr>
<td>- Hands And Feet Still When Teacher Is Teaching</td>
<td>- Playing With Objects When Teacher Is Teaching</td>
</tr>
</tbody>
</table>
1. Always Raise Your Hand
   - No talking, shouts-outs or making noises

2. Your Teacher Will Signal for You to Answer as a Class

3. Stay Seated During Instruction and Independent Work Time

4. Eyes on Books During Silent Reading.
Classroom H Handout

1. Always Raise Your Hand to Talk!
   - No Making Noises, Talking or Shout Outs

2. Your Teacher Will Signal for You to Answer as a Class

3. Stay Seated During Instruction and Independent Work Time
APPENDIX I

CHILDREN’S INTERVIEW RATING PROFILE
I am interested in learning your ideas about the Mystery Motivator Calendar. Below are some sentences. You may agree or disagree with the sentences. For each sentence, circle the smiley face that describes what you think. Remember, the most important thing is that you tell truth, it is OK to disagree.

*Instructions and questions were read individually to kindergarten and 1st grade participants. The scale was also completed one on one with the researcher to help ensure that students understood both the task and the items.

1. I thought the Mystery Motivator Calendar was fair.

2. I thought my teacher was NOT fair in assigning tallies marks on the calendar.

3. I thought the Mystery Motivator Calendar caused problems in my class.
4. There are better ways, different from the Mystery Motivator Calendar, to help kids behave in class.

5. The Mystery Motivator Calendar will be good for other children in other classrooms.

6. I liked having the Mystery Motivator Calendar in my class.

7. The Mystery Motivator Calendar could help other kids in other classes to behave.
Children’s Intervention Rating Profile (3rd and 4th Grade)

Developed by (Witt & Elliott, 1985)

I am interested in learning your ideas about the Mystery Motivator Calendar. Below are some sentences. You may or may not agree with the sentences. For each one, please circle the number that describes how much you agree or disagree with the sentence. Remember, the most important thing is that you truthfully let me know what you think. Please use the following guide:

1 = I agree very much
2 = I sort of agree
3 = I don’t agree or disagree
4 = I sort of disagree
5 = I disagree very much

1. I thought the Mystery Motivator Calendar was fair.

   1  2  3  4  5

2. I thought my teacher was too harsh assigning tallies marks on the calendar.

   1  2  3  4  5

3. I thought the Mystery Motivator Calendar might cause problems with my friends.

   1  2  3  4  5

4. There are better ways to handle classroom behavior than to use the Mystery Motivator Calendar.

   1  2  3  4  5

5. The Mystery Motivator Calendar will be good for other children in other classrooms.

   1  2  3  4  5

6. I liked using the Mystery Motivator Calendar to manage the classroom behavior.

   1  2  3  4  5

7. The Mystery Motivator Calendar would help other students manage their behavior in other classrooms.

   1  2  3  4  5
**Intervention Rating Profile - 20 (IRP -20)**

The purpose of this questionnaire is to obtain information that will aid in the selection of classroom interventions. General education teachers in general education settings will use this intervention. Please circle the number that best describes your agreement with each statement below. You are to rate the Mystery Motivator Calendar intervention, which was used to decrease disruptive behavior in your classroom. Developed by Witt & Martens (1983).

You are to rate the Mystery Motivator Calendar intervention, which was used to decrease disruptive behavior in your classroom. Developed by Witt & Martens (1983).

1. Most teachers would find the intervention suitable for the behavior problem described.

2. Most teachers find this intervention appropriate for behavior problems in addition to the one described.

3) The classroom’s behavior problem is severe enough to warrant use of this intervention.

4) This intervention should prove effective in changing a classroom’s problem behavior.

5) This would be an acceptable intervention for problem classroom behavior.

6) Overall, the intervention would be beneficial for the classroom.
7) I would be willing to use this intervention again in the classroom setting.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

8) This intervention would be appropriate for use before making a referral.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

9) This intervention would not result in negative side effects for the children in the class.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

10) This intervention would not result in risk to the children in the classroom.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

11) This intervention would not be considered a "last resort".

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

12) This intervention is practical in the amount of time required for parent contact.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

13) This intervention is practical in the amount of time required for contact with school staff.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |

14) This intervention is practical in the amount of time required for record keeping.

| Strongly Disagree | 1 | 2 | 3 | 4 | 5 | 6 | Strongly Agree |
15) This intervention is practical in the amount of out-of-school time required for implementation.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

16) This intervention would not be difficult to implement in a classroom of 30 students.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

17) This intervention is not disruptive to classroom functioning.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

18) It would not be difficult to use this intervention and still meet the needs of students in the classroom.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

19) Teachers are likely to use this intervention because it requires little technical skill.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

20) Teachers are likely to use this intervention because it requires little training to implement effectively.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree
REFERENCES


VITA

Eva Aleksandra Kowalewicz earned her doctoral degree in school psychology from Loyola University Chicago. In 2011-2012 she completed an Association of Psychology Postdoctoral and Internship Candidates (APPIC) accredited internship at Highland Park High School in Highland Park, Illinois. Throughout her tenure at Loyola she was involved in research projects centering on the following topics: academic interventions, adolescents and sexual health, social justice in school psychology programs, and students with emotional and behavioral disorders. She has presented and published in these areas.