Summer 6-13-2022

PROGRAMMING MAINTENANCE FOR THE MYSTERY STUDENT INTERVENTION

Sarah Litten

PROGRAMMING MAINTENANCE FOR THE MYSTERY STUDENT INTERVENTION

by

Sarah Litten

A Dissertation
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

Dr. Brad Dufrene, Committee Chair
Dr. D. Joe Olmi
Dr. Zachary LaBrot
Dr. R. Alexander Smith

August 2022

Published by the Graduate School

ABSTRACT

The current study evaluated the effects of the Mystery Student Intervention (MSI) in university and community-based preschool settings on the appropriate and disruptive behaviors in the classroom utilizing a randomized independent group contingency. This study extended the literature base of the MSI, which was previously conducted in Head Start classrooms by Pasqua (2019), and Pasqua and colleagues (2021) which determined the MSI to be effective at reducing disruptive behaviors in the classroom setting. A nonconcurrent multiple baseline design across three classrooms was used. This study also sought to further extend the literature by evaluating programmed maintenance using a partial sequential withdrawal method. The results of this study indicate that the MSI may be effective, consistent with Pasqua (2019) and Pasqua and colleagues (2021), but issues with baseline data limits internal validity. Based on the results, the effects of the MSI also generalized to other classroom activities in which the intervention was not being implemented for all classrooms. Additionally, in regards to maintenance, the results suggest that partial sequential withdrawal of the intervention may maintain the effects of the intervention over time. All classroom teachers found the MSI to be an acceptable and effective classroom intervention to a degree. This study adds to the literature base for group contingency interventions that are at least moderately effective in the preschool setting. Additionally, this study also contributes to the literature surrounding the maintenance of interventions and effective withdrawal methods.

ACKNOWLEDGMENTS

I would first like to acknowledge the efforts put forth by my graduate colleagues who helped me complete this project. This project would not have been able to be completed without your hard work and dedication and I am forever thankful for your efforts. I would also like to thank the preschool centers and teachers who participated for allowing me into their classroom's to complete this project. The time and effort you put forth will be thanked for years to come as this project is a stepping stone for future research and effective class-wide interventions. Lastly, I would like to thank my committee chair and committee members, Dr. Brad Dufrene, Dr. Joe Olmi, Dr. Zachary LaBrot, and Dr. R. Alexander Smith for their unwavering support throughout this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

AEB	Appropriately Engaged Behavior
DB	Disruptive Behavior
MSI	Mystery Student Intervention

CHAPTER I – INTRODUCTION

Currently, there are over 4.4 million children enrolled in preschool programs across the United States (U.S Department of Education, 2021). It is well known that the early years of life are critical for development. Research shows that early childhood education, such as preschool is positively associated with enhanced social-emotional development, academic achievement, and behavioral development as children age and transition to primary and secondary school (Carter et al., 2010; Love, 2010; Weiland & Yoshikawa, 2013). Children who attended preschool are less likely to face negative outcomes such as poor relationships, academic difficulties, absenteeism, increased risk of dropout, and juvenile delinquency (Barnett, 2008; Ehrlich et al., 2014; Reynolds et al., 2001; Schweinhart, 2013; Temple et al., 2000; Whitted, 2011). Although preschool attendance is associated with positive long-term outcomes, many preschool-aged children experience emotional and behavioral difficulties at this age (Egger & Angold, 2006). These difficulties may be associated with environmental factors such as race or ethnicity, parent's level of education, family structure, parent's occupation, low socioeconomic status, negative parent-child relationships, or exposure to violence (Andershed & Andershed, 2015; Hussar et al., 2020). As preschool children continue to exhibit behavioral difficulties in the classroom setting, this can impede upon valuable instructional time for many teachers (Jenson et al., 1998). Moreover, teachers may lack classroom management skills that effectively support children's inappropriate classroom behavior.

Siebert (2005) found that teachers reported they do not feel sufficiently trained in the area of classroom management. Due to teachers' difficulty with classroom behavior

management, many children have been suspended or expelled due to their behaviors exhibited in the classroom. Some of the disruptive behaviors that may be exhibited in the classroom include aggression, noncompliance, and property destruction (Bear et al., 2002). In a 2019 study by Zeng and colleagues, over 174,000 preschoolers were reported to have been suspended, and over 17,000 children are expelled annually due to disruptive classroom behaviors. This 2019 study further supports the findings from the 2005 Yale University Child Study Center study that found the national expulsion rate of preschool children is three times higher than that of K through 12 students (Gilliam, 2005). In contrast, multiple studies have found that preschoolers were less likely to be expelled when the teacher to child ratio was lower, children were only there for a few hours versus a full day, and when the teacher had external supports for classroom management and mental health needs (Phillips et al., 2000; Zinsser et al., 2019). A potential solution to this problem is consultation that includes providing preschool teachers with resources to implement evidence-based practices within their classrooms.

Fortunately, there are a variety of class-wide interventions that are evidence-based and ideal for teachers because they require limited training, resources, and time. Class-wide interventions are an effective method of targeting disruptive behaviors for reduction in classroom settings (Johnson et al., 1996). A large number of behavioral issues can be addressed by class-wide interventions and such interventions are less time consuming than individualized interventions (Simonsen, 2013). According to Walker and Shinn (2010), upwards of 80% of students may respond positively to class-wide interventions. Group contingencies are a type of class-wide intervention, are evidence-based, and are widely used in schools.

Group Contingencies

Group contingencies consist of the delivery of a reward based on a specific criterion that is met by the whole group or by a specific member of the group. Group contingencies consist of three different types: independent, interdependent, and dependent. In all group contingencies, the rewards are available, and the target behaviors are the same for all group members, but the manner in which the reward is delivered varies. In an independent group contingency, each person who meets the criterion receives the reward. The behavior of the other members of the group does not affect any individual's access to the reward. For example, each student must score 80% or greater on the spelling test to receive ice cream. In an interdependent group contingency, each individual's access to the reward is based on the group performing at or above a specific criterion. For example, if the class's average or a collective total score on a spelling test is 80% or greater, then the entire class receives ice cream. A dependent group contingency is a group contingency in which the delivery of reinforcement depends on the behavior of a specified student or a subgroup of students that must meet the criterion in order for all members of the group to receive reinforcement (Cooper et al., 2007). For example, a teacher may identify one student, and if that student scores 80% or greater on the spelling test, then the entire class will receive ice cream. Researchers have tested group contingencies for decades (Litow & Pumroy, 1975; Little et al., 2015) and as a result, there is a plethora of information available regarding the settings, persons, referral concerns, and outcomes of numerous group contingency studies.

Literature Reviews and Quantitative Syntheses of the Group Contingencies Literature

There have been a variety of reviews and quantitative syntheses of the group contingencies literature. Across those reviews, group contingencies in school settings have been described as effective in K-12 classrooms for a variety of student populations and referral concerns (e.g., Little et al., 2015; Maggin, et al., 2012). Additionally, researchers have evaluated the quality of experimental design and calculated effect sizes for the group contingencies in school settings. Maggin et al. (2012) evaluated design standards of the group contingency literature using What Works Clearinghouse standards for evaluating single-subject designs (WWC; Kratochwill et al., 2010). Based on their review, they determined that group contingencies implemented in school settings meets criteria for an evidence-based practice. Similarly, Maggin, Pustejovsky, and Johnson (2017) evaluated design standards and conducted a meta-analysis of the school-based group contingencies literature. They found that group contingencies in school settings meets criteria as an evidence-based practice. Moreover, Maggin et al. found an overall effect for reducing disruptive behavior (d = 1.95) and increasing academic engagement (d = 1.80).

Additionally, Little et al. (2015) provided further support for the use of group contingencies in school settings by examining 182 single case studies published between 1980 and 2010 and found strong overall and individual effect sizes. A variation of Cohen's (1988) d statistic was used and generated an effect size of 3.41, categorized as a strong effect. They examined in detail the types of group contingencies: dependent (d =3.75, n =11), independent (d = 3.27, n = 8), and interdependent (d = 2.88, n = 35). Their results provide further evidence of the effectiveness of group contingencies and indicated

that there may not be substantial differences in effectiveness across the three types of group contingencies.

Although researchers have conducted quantitative syntheses (i.e. meta-analysis) of the broad school-based group contingencies literature, there has not been a quantitative synthesis of the group contingencies used in preschool settings. However, Pokorski, Barton, and Ledford (2016) conducted a descriptive literature review of group contingencies in preschool settings. This review included 10 research studies (seven peer-reviewed publications and three dissertations) conducted between 1971 to 2013. It examined 28 variables across five areas of interest. These areas included study descriptors (i.e. participants, settings, implementers), topography and measurement of dependent variables, intervention characteristics (i.e. group contingency, training, and visual components), reward type (social, tangible, activity known and unknown), and reward selection, along with methodological rigor, outcomes, and credibility of results. They utilized Kratochwill et al. (2010) and Horner et al. (2005) single case design standards to analyze study quality, which are consistent with design standards set forth by the WWC. Inter-observer agreement (IOA), procedural fidelity, generalization, and social validity were examined for the purpose of assessing internal and external validity of the studies. Lastly, they utilized visual analysis of level, trend, and variability of the outcomes.

Child participants ranged in age from three to six years old. Nine of the 10 studies were conducted in regular preschool classrooms that included between seven to 20 children. The tenth study was conducted in a special education preschool setting. They found that group contingencies were more likely to be implemented in class sizes larger

than six. Nine studies reported collecting observational data for challenging classroom behaviors, and one study collected observation data for social skills. Three classrooms utilized standardized or norm-referenced assessments such as the Vineland Adaptive Behavior Scales (Sparrow, Balia, & Cicchetti, 1985) or the Conners' Teacher Rating Scale-Revised (Conners, 1997; Filcheck, 2004; Maus, 2007; Reitman et al., 2004). Of the 10 studies, five were reported to include "mystery" rewards (Filcheck & McNeil, 2004; Ling & Barnett, 2013; Maus, 2007; Murphy et al., 2007; Reitman et al., 2004), and two utilized rewards known to the students (Hunt, 2013; Swiezy, Matson, & Box, 1992).

In regards to outcome analysis, nine of the 10 studies attempted to demonstrate a functional relation between the group contingency intervention and improved child behavior. However, only 26% of those comparisons were able to demonstrate a successful functional relation and were successful. The findings for study rigor suggested low to moderate research quality. Only four of the 10 studies assessed generalization of findings, and only one study assessed maintenance (Filcheck, 2004; Herman & Tramontana, 1971; Swiezy et al., 1992). IOA for dependent measures was above 80% for nine of the 10 studies. Six studies assessed social validity or intervention acceptability. Five of the six found that teachers rated the interventions as very acceptable and one study (Reitman et al., 2004) reported variable acceptability. The studies did not meet Horner et al. (2005) recommendations for procedural fidelity, as only three of the 10 studies collected procedural fidelity data for 20% or more of the sessions and reported 80% or higher fidelity.

The overall results of this review suggested that relative to the broader group contingencies literature, far fewer studies have been conducted in preschool settings.

Moreover, although some preschool group contingencies studies included a functional relation between the intervention and improved child outcomes, many data sets did not include such a demonstration (Little et al., 2015; Maggin et al., 2012; Maggin et al., 2017). Finally, there are concerns regarding the methodological rigor of group contingency studies in preschool. As a result, additional research with experimentally rigorous designs and procedures that test group contingencies in preschool is needed.

Individual Group Contingency Studies in Preschool

Filcheck and colleagues (2004) compared the effects of an independent group contingency using the Level System to Parent Child Interaction Therapy (PCIT; Hembree-Kigin & McNeil, 1995; Herschell et al., 2002; McNeil & Filcheck, 2004) implemented by preschool teachers using an ABACC design. The Level System was an independent group contingency in which children would move their marker up for engaging in rule following behavior or down for engaging in rule violations. The children earned rewards if their marker was at a particular level at multiple points during the day. The Level System was implemented following baseline phases and was subsequently withdrawn prior to the implementation of PCIT. The PCIT intervention was introduced in two subsequent phases. For the first phase, teachers implemented Child Directed Interaction (CDI), which included teachers providing attention for children's appropriate behavior utilizing PRIDE skills. PRIDE skills encompass providing labeled praise, reflecting on the child's behaviors, imitating the appropriate behaviors a child engages in, describing the behaviors, and using a high level of enthusiasm during the interaction (Eyberg, 1988). This was followed by the second phase, Parent Directed Interaction

(PDI), which included teachers using the PRIDE skills taught during the CDI phase with the use of effective commands and time-out for rule violations.

During the Level System phase, children's' disruptive behaviors decreased and subsequently increased when the Level System was withdrawn. When PCIT was implemented, children's disruptive behaviors further decreased. Filcheck et al. (2004) did not re-implement the Level System with the independent group contingency after withdrawal. Therefore, without replication of treatment effects in the one classroom study, it is impossible to state that there was a functional relation between the Level System with an independent group contingency and reductions in children's disruptive behavior.

Similarly, Reitman et al., (2004) investigated the use of token economies utilized within a preschool setting to reduce problem behaviors. This study utilized an alternating treatments design (ATD) to evaluate the effect of individual and group contingencies with three participants. They first had the teacher identify the three most appropriate and inappropriate behaviors occurring in the classroom. These were identified as target behaviors and were worded as positive statements and posted in the classroom during the intervention (i.e. walking feet, hands and feet to yourself, listening to the teacher). Prior to beginning the intervention, children were taught to discriminate between appropriate and inappropriate behavior and to report the consequences of breaking the rules. To start the intervention, the teacher would explain the game and expected behaviors and role-play appropriate behaviors and consequences that would occur during the game. Consequences for appropriate and inappropriate behaviors were delivered in the form of a Velcro® ball attached to a token chart that the teacher would move up if the children

engaged in appropriate behaviors and down if the children engaged in inappropriate behaviors. The chart consisted of seven levels with the top three levels displaying smiley faces that increased gradually in size. Level one indicated the behaviors were excellent, level three indicated the behaviors were good and levels five through seven depicted sad faces gradually increasing in size indicating poor classroom behaviors. If the children placed within the top three levels (i.e. levels one to three), a buzzer would sound, and that child would have the opportunity to play the Rewards Target Game (RTG). The RTG consisted of a felt board covered with the numbers one through 10 depicting pool balls. A Velcro ball depicting a “cue ball” was then thrown at the board to determine the reward earned. Following the child's throw, a reward was delivered based on the number the ball landed on. Rewards consisted of small tangible items, dance parties, and other games.

One condition within the RTG was the *star intervention*, which is an individual contingency that consisted of selecting children to be "stars" within the group by randomly choosing names from a hat. The behaviors of the stars would determine whether the group gained or lost points on the token chart. At the beginning of the intervention, the star would announce the rules to the class and explain how their behavior would affect the movement of the cue ball on the token chart. If the star student engaged in a disruptive behavior, they were reprimanded by the teacher and would lose a level on the token chart. If the cue ball stayed at the neutral level, the star would gain the opportunity to play the RTG at the end of the session. The group intervention portion was similar to the star intervention in that a star was selected that was not one of the target children. The behavior of the other selected children within the classroom would determine if the RTG was earned. The results of this study indicate that the individual

contingency or the star intervention and group contingencies were effective at reducing problem behavior for the target children. There were no differences found between individual and group contingencies as a method of intervention. Although there were limitations in teacher acceptability, which was rated as variable, this study adds to the existing literature concerning the utilization and efficacy of group contingencies in preschool settings.

To add to the previous literature examining the use of group contingencies in preschool settings, Ling and Barnett (2013) conducted a study examining the effect of group contingencies during group learning activities. Two half-day preschool classrooms were selected for this study. A small group of children from each class was selected to participate in this intervention based on previous high rates of disruptive behavior displayed within the classroom. Examples of disruptive behaviors targeted were leaving the carpet during circle time, talking out of turn, and mild or playful aggression. Researchers also included teacher praise, teacher negative attention, and teacher-directed instruction as dependent variables within the study. An ABAB design with a multiple baseline element was utilized to assess the effectiveness of the intervention. The teachers were trained on how to implement the intervention while being provided direct feedback from the observers.

The interdependent group contingency component of the study consisted of the teacher explaining classroom rules to the children at the beginning of the intervention session (i.e. beginning of carpet time). The classroom rules were to sit with legs crossed, hands to selves, and listen to the teacher. The teacher reminded the children that they would have the opportunity to earn rewards. The teacher collected a frequency count of

disruptive behavior displayed in the classroom for the children participating in the study. If the children had less than the set goal number of disruptive behaviors, they would earn a reward. The teacher drew a card from a circle time box and provided access to the reward that was pulled. The results of this study showed an immediate decrease in disruptive behaviors in the classroom which was demonstrated in both group contingency phases. This study extended previous research by including teacher variables. Teachers' praise and direct instructions increased during intervention and teachers' negative attention showed a decrease from baseline levels. The use of randomized rewards may have contributed to and maintained the intervention effects for longer than a single reward. The randomized rewards may have been beneficial due to the likelihood that all children would alter their behaviors in order to gain access to an unknown reward. Ling and Barnett's (2013) results are consistent with previous research demonstrating the effectiveness of group contingencies in a preschool classroom (McNeil & Filcheck, 2004; Murphy et al., 2007; Pasqua, 2019; Pasqua et al, 2021; Reitman et al., 2004).

The literature base surrounding group contingencies in preschool settings is relatively limited. There is a small number of studies demonstrating the beneficial effects of group contingency interventions in preschool classrooms. However, given that being the case, there have been fewer studies that have tested novel adaptations such as unknown reinforcers, randomized rewards, or programmed maintenance of group contingency interventions with preschool children.

Unknown Reinforcers and Randomized Components

Within the group contingencies literature, there have been some interesting procedural variations with regard to the predictability of the delivery of the reward. In an

effort to control for students who may become reactive to specific components of interventions, components are often randomized, such as randomizing the behaviors observed, the reward, the type of group contingency, and the target student (Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Murphy et al., 2007; Theodore et al., 2004). This method could minimize reactive effects and maximize the effects of the group contingency. Recent methods that have gained popularity involve randomizing target students for intervention. This method of randomization can reduce the occurrence of retaliatory acts against target students if they fail to meet criterion. This method is also useful because it encourages the entire class to change their behavior (Williamson et al., 2009).

Some studies have included unknown or randomized reward components (e.g., Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Murphy et al., 2007; Reitman et al., 2004; Theodore et al., 2004). Kelshaw-Levering et al. (2000) investigated the use of randomized components within interdependent group contingencies in an elementary setting. This study consisted of 12 students who were selected by the researcher. However, the teacher reported she had difficulty managing the behaviors of all of the students within the classroom. This study is unique in that it featured a phase consisting of randomized reinforcers and a phase consisting of all randomized components such as individual, as in one person received a contingency or group contingency. Kelshaw-Levering et al. (2000) utilized an ABACBC multiphase time series design to assess the effects of randomized group reinforcement on disruptive behavior. The random reinforcement phase (RR+) utilized similar procedures to Mystery Motivator interventions (Kelshaw- Levering et al., 2000; Pasqua, 2019; Pasqua et al., 2021;

Theodore et al., 2004). The teacher began this phase by introducing the class rules. The teacher described the behaviors in which the students must engage in order to receive reinforcement (i.e., staying in their seat unless given permission to be out of their seat, raising hands before speaking, not talking to peers unless given permission, and doing what the teacher instructed them to do immediately). The instructional day was divided into three intervals, and the teacher recorded behaviors on a checklist provided by the researcher because the teacher wanted to deliver the reinforcers during natural school day breaks. The teacher listed types of reinforcers (e.g., extra recess, five to 15 minutes of free time, a special snack, points towards a party) and further explained to the students that if they met criterion, a reinforcer would be randomly selected from slips of paper in a jar.

The *all components randomized phase* consisted of similar procedures as the RR+ phase. However, the teacher did not tell the students the criterion for reinforcement. She explained to the students that the criterion was subject to change between time periods depending on the new criterion randomly selected by the teacher from a series of jars displayed to the students. Each jar was uniquely labeled with behaviors to be observed (e.g. off task), group or individual names, and reinforcers. The behavior jar included slips of paper describing the target behaviors and a number of the word "all" and numbers that range from zero to 36 representing the number of times the target behavior could occur during that interval. The group or individual jar included slips of paper with "whole class" or "individual student" written on them. If "whole class" was selected, then the teacher would record the whole class's behavior and vice versa for an individual student. The "names" jar featured each student in the classroom's name. The selected student

would be evaluated for the behavior selected from the behavior jar. If the student or the whole class met criteria, the entire class earned a reinforcer. The final jar contained the types of reinforcers available written on slips of paper.

The purpose of this study was to examine the effects of randomized group contingencies on reducing problem behaviors in an elementary classroom and to compare the effects of randomized reinforcers to randomized group contingencies. The overall results of this study suggested that randomized reinforcers and other randomized components within-group contingency interventions were effective for reducing students' disruptive behavior. Both treatment conditions resulted in decreased levels of disruptive behaviors. However, the randomization of multiple components proved to be slightly more effective than just randomizing rewards alone. Despite small methodological and procedural limitations such as teacher feedback to students during the intervention, which may have served as a prompt to the students to alter their behaviors, student selecting their rewards, the rules being posted during the intervention phases, and brief observation periods; this study demonstrated meaningful changes in student behavior. Furthermore, the results suggest randomized components within-group contingency interventions may be an effective class-wide behavior management strategy.

In a subsequent study, Theodore and colleagues (2004), investigated the effects of randomized reinforcers utilized with group contingencies within a high school setting to reduce disruptive behaviors. Theodore and colleagues (2004) utilized an ATD to assess the differential effects of interdependent, independent, and dependent group contingencies. The study incorporated design randomization in order to control for sequence effects and carry-over effects. This was done by counterbalancing the sequence

of treatments, meaning that the students received their treatment in a randomly assigned manner. This study featured three participants identified as having Oppositional Defiant Disorder (ODD) and a Serious Emotional Disorder. The teacher was given a randomly generated intervention schedule and one treatment condition was administered each day (independent, interdependent, or dependent). In order to ensure the students were able to discriminate the contingency for that day, the teacher would announce that day's intervention to the class and write in on the board. The students were notified that rewards were dependent on if they received five or fewer checks for following the new classroom rules (i.e., no voicing obscene words, following the classroom teacher's directions, orienting in the direction of the teacher or assignment, not talking to students who were working, and making no verbal putdowns regarding themselves, others or a particular situation). The classroom rules were placed on the blackboard in the students' view. Each treatment phase was employed for 15 days, and the final phase was the most effective treatment. The teacher was provided a list of the student's names and interventions in which she could record the student's behaviors. A jar was placed on the teacher's desk labeled reinforcers; the jar contained five slips of paper with possible reinforcers listed on each (i.e., late to class pass, free detention pass, soda, candy bars, and chips of their choice). If the student met the criterion for reinforcement, the teacher would select a slip of paper from the jar.

At the beginning of each treatment phase, the teacher would explain to the students the guidelines for the new phase. For the independent group contingency, the students were responsible for their own behavior. Therefore, all students had to meet criterion for the interdependent group contingency. During the interdependent group

contingency, each student needed to meet a specific criterion in order for the class to earnt the reward. More specifically, each student in the classroom needed to earn five or fewer check marks. For the dependent group contingency, one student was selected randomly by the teacher from a jar consisting of all of the student's names listed on slips of paper and remained unknown to the class. This student's name was not announced to the class, and this student's behaviors were observed during the intervention. If this student met their goal of five or fewer check marks, the whole class would earn a reward. The dependent group contingency is the only phase that featured a randomized component. The results of this study indicated that the group contingency was effective at reducing disruptive behavior within the classroom setting. The results of the varying treatment phases demonstrated the effectiveness of group contingencies as an intervention by the immediate change in behavior displayed as treatments were introduced. Overall, all three group contingencies were effective at reducing disruptive behaviors. However, the role of randomization in the dependent group contingency proved to be advantageous to behavior change. This follows suit with previous studies utilizing unknown or random reinforcers (Kelshaw-Levering et al., 2000). Despite design limitations with potential carry-over effects and potential confounds (e.g. order and carry-over effects), Theodore and colleagues (2004), provided another demonstration of the effectiveness of group contingencies with randomized components on reducing disruptive behaviors in school settings.

Mystery Student Intervention

In addition to making the reward and other aspects of the contingency random or unknown, researchers have randomized the target student that is eligible for earning a

reward (Kelshaw- Levering et al., 2000; Theodore et al., 2004). Pasqua and colleagues (2021) tested the Mystery Student Intervention (MSI), which included an independent group contingency and random selection of Head Start children that were eligible for earning the reward. Pasqua and colleagues (2021) utilized an ABAB design without maintenance or generalization phases. The MSI featured the use of "mystery students" whose names were randomly selected from a name bag and observed using a 90-second momentary time sampling interval-recording method by one of the classroom teachers. If the Mystery students met 60% appropriately engaged behavior they were provided with a reward for meeting their goal. The Mystery student's names were only revealed if they met criterion to earn a reward. The researchers coded for appropriately engaged behaviors and disruptive behaviors. The MSI was implemented during a predetermined time of day or activity in which the most disruptive behavior occurred. The results of this study showed disruptive behaviors immediately decreased and appropriate behaviors increased when the intervention was implemented.

Pasqua extended the research in MSI in 2019 with the inclusion of target child comparisons, generalization, and maintenance phases. This study utilized an ABAB design with maintenance and generalization phases across three classrooms. Three Head Start classrooms were selected for this study based on administrator or teacher reports of the need for behavioral management and high levels of disruptive behaviors or high numbers of behavior incidents reported to center administrators. The teachers identified appropriate behaviors they would like to see the children display in their classrooms and disruptive behaviors they would like to decrease in their classroom. The teacher selected one child to be the target student from each classroom. This selected child was one that

displayed disruptive behaviors more frequently than peers. This child was selected with the intention of comparing their behavior before and during intervention phases to determine if there was an intervention effect.

The teachers were then provided scripts detailing what their role was within the classroom during the intervention. Each classroom had two roles to be filled by the teacher and teacher assistant, the data collector who would record the child's behavior while wearing an interval timer or the instructor who delivers the rules of the game, provides reminders throughout the game, and delivers the rewards at the end of the game. The intervention took place during carpet time in the mornings, which lasted approximately 15 minutes. The teacher whose role was to collect data selected two children's names from a bag to be the Mystery students. The teacher did not announce the names of the Mystery students to the class. In contrast, the teacher told the children during the instructions that every child should engage in the appropriate behaviors during intervention. The teacher whose role was to be the data collector recorded data that were used for determining if the Mystery students met the criterion for the reward. Mystery students were revealed if they engaged in appropriate behaviors for 60% of the intervals.

The researchers coded for appropriately engaged behaviors and disruptive behaviors using an interval recording method using a momentary time sampling method. Observers coded behavior in the target setting in which the game was played (i.e., carpet time) and during randomly selected times and settings in which disruptive behavior was reported to also occur (e.g., center time) that served as a measure of generalization since the game was not played during those times. Maintenance data were collected two weeks after the intervention was terminated using the same observation methods used for

baseline and intervention. The results of this study were consistent with Pasqua and colleagues (2021), demonstrating decreased disruptive behavior and increased appropriately engaged behavior. For generalization, the findings were consistent with the previous literature and add to previous literature (Filcheck, 2004; Swiezy et al., 1992) which indicated that group contingencies would produce moderate effects in generalization settings in which the game was not played. The maintenance phase suggested that behaviors remained stable after the MSI intervention was removed; albeit, based on a limited number of maintenance sessions (i.e., approximately two sessions per class).

Group contingencies have been demonstrated as effective for improving children's behavior in preschool settings; albeit across fewer studies with K-12 students. Moreover, a limited number of studies have evaluated novel modifications such as randomized components or unknown reinforcers to group contingency studies in preschool classrooms (e.g., Pasqua, 2019; Pasqua et al., 2021). Unfortunately, the broader group contingencies literature as well the preschool group contingencies literature includes an insufficient number of studies demonstrating generalized and maintained effects of group contingency interventions (Filcheck, 2004; Kelshaw-Levering, et al., 2000; Pasqua, 2019; Pasqua et al., 2021; Swiezy et al., 1992; Theodore et al., 2004).

Group Contingencies and Maintenance

One way to assess the effectiveness of an intervention is to assess if the changes made during intervention are maintained over time after some or all of the intervention components have been removed. Maintenance is defined as the extent to which a learner continues to perform a behavior after all of or a part of the intervention that is responsible

for behavior change is no longer present (Cooper et al., 2007). Typically, maintenance is assessed during additional follow up phases after the treatment has concluded. There are three potential methods of withdrawing interventions and those are sequential-withdrawal, partial withdrawal, and partial-sequential withdrawal. A sequential withdrawal design involves one component of the treatment being withdrawn first and then withdrawing a second component of the treatment, and so on until all of the components of the treatment have been withdrawn. A partial withdrawal design involves one or all of the components of treatment being withdrawn from only one of the treatment phases in a multiple baseline design. A partial sequential design involves part or all of the treatment being withdrawn from one of the treatment phases in a multiple baseline design, followed by a subsequent withdrawal in a second baseline condition, and so on (Rusch & Kazdin, 1981). Similar to sequential withdrawal, another method of withdrawing an intervention is called fading. Fading is defined as systematically removing parts of an intervention over time (Cooper et al., 2007). For example, in Odom et al., (1992), preschool teachers were trained to provide prompts and visual feedback to their students to assess social interactions. The verbal prompts were faded first, followed by visual feedback, which was then followed by a maintenance phase with no verbal prompts or visual feedback available. By fading the verbal prompts and visual feedback gradually, the students were able to maintain similar levels of social interactions to when intervention was in place. Although the literature surrounding the use of maintenance in group contingency interventions is limited, a few studies have examined the effects of maintenance of behavior over time and with the withdrawal of intervention components over time (Pokorski et al., 2016).

Filcheck (2004) investigated the use of a whole class token economy to manage disruptive behaviors in preschool classrooms. This study utilized an ABAB withdrawal design with a 1-month follow up which examined the use of strategies already employed by the classroom teacher and a whole class token economy known as the Level System. Specifically, the researchers wanted to assess if appropriate classroom behaviors increased and inappropriate classroom behaviors decreased with the implementation of the Level System as compared to general classroom management strategies already used by the teachers such as verbal reprimands, redirection, yelling, time out, and removal from the class. The Level System consists of a seven level visual chart with the top three levels called the "sunny area" depicting suns with smiling faces, the middle level is considered a "neutral area", and the bottom three levels are the "cloudy levels". Each child is provided a shape with their name written on it and it is placed in the neutral area to start of each reward period. The Level System included the use of contingent rewards, attention, and praise for appropriate behaviors. The children's shapes are moved up or down contingent on engaging in appropriate or inappropriate behaviors. If they are moved up, labeled praise is provided stating the appropriate behavior observed. If a child engaged in inappropriate behaviors, they were given one warning before their shape was moved down. At the end of the reward period, all children that moved their shapes up were provided a reward such as a special snack or activity. The rewards were provided one to two times in the mornings and one to two times in the afternoon. Overall, the effects of the Level System proved to be an effective method for managing classroom behavior. However, at the one-month follow up, none of the teachers were recorded utilizing the intervention procedures. Despite the absence of the intervention procedures,

the teacher's classroom management ratings and skills maintained at follow up with a continued decrease in the use of time outs and the children's appropriately engaged behaviors maintained at high levels. In general, the research is limited assessing maintenance effects in preschool classrooms (Pokorski et al., 2016).

To add to the literature surrounding programmed maintenance for group contingencies, Dadakhodjaeva et al. (2019) conducted an ABC multiple baseline design assessing the effects of the Good Behavior Game (GBG) in middle school classrooms using GBG with Classroom Dojo. This study utilized a partial sequential withdrawal design in which the number of days the intervention was implemented decreased over subsequent phases. The reduced frequency phase was introduced after achieving stability in responding and treatment effects. During this phase, the teachers were instructed to implement the GBG one or two days per week rather than the initial five days per week. The number of times the GBG was implemented and the days in which the game was played was chosen randomly by the researcher, and the researcher notified the teacher the days the game would be played. On days the GBG was not played the teacher notified the students that they were not playing that day, but they might play the next day. The teacher reviewed the rules of the game and behavioral expectations, but did not divide the students into teams or provide rewards or points for rule following. The results of the maintenance phases in this study suggest that reviewing class rules daily might be sufficient for maintenance of improved behaviors.

CHAPTER II – PURPOSE

Emerging evidence suggests that group contingency interventions may be effective in preschool settings. Moreover, novel modifications to group contingency procedures in preschool classrooms such a randomizing when children are able to earn a reward may be effective for improving target students' behavior as well as the behavior of the entire class (Pasqua, 2019; Pasqua et al., 2021). This study will add to the small literature base testing the MSI by replicating treatment effects and systematically evaluating maintenance and generalization of the MSI. More specifically, this study will assess the effects of using a partial sequential withdrawal method to determine if class-wide disruptive behaviors maintain at decreased levels. Moreover, this study was conducted in university-based and community preschool settings, and as a result, will increase the external validity of the findings from Pasqua (2019) and Pasqua et al., (2021).

The following research questions were investigated:

1. Will the Mystery Student Intervention result in increases in class-wide appropriately engaged behavior?
2. Will the Mystery Student Intervention result in decreases in class-wide disruptive behaviors?
3. Will programmed maintenance phases utilizing a partial sequential withdrawal design result in maintained improvements in class-wide appropriately engaged behavior?
4. Will programmed maintenance phases utilizing a partial sequential withdrawal design result in maintained decreases in class-wide disruptive behavior?
5. Will the effects of the MSI generalize to other classroom activities?
6. Will teachers in a preschool setting rate the MSI as socially valid?

CHAPTER III – METHOD

Participants and Setting

The current study included three individual classrooms from university- and community-based preschool settings from the southeastern United States and their teachers. The university-based early classroom was housed within a child development center which consisted of nine classrooms grouped by age, ranging from three months to five years of age. The community based preschool was housed within a day care and after school care which consisted of 10 classrooms. The classrooms were grouped by age, ranging from six weeks to 12 years of age. For both settings, at the time of the study, there was no universal classroom management system in place. The classrooms were recruited based on referrals from administration or the teachers due to three or more children in the classroom displaying regular disruptive behaviors based on the teacher or the administrator's perception. Following the referral, in order to participate in the study: (a) teachers consented to participate in the study (see Appendix A for teacher consent form) and (b) a class-wide observation indicated disruptive behaviors (DB) that occur during 20% or more of the observed intervals during a 20-minute screening observation. Parents did not provide consent for their child's participation because this study did not include data for individual children, only aggregate data for the classroom's behavior. However, passive consent forms were sent home to the families, and they were given the option to have their child not participate in the study.

This project and all consent forms were reviewed and approved by the University of Southern Mississippi's Institutional Review Board (IRB) to ensure the project adhered

to ethical standards, followed federal regulations, and protected the rights of the human subjects involved (see Appendix B for IRB approval).

Classroom A

Classroom A was an early childhood classroom in a university-based child development center. The lead teacher was an African American female. She held a Bachelor's degree and has been teaching for over 20 years. Classroom A consisted of one lead teacher and a teaching assistant. The class was referred for the study due to high levels of disruptive behavior including yelling, leaving their designated area, playing with objects inappropriately, throwing objects, and aggression. There were 13 children in total, 10 males and three females' ages four to five years old. Previous classroom management strategies included the use of redirection and time out.

Classroom B

Classroom B was a community-based classroom. The lead teacher was an African American female. She held a Bachelor's degree and has been teaching for 38 years. Classroom B consisted of one lead teacher and a teaching assistant. The class was referred for the study due to high levels of disruptive behavior including aggression, yelling, playing with objects inappropriately, leaving designated area, and throwing objects. There were 19 children in total, 10 males and nine females' ages four to five years old. Previous classroom management strategies included the use of a conduct chart, redirection, time out, added structure, and office referrals.

Classroom C

Classroom C was a community-based classroom. The lead teacher was an African American female with a high school degree who had one year of teaching experience.

Classroom C consisted of one lead teacher and a teaching assistant. The class was referred for the study due to high levels of disruptive behavior including leaving their designated area, playing with objects inappropriately, yelling, aggression, and throwing objects. There were 18 children in total, 10 males and eight females' ages three to four years old. Previous classroom management strategies included the use of a conduct chart, redirection, time out, added structure, and office referrals.

Instruments and Materials

Class-wide Direct Behavior Rating (C-DBR)

A class-wide DBR was used as a method of measuring classroom behaviors as a whole (see Appendix C). Traditionally, DBR's are used for individual assessment; however, DBRs have been used previously to assess the class's behavior (Pasqua, 2019; Pasqua et al., 2021; Riley-Tillman et al., 2009). The C-DBR consisted of positively worded statements of the behaviors the teachers wanted the children to engage in based on each classrooms referral concerns. Each C-DBR was specific to each classroom and was constructed with each teacher's desired appropriate replacement behaviors.

The Behavior Intervention Rating Scale (BIRS; Elliot & Treuting, 1991)

The BIRS was used in the current study as a method of assessing social validity of the intervention (see Appendix D). Each teacher provided a rating of their perception of the intervention. The BIRS was completed by the teacher for each classroom at the end of the intervention. The BIRS is a 24-item questionnaire, each item is rated from 1 to 6; 1 indicates strongly disagree and 6 indicates strongly agree. The factors measured within the BIRS are acceptability, effectiveness, and time of effectiveness (Elliott & Treuting, 1991). Factor 1, acceptability of the intervention, refers to the teachers' overall

acceptance of the intervention procedures. Factor 2, effectiveness, refers to the effectiveness of the intervention, maintenance, and generalization phases. Factor 3, time to effectiveness, refers to how quickly behaviors improve, and the positive changes associated. All three factors have a combined variance of 73.6 %. Factor 1 accounts for 63% of the variance. Factor 2 accounts for 6% of the variance and factor 3 accounts for 4.3% of the variance. The BIRS has an internal consistency reliability of .97 with the three factors acceptability, effectiveness, and time subscales yielding alphas of .97, .92, and .87 (Elliott & Treuting, 1991).

The Problem Identification Interview Form (PII; Kratochwill & Bergan, 1990)

The PII was utilized for the initial teacher interview (see Appendix E). The PII is commonly used in consultation research due to its adherence to Kratochwill and Bergan's (1990) four stages of behavioral consultation. The goal of the PII is to operationally define the problem. If the problem at hand is not correctly identified and defined, then the intervention may be largely ineffective due to the lack of specificity of the behavioral definition (Andersen et al., 2010). The PII was used to identify the problem behaviors, as well as replacement behaviors the teachers wanted the children to engage in, and to operationally define replacement behaviors.

Materials

MotivAider ©

A MotivAider is an electronic device that is small in size and can be clipped onto clothing with minimal noticeability. This device provides tactile prompts at set intervals. For this study, the MotivAider was set to 90-second intervals. The MotivAider was

utilized in this study by the teacher who was assigned the task of monitoring the "Mystery" students.

Mystery Student Teacher Observation Form

Teachers completed a data sheet for each intervention period as a way of tracking children's behaviors. The teachers coded for appropriately engaged behavior (AEB) and disruptive behavior (DB); these definitions were listed at the bottom of the data sheet. Observations were conducted utilizing a 90-second interval momentary time sampling (MTS) method. The data sheets were used to determine if the Mystery student met their behavior criterion and could be revealed at the end of the intervention period. (see Appendix F).

Prize Box

The prize box was a brightly colored box shaped like a treasure chest. The box consisted of toys and edible rewards, both small and medium in size that were approved by teachers. The prize box was used when the mystery student or students met the criterion for reinforcement.

Name Bag

A bag filled with the names of all students enrolled in the class was used at the beginning of the intervention. The name bag was used to randomly select children to be the Mystery students for that day. Two names were drawn from the bag prior to intervention beginning that day. Names were not re-entered into the bag for that week but were re-entered for the following week.

Dependent Measures and Data Collection Procedures

Dependent Measures

This study consisted of two dependent variables, levels of DB and levels of AEB. The first dependent variable of this study was the level of DB in which the children in each classroom engaged. For DB, each behavior was operationally defined for each classroom during a brief interview following referral between the researcher and teacher. For all three classrooms DB included off task, out of area, inappropriate vocalizations, aggression, and inappropriate use of objects. Off task was defined as the child's attention was directed away from the assigned task or teacher. Out of area was defined as the child had one or more body parts outside of their designated area. Inappropriate vocalizations were defined as the child was engaging in vocalizations without teacher permission such as yelling, talking above a conversational tone, singing, or crying. Aggression was defined as making contact with another person's body by hitting, kicking, scratching, biting, or pinching with the hands, feet, or mouth. Inappropriate use of objects was defined as manipulating an object without teacher permission, engaging with an object that is not related to the task, or using the object inappropriately.

The target behaviors identified were appropriate replacement behaviors for the disruptive behaviors the children engaged in during a specific activity that the teachers identified as having the most disruptions. The second dependent variable was the level of AEB. The definitions of these behaviors encompassed both active and passive engagement. AEB was defined as the child attending to the teacher or being activity involved during the assigned activity, transitioning appropriately in the classroom,

keeping hands and feet to self, using an appropriate voice, and engaging with objects appropriately.

Data Collection Procedures

Observers included graduate students in psychology that were previously trained to conduct the observation procedures in this study and had demonstrated 90% or greater agreement with a previously established observer. Direct observations were conducted to collect AEB and DB data. The observations took place during an activity in which the teachers reported the most disruptive behavior to occur. The duration of observations was determined by the length of the activity and were between 10 and 20 minutes. Momentary Time Sampling (MTS) method was used, and observers recorded the behavior as having occurred if it occurred at the end of the interval. MTS is an accurate time-sampling as it yields results similar to continuous recording methods (Prykanowski et al., 2018). The data collection sheet for the observers was divided into intervals of 10 seconds (see Appendix G). The observers used an audio recording to signal to observe a child at the end of the interval. The observer then recorded if the child was engaging in AEB or DB by circling the initials of the behaviors on the data sheet. Observers used a fixed rotation method to obtain a sample of the class's behavior (Prykanowski et al., 2018). The fixed rotation method consisted of the observers rotating observations of the children in the classroom in a fixed order by looking at a different child at the end of each 10-second interval. Direct observation data were reported as the percentage of intervals in which AEB and DB occurred. One of the target behaviors was recorded for each interval. Observers arrived to the classroom at least 10-minutes prior to beginning an observation, did not interact with children or the teacher, and sat in an unobtrusive location to

minimize reactivity. During each phase, generalization data were collected at randomly selected instructional times in which the MSI was not implemented. These probes were collected in the same manner as the observations during the intervention. This included any time period in which the children were engaged in activities in the classroom.

Experimental Design and Data Analysis

This study included a non-concurrent multiple baseline design across classrooms. The following phases were included in the multiple baseline design: (a) baseline, (b) MSI, and (c) maintenance. The maintenance phase included sequential withdrawal of the MSI. Each phase included at least five data points, with a minimum one-data point stagger between classrooms prior to implementation of the MSI. There were three total classrooms, which allowed for three demonstrations of treatment effect and the researcher controlled manipulation of the independent variable. As a result, the design met design standards for single subject research designs described by Kratochwill et al. (2021) and WWC (2020).

For phase change decisions, class-wide level of DB was the primary dependent variable and phase changes were based on variability within the data. Classrooms that demonstrated high stable, highly variable, or upward trending DB after a minimum of five sessions entered the MSI phase. The second classroom with a high stable or upward trending DB and at least one more baseline data point than the first class that entered MSI then entered MSI. The same decision-making rules were used to determine the third classroom that enter the MSI phase. For the MSI phase, phase changes were made when there was moderate to low levels of DB with minimal to no overlap with baseline and were stable or decreasing; and when that occurred, that classroom entered the

maintenance phase. A minimum of seven data points were collected for each classroom's intervention phase in order to allow for a trend in data to emerge.

Data analysis included visual analysis of graphed data and calculation of effect sizes, which included baseline corrected tau (BCT) and Hedge's g. The researcher visually analyzed data for level, trend, variability, immediacy of effect, consistency of effect, and overlap across adjacent conditions (Horner et al., 2005). BCT (Parker et al., 2011) is a non-parametric effect size that can be used to evaluate overlap across adjacent phases. BCT is able to test for a significant baseline trend in the unintended direction and if a significant trend is detected, then the tau value is corrected for in the calculation. Parker and colleagues (2011) and Vannest and Ninci, (2015) reported effect sizes as very large, large, moderate, and small. A very large effect size is considered .80, a large effect size is considered .60 or greater, a moderate effect size is .20 to .60, and a small effect size is reported as anything below .20 (Parker et al., 2011; Vannest & Ninci, 2015). These effect sizes are based on Tau-U which is a predecessor to BCT (Parker et al., 2011). In addition to BCT, a secondary measure of effect size was calculated in order to provide a weighted effect size utilizing Hedge's *g*. Hedge's *g* is a commonly used standardized mean difference summary measure, interpreted in terms of the number of standard deviations of difference between two groups (e.g. baseline and intervention) (Hedges & Olkin, 1985). Finally, a Pearson's R correlation was calculated to evaluate the strength of the correlation between directly observed levels of AEB and class levels of AEB that the teachers rated. Correlation coefficients of +/- .1 to .3 are considered to have small or weak associations, coefficients yielding +/- .3 to .5 are considered to have medium or

moderate associations, and coefficients yielding +/- .5 to 1.0 are considered large or strong associations (Cohen, 1988).

Procedures

Initial Consultation with Teachers

The researcher met with the teachers via video conference (i.e. Zoom) prior to baseline to identify concerns and operationally define AEB and DB during the PII. Operational definitions informed observation procedures by clarifying the behaviors that were appropriate and disruptive to ensure coding of each behavior was done so with consistency between observers and the teachers when recording Mystery student data and the C-DBR. Subsequently, the researcher trained the teacher to complete C-DBR based on training procedures described by Pasqua (2019). First, the researcher met with the teacher and reviewed the C-DBR form. The researcher explained that the purpose of the C-DBR was to provide an easy method of measuring teachers' perceptions of the occurrence of child behaviors. The researcher explained that the C-DBR is a method of assessing aggregate classroom behavior. The researcher then explained the format of the C-DBR, which consisted of positively worded statements of the behaviors the teacher wanted the children to engage in. Next, the researcher explained the anchors of the C-DBR. The scale of the C-DBR ranged from one to 10, with a score of one indicating that children never engaged in the behavior, five indicating the children occasionally engaged in the behaviors, and ten indicating the children always engaged in the behavior. During the training process, the researcher discussed operational definitions for DB and AEB with the teacher and provided examples and non-examples of the behaviors. Next, the researcher provided the teacher with an opportunity to practice using the C-DBR. For this

portion of the training, the researcher used short videos of children in a classroom setting retrieved from an online database. The teachers observed the children in this video and rated the children's behavior after watching the video. Lastly, the researcher provided feedback to the teacher regarding the use of the Mystery student teacher observation form and C-DBR (see Appendix H).

Finally, the researcher and the teacher collaborated to create a prize box for the children. After collaborating with the teacher on what prizes should be included, the researcher provided the teacher with a list of approximately 20 items that may be included in the prize box, and the teacher selected five to 10 items that were included in the prize box. The researcher created prize boxes and distributed them to teachers on the first day that the intervention was implemented.

Baseline

During the baseline phase, the researcher instructed teachers to conduct class in their typical manner. The observers observed children from an observation room that includes two-way mirror glass. The researcher and the observers did not provide any feedback to teachers or children.

Teacher Training

The researcher trained teachers via video conference to implement the MSI using a behavior skills training approach (Miltenberger et. al., 2004), which includes instructions, modeling, teacher practice, and feedback from the researcher. The researcher provided the teacher with an intervention script that included all the intervention steps and the researcher had a training script that included all of the training steps (see Appendices I and J). The teacher was required to implement the intervention with 100%

integrity during training with the researcher prior to implementation with the class. If the teacher missed any steps, the researcher provided performance feedback and allowed the teacher to practice implementation again. Prior research indicates that teachers complete training for the MSI in one session (Pasqua, 2019; Pasqua et al., 2021). If a teacher's procedural integrity fell below 80%, follow up training was provided. The teachers did not require re-training as procedural integrity of approximately 80% was maintained by each teacher.

Mystery Student Intervention

The classroom teachers were provided the option to choose one of two roles in the intervention: the teacher or the data collector. Classroom A's teachers selected roles and due to staffing constraints for classroom's B and C, the lead teacher assumed the role of both the teacher and data collector. Upon choosing a role, they were then provided additional training in the form of modeling, role-playing, and feedback for intervention implementation and data collection procedures. During this training, the researcher collected procedural fidelity data to ensure that the intervention and all procedures were implemented as planned. The researcher conducted additional fidelity checks during the intervention. At the beginning of the intervention, the researcher provided a script to the teachers and the teacher introduced the MSI to the children. The teacher indicated to the children that the class would playing a game. The teacher indicated the time the game would be played and for how long, described and modeled the appropriate behaviors expected to be shown by the children and encouraged each child to engage in the appropriate behaviors. The teacher provided reminders periodically to the students about the ongoing game.

The teacher, whose role was data collector, drew the names of two target students out of the name bag and recorded them on the teacher's data sheet for observation. The children were allowed to know who the Mystery students were; therefore, observations were conducted in a discrete manner, with the teacher scanning the whole classroom periodically.

The intervention began for the selected activity once all instructions were provided and both teachers were ready. For the intervention period, the teacher who assumed the role of data collector placed the MotivAider© set for 90-second intervals on clothing near their hip and held a clipboard and data sheet and observed the children in class in a specific order in which the Mystery students were observed in every other interval. The data collector recorded a plus sign for AEB or a minus sign for DB on the data sheet. When the MotivAider vibrated at the end of the 90- second interval, the data collector looked at one Mystery student for three seconds and record his or her behavior. After that interval, the data collector recorded the second Mystery student's behavior after the next 90-second interval and continued rotating until the activity period concluded.

At the end of the activity period, the teachers together reviewed the data and determined if the Mystery students could be revealed. In order to reveal the Mystery students, 60% of the observed intervals for a single student must have been scored for AEB. If both students met the criterion, then both students were revealed. If only one Mystery student met the criterion, then only that child’s name was revealed. Upon revealing the Mystery students, the teachers provided examples of the appropriate behaviors that were observed during the selected activity. This was done by restating the operational definition of appropriately engaged behavior. The mystery students were

allowed to select one prize from the prize box for meeting criterion. If both mystery students did not meet criterion, then the teacher announced that both mystery students did not meet the criterion to be revealed and provided examples of the disruptive behaviors that were observed, provided examples of the appropriate behaviors that could have taken place, and proceed to tell the children they can always try again the next day.

Maintenance Phase

This study used a partial sequential withdrawal design modeled after the procedures described by Dadakhodjaeva et al. (2019). That is, during maintenance, the MSI was implemented one to two days per week, with a maximum of two days versus the MSI being implemented four to five days a week for intervention. The maintenance phase began based on stability of data in the MSI phase; therefore, the maintenance phase in any given classroom may not begin on a Monday. As a result, randomly assigning days in which the game is played during a maintenance week that begins on a Wednesday may result in playing the game only once during that week. During the maintenance phase, the researcher randomly chose which days would include the MSI and which days would not and communicated that information to the teacher. On days in which the MSI was implemented, the teacher implemented the MSI in the identical manner as was done during the MSI phase. On days in which the MSI was not implemented the teacher told the class that there would not be Mystery students that day, but that they may play the MSI game again on another day that week. Additionally, the teacher reviewed expected behaviors with the children in an identical manner that was done during the MSI phase. The maintenance phase lasted at least two weeks and data were collected on each day regardless of whether the game was played or not.

Interobserver Agreement

Interobserver agreement (IOA) was collected for at least 20% of the observations by phase and participant, which meets single subject research design standards described by Kratochwill et al. (2021). IOA was calculated using the interval-by-interval method; the number of intervals with agreed upon codes was divided by the total number of intervals and multiplied by 100. If observer IOA fell below 80%, the observers met to review the operational definitions and re-training was conducted.

For classroom A, IOA was recorded for 47.6% of all observations with an average agreement of 99.0% (range = 96.0% - 100%). IOA was collected for 40.0% of baseline observations and agreement averaged 100% (range = 99.0% - 100%). For the treatment phase, 42.8% of observations were coded for IOA and had an average agreement of 99.0% (range= 98.0% -100%). For the maintenance phase, 55.5% of observations were coded for IOA and had an average agreement of 97.0% (range = 96.0% - 100%).

For classroom B, IOA was recorded for 62.5% of all observations with an average agreement of 97.0% (range = 89.0% - 100%). IOA was collected for 27.3% of baseline observations and agreement averaged 92.0% (range = 89.0% - 100 %). For the treatment phase, 41.6% of observations were coded for IOA and had an average agreement of 99.6% (range= 98.0% -100%). For the maintenance phase, 44.4% of observations were coded for IOA and had an average agreement of 97.5% (range = 94.0% - 100%).

For classroom C, IOA was recorded for 30.0% of all observations with an average agreement of 91.4% (range = 78.0% - 100 %). IOA was collected for 33.3% of baseline observations and agreement averaged 92.0% (range = 90.0% - 95.0%). For the treatment phase, 22.2% of observations were coded for IOA and had an average agreement of

99.0% (range= 78.0% -100%). A review of operational definitions and re-training was conducted after one IOA observation fell below the minimum threshold of 80%. For the maintenance phase, 22.2% of observations were coded for IOA and had an average agreement of 89.0% (range = 85.0% - 93.0%).

Procedural Integrity and Treatment Integrity

Observers evaluated procedural integrity for teacher training procedures using a procedural integrity checklist (see Appendix K). The procedural integrity checklist included the following items: (1) the trainer explained the two different roles of teacher involvement, (2) trainer reviewed the script with the teacher, (3) trainer explained the materials used for the intervention, (4) trainer explained operational definitions of AEB, (5) trainer demonstrated how to use of the MotivAider, (6) trainer demonstrated how to complete the data sheet, (7) the trainer modeled the script for the teacher, (8) the trainer role-played the intervention with the teacher, (9) the trainer provided feedback for teacher implementation of the intervention, and (10) the trainer ensured the teacher's understanding of the intervention. Observers completed procedural integrity checks for all training sessions. Procedural integrity for teacher training procedures for classroom A, B, and C were 100%. All teachers were able to implement the procedures with 100% integrity at the end of training.

Observers evaluated treatment integrity for 100% of sessions by phase and participant through the use of treatment integrity checklists. Treatment integrity checklists for baseline, maintenance, and generalization included the following components: (1) the observer will sit in an unobtrusive location in the classroom, (2) the observer will remind the teacher to use their typical classroom behavior management

strategies and to not implement any of the intervention components, and (3) the observer will not provide any feedback to the teacher (see Appendix L).

The treatment integrity checklist for MSI included the following components performed by the teacher: (1) the teacher turned on and programmed the MotivAider, (2) the teacher introduced the game and randomly selected students' names out of a bag, (3) the rules of the game are explained and behavioral expectations are described and modeled by the teacher, (4) the teacher reminded the class of the ongoing game (2-5 times), (5) the teacher completed the data collection sheet, (6) the teacher announced when the game was over and made a reward decision, (7) if criterion was met, the teacher revealed Mystery students, and explained why the Mystery Students earned a reward, (8) the teacher provided rewards if criterion is met, (9) if criterion was not met, the teacher did not reveal Mystery students and explained why the reward could not be given, (10) the teacher reminded class of tomorrow's game, (11) the teacher completes C-DBR (see Appendix M). IOA for treatment integrity was calculated for 20% of sessions in which treatment integrity checks are conducted. IOA for treatment integrity was calculated by dividing the number of components scored as having occurred by the total number of components and multiplied by 100.

For classroom A, IOA was obtained for 40.0% of baseline treatment integrity sessions with 100% agreement for all sessions. Teacher treatment integrity for baseline averaged 100%. For baseline generalization, IOA was recorded for 100% of sessions with 100% agreement. Teacher treatment integrity for baseline generalization averaged 100%. For the MSI, IOA for treatment integrity was recorded for 40.0% of sessions with 100% agreement. Teacher treatment integrity for the MSI averaged 89.0% (range = 66.0% -

100%). Retraining and feedback were provided for one session. For MSI generalization, IOA was recorded for 33.3% of sessions with 100% agreement. Overall treatment integrity for the MSI generalization was 100%. For maintenance, IOA was recorded for 33.3% of sessions with 100% agreement. Teacher treatment integrity for maintenance averaged 96.0% (range = 75.0% - 100%). For maintenance generalization, IOA was recorded for 33.3% of sessions with 100% agreement. Overall maintenance generalization treatment integrity averaged 100%.

For classroom B, IOA was obtained for 27.3% of baseline treatment integrity sessions with 100% agreement for all sessions. Teacher treatment integrity for baseline averaged 100.0%. No IOA was recorded for baseline generalization. However, teacher treatment integrity averaged 100%. For the MSI, IOA for treatment integrity was recorded for 35.7% of sessions with 100% agreement. Teacher treatment integrity for the MSI averaged 87.4% (range = 80.0% - 90.0%). For MSI generalization, IOA was recorded for 33.3% of sessions with 100% agreement. Overall treatment integrity for the MSI generalization was 100%. For maintenance, IOA was recorded for 44.4% of sessions with 100% agreement. Teacher treatment integrity for maintenance averaged 95.6% (range = 80.0% - 100%). For maintenance generalization, IOA was recorded for 33.3% of sessions with 100% agreement. Overall maintenance generalization treatment integrity averaged 100%.

For classroom C, IOA was obtained for 33.3% of baseline treatment integrity sessions with 100% agreement for all sessions. Teacher treatment integrity for baseline averaged 100%. For baseline generalization, IOA was recorded for 50.0% of sessions with 100% agreement. For the MSI, IOA for treatment integrity was recorded for 22.2%

of sessions with 100% agreement. Teacher treatment integrity for the MSI averaged 88.0% (range = 81.0% - 91.0%). For MSI generalization, IOA was recorded for 25.0% of sessions with 100% agreement. Overall treatment integrity for the MSI generalization was 100%. For maintenance, IOA was recorded for 22.2% of sessions with 100% agreement. Teacher treatment integrity for maintenance averaged 96.7% (range = 80.0% - 100 %). For maintenance generalization, IOA was recorded for 25.5% of sessions with 100% agreement. Overall maintenance generalization treatment integrity averaged 100%.

CHAPTER IV – RESULTS

Figure 1 depicts observed levels of AEB and DB for each classroom. Hedge's g and BCT effect sizes for each class are presented in Table 1. Teacher-rated levels of AEB obtained from C-DBR's and observed levels are reported in Table 2. Table 3 depicts the results of the social validity measure, the BIRS, including the factor and total scores for each classroom.

Visual Analysis of Appropriately Engaged and Disruptive Behavior

Classroom A

Figure 1 depicts observed levels of AEB and DB for each classroom. Children in classroom A (top panel) displayed DB at an average of 33.0% (range = 25.0% - 40.0%) intervals observed during baseline. DB data were stable throughout the phase at low moderate levels. AEB was observed to occur at an average of 67.0% (range = 60.0% - 75.0%) of observed intervals in baseline. AEB remained stable at high moderate levels throughout the phase. The mean percentage of DB and AEB in the generalization setting during baseline was 39.0% (range = 30.0% - 48.0%) and 61.0% (range = 52.0%- 70.0%). During the MSI, a decreasing trend for DB was observed from baseline levels immediately. DB was observed to continue to trend downward and stabilize at low levels. The average level of DB observed during treatment was 14.7% (range= 6.0%- 20.0%). The number of intervals observed throughout the treatment phase for DB was lower than levels observed in baseline. During the MSI phase, none of the DB data points were observed to overlap with baseline DB data points. Hedge's g and BCT scores indicated very large to moderate effects for DB (g =3.4 and BCT= 0.3). Levels of AEB increased from baseline with the implementation of the MSI and an increasing trend was observed

until AEB stabilized at high levels. The average levels of AEB observed was 85.2% (range= 80.0% - 90.0%). During the MSI phase, none of the AEB data points were observed to overlap with baseline AEB data points. Hedge's g and BCT scores indicated a very large to moderate effects for AEB (g =3.4 and BCT= -0.3). For generalization, the mean levels of DB and AEB in the generalization setting was 14.3% (range = 0.0% - 23.0%) and 85.6% (range = 77.0% -100.0%). Overall, the MSI demonstrated a consistent, favorable effect for DB and AEB in the target and generalization setting for classroom A.

During maintenance, DB remained stable at low levels with and without the MSI implemented. DB occurred at an average of 8.3% (range = 5.0% - 18.0%) of all intervals observed. DB occurred at an average of 6.3% (range = 5.0% - 8.0%) with the MSI and 9.3% (range = 5.0% - 10.0%) without the MSI and only behavioral expectations stated. The children engaged in AEB during maintenance on average 91.6% (range = 82.0% - 95.0%). AEB occurred at an average of 93.6% (range = 82.0% - 95.0%) with the MSI and 90.6% (range = 90.0% - 95.0%) without the MSI and only behavioral expectations stated. Throughout this phase, AEB remained stable at high levels. Overlap in the data were observed during maintenance when compared to treatment, which is expected when demonstrating maintenance of behavior change following treatment. No overlapping data were observed for maintenance when compared to baseline levels of DB and AEB. Mean percentage of children's display of DB and AEB in the generalization setting during maintenance was 16.0% (range = 6.0% -27.0%) and 84.0% (range = 73.0%- 94.0%). Overall effect sizes for generalization indicated very large to large effects for DB and (g =1.9 and BCT= -0.8) and (g =1 and BCT= 0.8) for AEB. Overall during maintenance, the

partial sequential withdrawal method produced a consistent, favorable effect for DB and AEB in the target and generalization setting.

Classroom B

Children in classroom B (middle panel) displayed DB at an average of 15.0% (range =0% - 51.0%) intervals observed during baseline. DB data were relatively stable at low to moderate levels with some variability. AEB was observed to occur at an average of 85.0% (range = 49.0% - 100.0%) of observed intervals in baseline. AEB was observed at relatively moderate to high levels with some variability. The mean percentage of DB and AEB in the generalization setting during baseline was 7.6% (range= 5.0% -9.0%) and 92.3% (range= 91.0%- 95.0%). During the MSI, DB was observed to increase from baseline levels immediately, followed by a variable descending trend, and then stabilized at low levels. The average level of DB observed during treatment was 17.2% (range= 0%- 37.0%). Overlap in the data was observed for DB during the MSI phase for all data points. However, Hedge's g and BCT scores indicated small to moderate effects for DB (g =0.1 and BCT= 0.3). Levels of AEB decreased from baseline with the implementation of the MSI, and an upward trend was observed until data stabilized at high levels. The average levels of AEB observed was 82.4% (range= 63.0% -100.0%). Overlap in the data was observed for DB during the MSI phase for all data points. However, Hedge's g and BCT scores indicated small to moderate effects for AEB (g =0.1 and BCT= -0.4). For generalization, the mean levels of DB and AEB in the generalization setting was 7.3% (range= 0.0% -20.0%) and 92.6% (range= 80.0% -100.0%). Overall, the MSI demonstrated a consistent, favorable effect for DB and AEB in the target and generalization setting for classroom B.

During maintenance, DB remained variable at a moderate to low levels with and without the MSI implemented. DB occurred at an average of 36.6% (range = 12.0% - 61.0%) of all intervals observed. DB occurred at an average of 37.6% (range = 12.0% - 60.0%) with the MSI and 36.1% (range = 16.0% -61.0%) without the MSI and only behavioral expectations stated. The children engaged in AEB on average, 63.4% (range = 39.0% -88.0%) of the observed intervals. AEB occurred at an average of 63.2% (range = 40.0% - 88.0%) with the MSI and 62.3% (range = 39.0% -85.0%) without the MSI and only behavioral expectations stated. Throughout this phase, AEB was variable at high moderate levels. Mean percentage of children's display of DB and AEB in the generalization setting during maintenance was 7.0% (range = 2.0% -77.0%) and 93.0% (range = 83.0%- 98.0%). Overall effect sizes for generalization indicated small effects for DB and (g =0.03 and BCT= -0.02) and (g =0.03 and BCT= 0.02) for AEB. Overlap in the data was observed for DB during the maintenance for five of the data points, and overlap was observed for all data points for AEB when compared to treatment, which is expected when demonstrating maintenance of behavior change following treatment. Overlap in the data for DB and AEB was observed for maintenance when compared to baseline levels. Overall during maintenance, the partial sequential withdrawal method produced variable effects for DB and AEB in the target and generalization setting.

Classroom C

Children in classroom C (bottom panel) displayed DB at an average of 6.0% (range =0% - 22.0%) intervals observed during baseline. DB data were stable throughout the phase at low moderate levels. AEB was observed to occur at an average of 94.0% (range = 78.0% -100.0%) of observed intervals in baseline. AEB remained stable at

moderate to high levels throughout the baseline phase. The mean percentage of DB and AEB in the generalization setting during baseline was 7.8% (range = 0% -24.0%) and 87.5% (range = 70.0%- 100.0%). During the MSI, DB was observed to decrease from baseline levels immediately and continued at a decreasing trend until stabilizing at low levels. The average level of DB observed during treatment was 5.2% (range= 0%-22.0%). Overlap in the data was observed for DB during the MSI phase for all data points. Hedge's g and BCT scores indicated small effects for DB (g =0.1 and BCT= -0.06). Levels of AEB increased from baseline immediately with the implementation of the MSI, continued to trend upward, and then stabilized at high levels. The average levels of AEB observed was 94.6% (range= 78.0% -100.0%). Overlap in the data was observed for AEB during the MSI phase for all data points. However, Hedge's g and BCT scores indicated a very large to moderate effects for AEB (g =0.08 and BCT= -0.06). For generalization, the mean levels of DB and AEB in the generalization setting was 1.7% (range = 0.0% -3.0%) and 98.2% (range = 97.0% -100.0%). Overall, the MSI demonstrated a consistent, favorable effect for DB and AEB in the target and generalization setting.

During maintenance, DB remained stable at low levels with and without the MSI implemented. An increase in level towards the end of the phase was observed with the reduced frequency of the MSI. It should be noted that for classroom C, maintenance started one week after the conclusion of the MSI due to classroom schedule changes. DB occurred at an average of 10.6% (range =0% - 31.0%) of all intervals observed. DB occurred at an average of 8.0% (range = 3.0% - 15.0%) with the MSI and 12.2% (range = 0% -20.0%) without the MSI and only behavioral expectations stated. The children

engaged in AEB during maintenance on average 89.3% (range = 69.0% - 100%) of the time. AEB occurred at an average of 92.0% (range = 85.0% - 97.0%) with the MSI and 87.8% (range = 69.0% -100.0%) without the MSI and only behavioral expectations stated. Throughout this phase, AEB was high and stable with some variability as the frequency of the MSI was reduced to once per week. Mean percentage of children's display of DB and AEB in the generalization setting during maintenance was 1.7% (range = 2.0% -5.0%) and 98.2% (range = 95.0%- 100.0%). Overall effect sizes for generalization indicated very large to moderate effects for DB and (g =0.8 and BCT= -0.4) and (g =1.1 and BCT= 0.5) for AEB. Overlap in the data was observed for DB during the maintenance for all, but one data point, and overlap was observed for all data points for AEB when compared to treatment, which is expected when demonstrating maintenance of behavior change following treatment. Overall, during maintenance, the partial sequential withdrawal method produced consistent, favorable effects with minor variability for DB and AEB in the target and generalization setting.

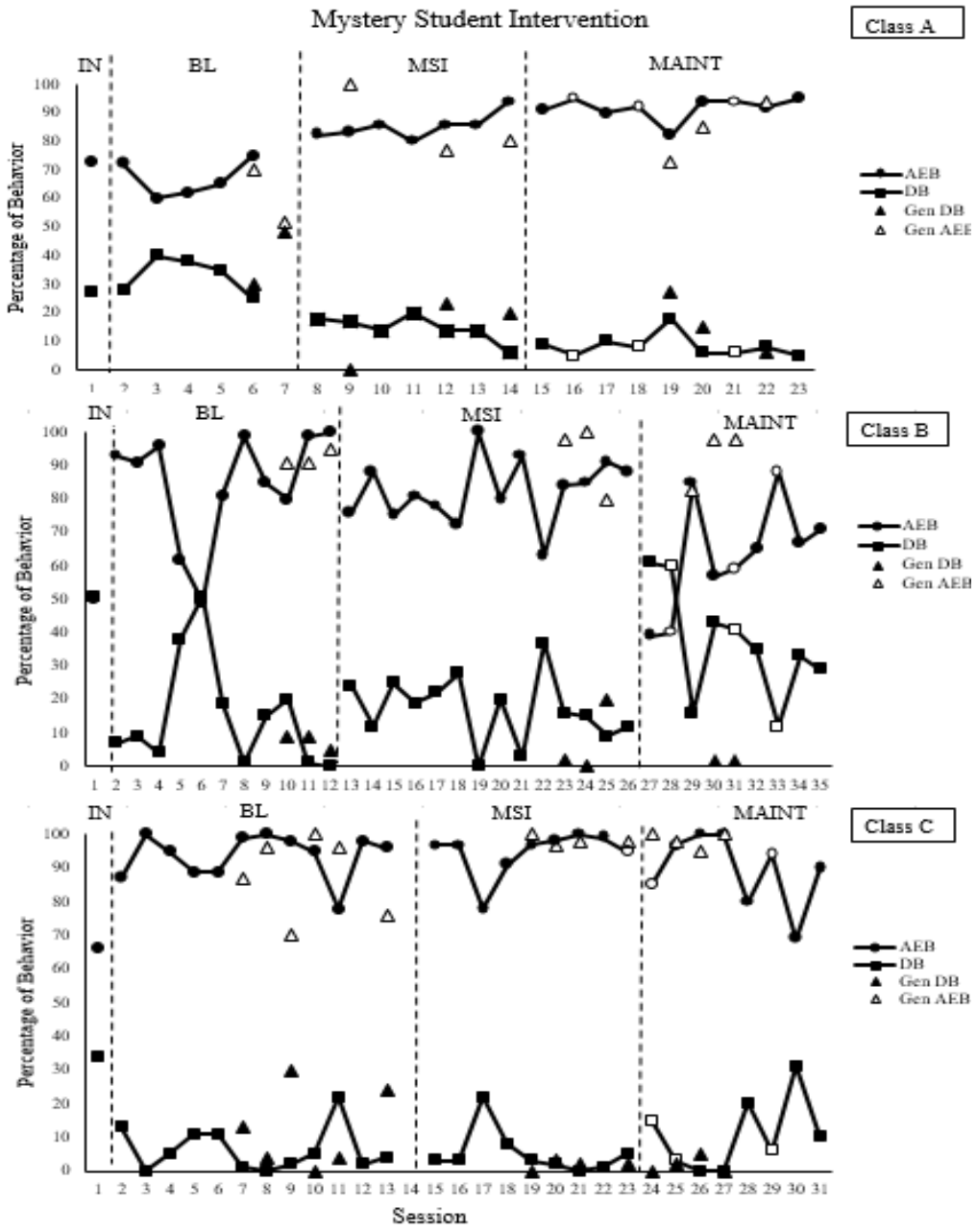


Figure 1. Class-wide Observation Data

Note: Open symbols during MAINT indicate days during which the MSI was not implemented. DB = disruptive behavior; AEB = academic engagement behavior; MSI = Mystery Student Intervention.

Table 1 *Class-wide AEB and DB from baseline to intervention Hedge's g and BCT Effect Sizes*

	Hedge's g				Baseline Corrected Tau (BCT)			
Class	DB	DB-GEN	AEB	AEB-GEN	DB	DB-GEN	AEB	AEB-GEN
A	3.4	1.9	3.4	1.9	0.3	-0.8	-0.3	0.8
B	0.1	0.03	0.1	0.03	0.3	-.02	-0.4	0.2
C	0.1	0.8	0.08	1.1	-0.06	-0.4	0.06	0.5

Note. A very large effect size is considered .80, a large effect size is considered .60 or greater, a moderate effect size is .20 to .60, and a small effect size is reported as anything below .20 (Parker et al., 2011; Vannest & Ninci, 2015)

Correlations between Direct Observations and C-DBR

Table 2 depicts the Pearson's R correlation of teacher-rated AEB and observed levels of AEB for each classroom. The Pearson's R correlation coefficient was computed to assess the relationship between a teacher's ratings of AEB and observed levels of AEB. Teacher ratings of AEB were provided on the C-DBR, the teacher ratings were then calculated by totaling the number of points earned and dividing the total number of points possible and multiplied by 100 to provide a percentage.

Classroom A

For classroom A, teacher-rated levels of AEB were not collected in baseline due to the C-DBRs not being returned due to researcher error. During MSI implementation, teacher ratings were an average of 79.5% (range = 60.0% - 95%) in the target setting. Teacher-rated levels of AEB in the maintenance phase were 96.25% (range= 90.0% - 100%) in the target setting. The overall comparisons of teacher rated levels of AEB and observed levels of AEB yielded a strong positive correlation between the two variables, r = .65, n = 10 and a statistically significant relationship (p = .038). For intervention, the

comparisons of teacher rated levels of AEB and observed levels of AEB yielded a strong positive correlation between the two variables, r = .70, n = 7. However, a statistically significant relationship was not found (p =.118). The comparisons of teacher-rated levels of AEB and observed levels of AEB during maintenance could not be computed due to the sample size being insufficient to calculate a correlation coefficient.

Classroom B

For classroom B, teacher-rated levels of AEB averaged 65.8% (range = 57.5% - 77.5%) in the target setting during baseline. During MSI implementation, teacher ratings increased to an average of 88.2% (range = 62.5% - 100%) in the target setting. Teacher-rated levels of AEB in the maintenance phase were 98.5% (range = 97.5% - 100%) in the target setting. The overall comparisons of teacher rated levels of AEB and observed levels of AEB yielded a weak negative relationship between the two variables, r = -.36, n = 24 and there was not a statistically significant relationship (p = .083) found. Therefore, these results should be interpreted with caution. For intervention, the comparisons of teacher rated levels of AEB and observed levels of AEB yielded a weak positive correlation between the two variables, r = .18, n = 14 and there was not a statistically significant relationship found (p = .528). The comparisons of teacher rated levels of AEB and observed levels of AEB during maintenance yielded a moderate positive correlation between the two variables, r = .42, n = 7 and there was not a statistically significant relationship found (p = .343).

Classroom C

For classroom C, teacher-rated levels of AEB averaged 79.6% (range = 70.0% - 97.5%) in the target setting during baseline. During MSI implementation, teacher ratings

increased to an average of 69.7% (range = 57.5% - 80.0%) in the target setting. Teacher-rated levels of AEB in the maintenance phase were 73.3% (range = 67.5% - 77.5%) in the target setting. The overall comparisons of teacher rated levels of AEB and observed levels of AEB yielded a weak positive correlation between the two variables, r = .15 n = 22 and a statistically significant relationship found (p = .477) was not found. Therefore, these results should be interpreted with caution. For intervention, the comparisons of teacher rated levels of AEB and observed levels of AEB yielded a weak negative correlation between the two variables, r = -.38, n = 9 and there was not statistically significant relationship found (p = .310). However, a statistically significant relationship was not found (p =.118). The comparisons of teacher rated levels of AEB and observed levels of AEB during maintenance yielded a strong positive between the two variables, r = .83, n = 6 and there was statistically significant relationship (p = .039).

Table 2 *Pearson's correlation coefficients for levels of the levels of teacher-rated AEB and AEB directly observed across phases (i.e., C-DBR)*

Classroom	C-DBR
A	.65
B	-.36
C	.15

Note. Correlation coefficients of +/- .1 to .3 are considered to have small or weak associations, coefficients yielding +/- .3 to .5 are considered to have medium or moderate associations, and coefficients yielding +/- .5 to 1.0 are considered large or strong associations (Cohen, 1988).

Social Validity of MSI

Social validity was assessed using the BIRS. Each teacher completed the BIRS to assess the acceptability of the MSI as a classroom intervention. Table 3 depicts the

average scores across each teacher for each factor and total scores. The mean scores were 4.7, 4.1, and 3.3 indicating the MSI was somewhat favored for classroom A and B's teachers and less favored by classroom C's teacher. For the acceptability factor, classroom A and B's teachers rated the highest scores for the MSI, therefore indicating they felt the MSI was an acceptable classroom intervention. However, classroom C's teacher indicated the MSI was not an acceptable classroom intervention. Anecdotal evidence would suggest classroom C's teacher may have not favored the MSI as she was new to the classroom and did not have prior experience with classroom interventions and had limited experience teaching. For the effectiveness factor, all three teachers did not indicate the classroom intervention was effective. In addition, all three teachers indicated the MSI did not produce results soon after implementing the MSI, therefore, providing low ratings for the time of effectiveness factor.

Table 3 *BIRS Factor and Total Scores*

Teacher	Acceptability	Effectiveness	Time of Effectiveness	Total
A	5.4	3.8	3.5	4.7
B	4.4	3.7	3.5	4.1
C	3.6	2.8	2.0	3.3
Total	4.5	3.4	3.0	4.0

Note. Acceptability, refers to overall acceptance of the intervention procedures. Effectiveness, refers to the effectiveness of the intervention, maintenance, and generalization phases. Time of effectiveness, refers to how quickly behaviors improve, and the positive changes associated (Elliot & Treuting, 1991).

CHAPTER V – DISCUSSION

Research shows that early childhood education is positively associated with enhanced social-emotional development, academic achievement, and behavioral development as children age and transition to primary and secondary school (Carter et al., 2010; Love, 2010; Weiland & Yoshikawa, 2014). However, many preschool aged children experience emotional and behavioral difficulties at this age due to a variety of risk factors (Egger & Angold, 2006). As preschool children continue to exhibit behavioral difficulties in the classroom setting, this can interfere with valuable instructional time for many teachers and many teachers reported they do not feel sufficiently trained on classroom management and this can be attributed to an increase in behavior problems within classrooms (Jenson, Reavis, & Rhode, 1998; Siebert, 2005). Fortunately, a vast amount of research has been conducted examining class-wide interventions that are evidence-based and are ideal for teachers because they require limited training, resources, and time and are effective at reducing disruptive behaviors in the classroom (Johnson, Stoner, & Green, 1996). The results of the current study further extends the current literature base surrounding the use of randomized independent group contingency interventions in classroom settings, and specifically to the limited literature base examining the use of group contingency interventions in preschool settings. The discussion of the results for this study are organized by research question, implications of the current study, and limitations.

Research Question 1 and 2

The first and second research questions in this study addressed the effects the MSI would have on AEB and DB in the classroom setting. Visual analysis and evaluation of

effect sizes ranged from small to very large. An immediate increase in AEB was observed for classroom A, with variability in the data for classroom B and C. However, levels of AEB were observed to remain at stable high levels for classroom A and C, and remain at a high level with some variability for classroom B. A functional relationship was demonstrated for classroom A; however, due to high levels of AEB in baseline for classroom's B and C, a conclusion about intervention effects cannot be made due to the limited changes seen between baseline and intervention. Therefore, the results of this study suggest that the MSI was effective at increasing AEB on average from baseline levels for one classroom with null effects seen for classroom's B and C. Moreover, the C-DBR scores provided by the teachers yielded higher scores for intervention than in baseline, therefore, indicating they also perceived an increase in AEB. A functional relationship was demonstrated for classroom A; however, due to low levels of DB in baseline for classroom's B and C, a conclusion about intervention effects cannot be made due to the limited changes seen between baseline and intervention. Therefore, the results of this study also support the notion that the MSI was an effective intervention for decreasing DB in the classroom for one classroom with null effects seen for classroom's B and C due to low levels of DB observed during baseline. DB was observed to decrease immediately for classroom A and C, with increases seen for classroom B. However, DB stabilized to low levels for all three classrooms with some variability. Moreover, the effect sizes for DB ranged from very large to small indicating some effect on class-wide DB. These findings extend the previous MSI literature and other literature assessing the use of group contingency interventions in preschool settings in that the MSI produced small effects on classroom behavior (Pasqua, 2019; Pasqua et.al, 2021).

Research Question 3 and 4

The third and fourth research question addressed the use of programmed maintenance using a partial sequential withdrawal design and if it would result in maintained improvements of AEB and DB. The results of this study support the notion that programmed maintenance is effective at maintaining AEB at high levels for two classrooms and maintaining AEB at moderate levels for the third classroom. The results of this study also support the notion that programmed maintenance is effective at maintaining DB at low levels for two classrooms and maintaining DB at moderate levels for the third classroom. However, these results should be interpreted with caution due to the minimal changes from baseline AEB and DB noted for classroom's B and C. On days in which the MSI was implemented we saw high levels of AEB and subsequent low levels of DB for all three classrooms overall. For two of the three classrooms we saw maintained high rates of AEB and low rates of DB with some minor variability on days in which the MSI was not implemented. The results of the C-DBR suggest that teachers observed changes in the levels of AEB and DB during maintenance, as the C-DBR scores remained high. These findings are consistent with previous literature examining the use of partial sequential withdrawal and indicate that reviewing class rules daily might be sufficient for maintenance of improved behaviors (Dadakhodjaeva et al., 2019).

Research Question 5

The fifth research questions addressed whether or not the effects of the MSI would generalize to other classroom activities or settings in which the intervention was not being implemented. The effects of the MSI were observed to generalize to other activities and settings for all three classrooms. Average levels of DB in the generalization

setting in baseline for all three classrooms were higher than during the MSI. Effect size calculations yielded very large to small effects of the MSI on DB in the generalization settings during intervention. The effects of the MSI were observed to carry over during maintenance observations, yielding a low percentage of intervals in which DB was observed for all three classrooms. These results are consistent with the previous MSI literature and add to the literature that has examined generalized behavior change using group contingencies in preschool settings (Pasqua et.al, 2021). Future research should further assess the generalized effects of group contingency interventions in other school settings and assess the techniques that are most effective at increasing generalization.

Research Question 6

The sixth research question sought out evaluate the degree to which the classroom teachers would rate the MSI as socially valid. The mean scores for acceptability were 5.4, 4.4 and 3.6. The outcomes of the socially validity measure contribute to the literature assessing group contingencies as acceptable classroom interventions (Filcheck, 2004; Ling & Barnett, 2013; Murphy, et. al., 2007; Pasqua, 2019; Pasqua et al., 2021; Pokorski, et al., 2016; Reitman Murphy, Hupp, & O'Callaghan, 2004). These findings are also consistent with previous literature specifically examining the use of the MSI in preschool settings (Pasqua, 2019; Pasqua et al., 2021). While the factor for acceptability may be consistent with previous research, the factors examining effectiveness and time of effectiveness were not rated as high as previous literature. For classroom's B and C, baseline levels of DB were low and levels of AEB were high, which in turn limited the observable treatment effect of the MSI. Therefore, the baseline levels of DB and AEB may have impacted the ratings of effectiveness as the teachers may not have been able to

observe large changes in behavior across phases. It should also be noted that the previous studies were conducted before the COVID-19 pandemic in 2020. Therefore, many classroom settings have been adjusted to meet the Center for Disease Control and Prevention (CDC) guidelines, there is high turnover rate for teachers and support staff leaving the student to teacher ratio higher than expected, and many children have been out of daycares and structured settings for some time. Therefore, the factors for effectiveness and time of effectives should be interpreted with caution as many teachers are still combatting the effects of the pandemic.

It should also be noted that community-based preschools may also lack additional resources for classroom behavior management supports and individual student supports, consequently many students within community-based settings may have undiagnosed behavioral or mental health disorders which may lead to increases in disruptive behavior in the classroom setting. Therefore, impacting the effectiveness and time of effectiveness of the MSI, as some children may need extended exposure to class-wide interventions or more individualized behavioral interventions in order to be successful in the classroom setting.

Implications, Limitations, and Future Directions

The current study extends the literature surrounding the use of the MSI in preschool settings in a variety of ways. Previous literature has examined the use of the MSI using a teacher and a teaching assistant. Due to staffing issues, a teaching assistant was not able to be trained for classroom B and C. Therefore, the current study extends the previous literature by examining the use of the MSI in a classroom with one teacher present during intervention. Based on the high levels of integrity and teacher

acceptability of the MSI, a conclusion can be made that the MSI can feasibly be implemented by one classroom teacher. Future research should continue to examine the use of the MSI using limited classroom personnel.

The use of programmed maintenance with partial sequential withdrawal proves to be efficacious in the preschool settings for maintaining effects of a class-wide intervention. However, research should continue to examine its use in preschool settings to determine an appropriate intervention dosage that allows for continued maintenance of treatment effects. In addition, it may be beneficial to examine the use of the MSI when utilized multiple times per day versus once per day to determine the extent to which the frequency of the MSI impacts classroom behaviors. The study also highlights that some preschool classrooms may benefit from the added structure during periods of the day in which disruptive behaviors are likely to occur. The MSI proves to be an effective classroom management tool and is useful for certain activities during the day. Additional research should examine the use of the MSI across the day and across additional activities.

This study highlights the differences specific locations for preschools may have in terms of resources, classroom structure, and teacher experience. One classroom in the current study was university based and received student, teacher, and classroom supports from within the university on a consistent basis. Two of the three preschools in the current study were community based, meaning they were not tied to a school district and did not have consistent access to outside resources. Classroom structure was remarkably different between the university and community-based settings. This may have been due to staffing difficulties, inside and outside resources for classroom supports, and years of

experience in the classroom. Although, there were marked differences between the two settings, the current study extended the current literature of the MSI by examining its use in two new preschool settings with a variety of differences as compared to Head Start classrooms. The results suggested the MSI is effective in these new preschool settings and can be implemented with high integrity and produce behaviors change even with limited classroom personnel. Future research should continue to assess the effectives of the MSI with modifications specific for each classroom setting.

While the current study examining the use of the MSI in preschool settings resulted positive changes and extended the current literature, it does not go without its limitations. Therefore, results should be interpreted with caution. The first limitation worth mentioning as it pertains to data collection, it that it is unknown if the intervention had an effect on every child's behavior in the classroom, due to aggregate data being utilized. A second limitation that supports the difficulty with data collection methods is the use of the fixed-rotation method. The fixed-rotation method for data collection may have resulted in some disruptive behavior in the classroom to be missed. Due to the observers looking at one specific child at the end of each interval, it is likely that disruptive behaviors were still occurring in the classroom amongst other children. Therefore, it does not accurately portray the level of disruptive behaviors occurring in the classroom at one time. Future research should consider examining target student behaviors as previously done by Pasqua and colleagues (2021) and utilize a more robust observation method such as momentary time sampling for whole class behavior. Additionally, a third limitation is that we do not have data on specific disruptive behaviors that were occurring in the classroom because the data were collected in

aggregate form meaning that disruptive behaviors were recorded as a whole and not recorded by individual behavior. Future research should consider collecting data on specific behaviors to address the level of change in more intense disruptive behaviors versus less intense disruptive behaviors.

A fifth limitation that is worth noting is that we are unable to determine what part of the MSI is responsible for behavior change. Therefore, future research should consider conducting a component analysis to determine which aspects of the MSI are responsible for behaviors change. This could also be extended into the programmed maintenance phase to determine what frequency of the MSI is the most effective and producing and maintaining behavior change and what components of the MSI are responsible for maintaining the behavior change. This may also contribute to the social validity of the intervention by simplifying the MSI to ensure teachers are able to implement the intervention with high integrity and consistency over time.

A sixth limitation worth mentioning is the extent to which the children perceived the intervention is unknown. The current study assessed teacher acceptability, but did not address children's perceptions of the MSI. Anecdotal evidence would suggest the children enjoyed earning prizes for engaging in AEB; however, no conclusions can be reached based on this evidence. Future research should evaluate children's perceptions of the MSI as a socially valid intervention. A seventh limitation worth noting, is the population in which the MSI was implemented with. The population in which the MSI was implemented with leads to concerns with external validity as it has only been evaluated in preschool settings with typically developing children. Future research should

consider evaluating the effectiveness of the MSI with elementary, middle, or high school settings and with those with neurodevelopmental disorders.

The last limitation worth noting is that the data collected during baseline for classroom's B and C portrayed high levels of AEB and low levels of DB. Due to the high levels of AEB observed during baseline, it was difficult to substantially increase AEB and decrease DB; and as a result, the magnitude of treatment effect was limited. High levels of AEB may be attributed to the classroom structure and level of demands placed on the children in baseline which were then altered with the introduction of the MSI. More specifically, during baseline for classroom's B and C, the demands placed on the children were minimal; therefore, resulting in inflated levels of AEB as there were limited opportunities to engage in DB when a clear demand was not placed. These high levels of AEB and low levels of DB seen in baseline indicate a "ceiling effect" and "floor effect" which account for the minimal changes observed from baseline to intervention. Future studies should look at classrooms with more problematic behaviors and the level of demands placed on the children in the classroom.

Several classroom limitations are worth noting, the first being is the lack of consistency in teacher presence in the classroom for classroom B and C. It should be noted that the teachers for classroom B and C left early on certain days of the week, thus leaving the classroom with another teacher who may have been unfamiliar to the children, or the children were sent outside with another class for the remainder of the day. This in turn shifted the children's schedules day to day. In addition, it was noted by observers that the class schedule was not followed consistently. For example, each classroom had set times in which the children would be engaged in certain activities.

However, on many occasions, the classroom schedule had been shifted and the children were engaged in free play or unstructured activities during structured activity times. The last classroom limitation worth mentioning is the lack of classroom support in the classroom for class B and C. Both classrooms had a teacher assistant assigned; however, on many occasions the teacher assistant was not available or was assigned to another duty within the preschool. This severely limited the lead teacher in terms of classroom management, as the teacher to child ratio was high for both classrooms. It should also be noted that on days in which a teacher assistant was not available, the classroom schedule was also altered. The three classroom limitations mentioned above limited the ability to maintain classroom structure on a day to day basis.

In addition to the classroom limitations mentioned, it should also be noted that the community-based preschool demonstrated a significant need for classroom behavior management training and additional outside supports across classrooms. This may have contributed to variability in data, as the classroom teachers were often experiencing high levels of DB outside of the specified intervention and generalization times which may have carried over and impacted teacher ratings on the C-DBR and BIRS. Moreover, due to the lack of classroom management within the classrooms, teachers often placed little demands on the children which also may have impacted the level of DB displayed during baseline for classroom's B and C.

CHAPTER VI – CONCLUSION

The current study provides additional evidence on the use of the MSI in preschool settings (Pasqua, 2019; Pasqua et al., 2021). This study also further extended the use of the MSI by evaluating programmed maintenance using partial sequential withdrawal methods. Results indicate the MSI is effective at reducing DB in a university-based preschool setting with null effects observed for community-based preschool settings. More specifically, the results of the current study indicate the MSI was effective at decreasing DB and subsequently increasing AEB for one of the three classrooms with marginal changes noted in second and third classroom. Behaviors were maintained with the use of programmed maintenance for two of the three classrooms at high levels and moderate levels for the third classroom. Based on the results, the effects of the MSI generalized to other classroom activities in which the intervention was not implemented. Additionally, the use of programmed maintenance using partial sequential withdrawal methods may maintain the effects of the intervention over time. However, the results should be interpreted with caution due to issues with internal validity for two of the three classrooms. All classroom teachers found the MSI to be an acceptable and effective classroom intervention to a degree. Additional evaluation is warranted to further assess the effects of the MSI in novel preschool settings and the use of programmed maintenance while taking into account the limitations set forth in the current study. In sum, this study adds to the limited literature base for group contingency interventions implemented in the preschool setting. Additionally, this study also contributes to the literature surrounding maintenance of interventions and effective withdrawal methods.

APPENDIX A – Teacher Consent Form

Teacher Consent Form

Title of Study: Improving Behaviors in Preschool Classrooms with Mystery Student Intervention Assessment of Effectiveness, Generalization, and Programmed Maintenance

Study Site: The University of Southern Mississippi Center for Child Development

Name of Researcher & University Affiliation: Sarah Litten, M.S. BCBA.; The University of Southern Mississippi

Dear Teacher,

Hello, my name is Sarah Litten, and I am graduate student at University of Southern Mississippi in the School Psychology Doctoral Program. I am currently conducting my thesis, which will assess the effectiveness of a classroom based behavioral intervention. This study is being conducted under the supervision of Dr. Brad Dufrene.

Your classroom has been referred for inclusion in this study for elevated levels of class-wide disruptive behaviors. This intervention aims to reduce class-wide disruptive behavior and aid in classroom management procedures.

Please consider the following when deciding if you will participate in this study:

Purpose:

The purpose of this study is to assess the effectiveness of a class-wide intervention designed to reduce classroom disruptive behavior and subsequently increase appropriately engaged behavior. The Mystery Student Intervention is a classroom management strategy that utilizes reinforcement procedures to produce behavior change in the classroom setting. This study will also assess intervention effectiveness in the form of a rating scale consisting of various questions about components of the intervention.

Procedure:

If you choose to participate in this study, you will be trained to perform a series of classroom management strategies. Prior to intervention implementation, you will complete an initial consultation with the researcher to identify target behaviors and times in which they occur and complete separate trainings on intervention implementation and data collection procedures.

Following the initial consultation, screenings will be conducted to verify that the classroom meets inclusion requirements for the study. During this time, you will be asked to continue your normal classroom management procedures in terms of dealing with disruptive behaviors. If your classroom does not qualify for participation in the study, you will be referred for alternative classroom supports.

If your classroom does qualify for participation, you will be required to attend a training session with me to explain and practice the steps of each intervention. Upon displaying 100% of the steps successfully and being able to accurately describe each step, the intervention phase will begin.

The Mystery Student Intervention (MSI) is an intervention led by the teacher aimed as reducing disruptive behavior in the classroom setting. During the intervention, you will be asked to record data on two randomly selected target students behavior. You will be equipped with a MotivAider® which is a small electronic device that provides a tactile prompt for data collection. When prompted by the MotivAider®, you will look at the target student and record a plus (for appropriate behavior) or a minus (for disruptive behavior) on a data sheet. At the end of the intervention period, you will refer to the data sheet to determine if the students met criteria to earn a reward. The overall goal of the intervention is for all students in the classroom to engage in lower levels of disruptive behaviors and higher levels of appropriately engaged behaviors.

Additionally, trained graduate students from the USM School Psychology program or an undergraduate researcher will conduct classroom observations utilizing a fixed rotation method multiple times per week. The observations will not interfere with the class routine. The observers will be collecting data on disruptive behaviors and appropriately engaged behavior as well as procedural integrity with the implementation of the intervention. You will be provided feedback on intervention implementation as needed.

Benefits:

Agreeing to participate in this study may offer several benefits for you and your students. By participating in this study, you will be trained on the implementation of a new intervention technique that can be used with other students. An additional benefit is the expected decrease in inappropriate behaviors and the increased appropriate behaviors by your students.

Risks:

There appear to be very few risks for either you or your students participating in this study. The greatest discomfort for you may be related to implementing a new procedure in the classroom. To reduce discomfort, I and/or other trained graduate students will provide training, materials, and will be available to answer any questions you may have. Your students should not experience any discomfort from the implementation of the recommended intervention.

Will this information be kept confidential?

Your name and behavior information will be kept confidential. To protect your and the student's privacy, you will be assigned a number. This number will be placed on all paperwork. At no time will any paperwork contain your name. Please note that these records will be held by a state entity and therefore are subject to disclosure if required by law.

Who do I contact with research questions?

If you should have any questions about this research project, please feel free to contact Sarah Litten, M.S. BCBA at 251-554-8108 or Dr. Brad A. Duffrene at 601-266-5256. If you have any questions regarding your rights as a research participant, please feel free to contact the USM Institutional Review Board at 601-255-5509.

What if I do not want to participate?

Please understand that your participation is voluntary, your refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled, and you may discontinue your participation at any time without penalty or loss of benefits.

What if I DO want to participate?

If you would like to participate, please sign the bottom of this sheet. You may keep the second copy for your records.

_____ _____

Participant Signature Date

_____ _____

Investigator Signature Date

APPENDIX B –IRB Approval Letter

Office *of*
Research Integrity

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

118 COLLEGE DRIVE #5116 • HATTIESBURG, MS | 601.266.6756 | WWW.USM.EDU/ORI

NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

The project below has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services regulations (45 CFR Part 46), and University Policy to ensure:

- The risks to subjects are minimized and reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered involving risks to subjects must be reported immediately. Problems should be reported to ORI via the Incident submission on InfoEd IRB.
- The period of approval is twelve months. An application for renewal must be submitted for projects exceeding twelve months.

PROTOCOL NUMBER:	21-362
PROJECT TITLE:	Programming Maintenance for the Mystery Student Intervention
SCHOOL/PROGRAM	School of Psychology
RESEARCHERS:	PI: Sarah Litten Investigators: Litten, Sarah~Dufrene, Brad~
IRB COMMITTEE ACTION:	Approved
CATEGORY:	Expedited Category
PERIOD OF APPROVAL:	07-Dec-2021 to 06-Dec-2022

Donald Sacco

Donald Sacco, Ph.D.
Institutional Review Board Chairperson")

APPENDIX C – Class-wide Daily Behavior Rating

Class-Wide Daily Behavior Rating

Completed by: _____ Date: _____ Target Activity:_____

Instructions: Rate the following behaviors of your class on a scale of 1-10 as a whole for the target activity.

1. **The children kept their hands and feet to themselves.**

Circle the degree to which the children engaged in the behavior.

1	2	3	4	5	6	7	8	9	10
Never				Occasionally					Always

2. **The children were actively involved in the carpet time activity.**

Circle the degree to which the children engaged in the behavior.

1	2	3	4	5	6	7	8	9	10
Never				Occasionally					Always

3. **The children waited to be called on or given permission by the teacher before talking.**

Circle the degree to which the children engaged in the behavior.

1	2	3	4	5	6	7	8	9	10
Never				Occasionally					Always

4. **The children refrained from playing with objects unrelated to the ongoing activity during carpet time.**

Circle the degree to which the children engaged in the behavior.

1	2	3	4	5	6	7	8	9	10
Never				Occasionally					Always

APPENDIX D – Behavior Intervention Rating Scale (BIRS)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
This was an acceptable intervention for the students' problem behavior(s).	1	2	3	4	5	6
Most teachers would find this intervention appropriate for behavior problems in addition to the one(s) described.	1	2	3	4	5	6
The intervention proved effective in changing the students' problem behavior(s).	1	2	3	4	5	6
I would suggest the use of this intervention to other teachers.	1	2	3	4	5	6
The students' behavior problem(s) were severe enough to warrant use of this intervention.	1	2	3	4	5	6
Most teachers would find this intervention suitable for the behavior problem(s) described.	1	2	3	4	5	6
I would be willing to use this in the classroom setting again.	1	2	3	4	5	6
The intervention would *not* result in negative side effects for students.	1	2	3	4	5	6
The intervention would be appropriate intervention for a variety of students.	1	2	3	4	5	6
The intervention is consistent with those I have used I have used in classroom settings.	1	2	3	4	5	6
The intervention was a fair way to handle the students' problem behavior(s).	1	2	3	4	5	6
The intervention is reasonable for the behavior problem(s) described.	1	2	3	4	5	6
I like the procedures used in the intervention.	1	2	3	4	5	6

Adapted from Elliott, S., & Von Brock Treuting, M. (1991).

APPENDIX E – Problem Identification Interview

Problem Identification Interview

Teacher: _____ **Date:** _____

1. What problem behavior(s) commonly occur within the classroom?

2. Operationally define the problem behavior(s).

3. Which behavior causes the most disruption to the class?

4. How often do the behavior(s) occur?

5. Do the behaviors occur more often during a specific activity/ time of day?

6. What methods have you tried previously to correct the behavior(s)?

7. What are some replacement behaviors for the problem behavior(s) described above?

Adapted from Behavioral Consultation in Applied Settings, An Individual Guide (Kratochwill 1990).

APPENDIX F – Mystery Student Teacher Observation Form

Mystery Student Teacher Observation Form

Completed By: _____ Date: _____

Target Activity: _____

Instructions: When prompted by the interval timer every 1.5 minutes, look up at the Mystery students and mark a + if they are engaged in appropriate behaviors and a – if they are engaged in disruptive behavior.

Mystery Student 1:

1	2	3	4	5	6	7	8	9	10

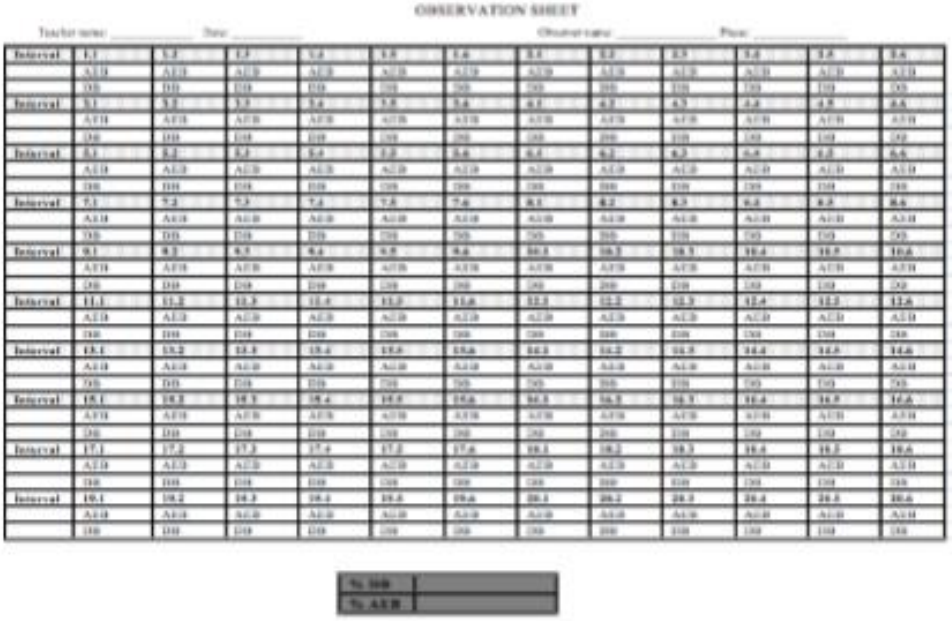
Mystery Student 2:

1	2	3	4	5	6	7	8	9	10

Disruptive Behaviors
Off task behavior: student's attention is directed away from the assigned task or the teacher
Out of area: the student has one or more body parts outside of their designated area. When transitioning centers, the student must go straight to the next area without engaging with any objects or peers along the way, or stopping at any other areas unrelated to the task.
Inappropriate vocalizations: verbalization made without teacher permission such as yelling, talking above typical conversational tone, or crying.
Aggression: making contact with another person's body by hitting, kicking, scratching, pinching with the hands or feet, or biting.
Inappropriate playing with objects: manipulating an object without teacher permission or that is not related to the task, or using the object to strike another person, or throw.

Appropriately Engaged Behavior
Active engagement: student is involved or attending to the teacher during the assigned activity, transitions appropriately in the classroom, keeps hands and feet to self, uses an appropriate voice, engages with objects related to the task, appropriately.

APPENDIX G – Observation Form



OBSERVATION SHEET

Teacher name: _____ Date: _____ Observer name: _____ Place: _____

Interval	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.10	3.11	3.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	4.10	4.11	4.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	5.10	5.11	5.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	6.10	6.11	6.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	7.10	7.11	7.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	8.10	8.11	8.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	9.10	9.11	9.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	10.10	10.11	10.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	11.10	11.11	11.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	12.10	12.11	12.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	13.10	13.11	13.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	14.10	14.11	14.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	15.1	15.2	15.3	15.4	15.5	15.6	15.7	15.8	15.9	15.10	15.11	15.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	16.1	16.2	16.3	16.4	16.5	16.6	16.7	16.8	16.9	16.10	16.11	16.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	17.1	17.2	17.3	17.4	17.5	17.6	17.7	17.8	17.9	17.10	17.11	17.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR
Interval	18.1	18.2	18.3	18.4	18.5	18.6	18.7	18.8	18.9	18.10	18.11	18.12
	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR	OR

% OR	
% A/D	

Disruptive Behaviors

- **Out-of-area:** the student has one or more body parts off of the designated rug or is standing up without teacher permission
- **Inappropriate vocalizations:** verbalization made without teacher permission such as speaking, yelling, humming, singing, and/or whispering
- **Aggression:** making contact with another person's body in a hitting, kicking, pinching, or biting manner with hands, feet or mouth.
- **Playing with Objects:** Manipulating any object without teacher permission or that is not related to the task.

Appropriately Engaged

- Active Engagement: student is involved in or attending to (e.g. looking at or writing on) independent seatwork/carpet time, teacher instruction, designated classroom activities, and/or engaging in task related (permissible) vocalizations with teachers and/or peers.
- Passive Engagement (e.g., listening to the teacher, reading silently, looking at instructional materials).

IOA: _____

APPENDIX H – Procedural Integrity Checklist for DBR Training

Teacher Name: _____ Date:_____ Observer:_____

Procedural Integrity Steps	
The trainer explains what DBR is and why it is useful.	✓ X N/A
The trainer reviews the steps for completing the DBR form with the teacher.	✓ X N/A
The trainer explains the operational definitions of the target behaviors.	✓ X N/A
Steps completed	/
Percentage of Steps completed	

APPENDIX I – Scripted Teacher Instructions

1) Introduction of the "Mystery Student" Game

   "Alright class, today we are going to play a game! The game is called, Mystery Student. This is how it works; everyone's name is in my special bag and I am going to pick two names out of this bag. The two names I pick will be my Mystery Students! Anyone of you can be a Mystery Student, however, I can't tell you who the Mystery Students are. "

2) State game parameter and describe and demonstrate behavioral expectations.

   - "Both Mystery Students' behaviors will be closely watched closely during _____(activity) to make sure they (EXPLAIN & DEMONSTRATE BEHAVIORS).

3) Explain reward procedures.

   - "Both of my Mystery Students will have a chance to pick a reward from the prize box but only if they follow those rules/expectations."

4) Throughout the intervention phase, frequently (2-4 times) remind the class of the ongoing game.

   - "Remember class, I still have two Mystery Students who are being watched for good behavior."

5) End the game and inform the children that it is time to reveal the Mystery Students.

   - "Okay class it is time to reveal the Mystery Student."

6) If criterion is met, reveal Mystery Student(s), give them reward(s), and announce the tomorrow's game.

   - "Great job _____ and _____, you (both) get to pick a prize because you were the Mystery Students and you both exhibited good behavior like, (give behavior examples)"

   - "Tomorrow, everyone will have another chance at being a Mystery Student but remember, everyone should try their best to follow the expectations."

APPENDIX J – Scripted Training Instructions

Scripted Training Instructions

1. There are two different roles for this intervention
    - The "teacher" role
        - "This person will be responsible for telling the rules of the game to the class. This teacher will be responsible for notifying the class when the game is played and providing reminders to the students about the game. This person is also responsible for notifying the class how long the game is played."
    - The "data collector" role
        - "This person is responsible for drawing two students' names from a jar to serve as mystery students and for data collection."
2. Review the teacher script.
3. Explain operational definitions of AEB.
    - Provide examples and non-examples
4. Explain how to use the MotivAider
    - "Turn on the MotivAider and press the arrow to set it to a 90-second interval. It will vibrate to signal the interval has ended."
5. Explain the data sheet
    - "The data sheet consists of two rows (one for each target student) and has operation definitions of AEB and DB at the bottom for your convenience. Place your MotivAider on your hip and set the interval for 90-seconds. When collecting data you will collect in a discrete manner only looking at the students for approximately three seconds. You will then mark a plus if the student is engaging in AEB or a minus of the student is engaging in DB. This observation should last as long as the activity."
6. Model the intervention steps
    - Explain the rules to the class in a neutral tone (i.e. follow teacher script)
7. Role- Play
    - "Next, we are going to practice the intervention!"
8. Provide feedback
    - If the teacher has mastered the training
        - "Great job! You are really getting the hang of this!"
    - If the teacher has not mastered the training
        - "I like how you are XXXX but let's work on XXX."
9. "How are you feeling about this intervention?"
    - "is there anything I can clarify or go over for you again before you implement this on your own?"

APPENDIX K – Procedural Integrity for Intervention Training

Teacher Name: _____ Date: _____

Phase: _____ Observer: _____

Procedural Integrity Steps	
The teacher explained the two different roles of teacher involvement	Y N
The trainer reviewed the teacher script with the teacher(s)	Y N
The trainer presented and explained the materials needed for intervention implementation • Prize Box • Interval timer • Data sheet • Bag of student names	Y N
The trained explained the operational definitions of AEB	Y N
The trained explained and demonstrated how to use the interval timer	Y N
The trained explained and demonstrated how to complete the data sheet	Y N
The trainer modeled the intervention script for the teacher(s)	Y N
The trainer role-played the intervention with the teacher(s), allowing for them to practice implementing the steps of the game	Y N
The trainer provided appropriate feedback contingent upon teachers mistakes during the role-play implementation session	Y N
The trainer ensured the teachers had full understanding of the intervention components.	Y N
Steps completed	/
Percentage of steps completed	

APPENDIX L – Procedural Integrity for Baseline, Maintenance, and Generalization

Teacher Name:_____ Date:_____ Observer:_____

Procedural Integrity Steps	
The observer sits in an unobtrusive location in the classroom.	✓ X N/A
The observer reminds teacher to use typical behavior management strategies and to not implement any components of the intervention.	✓ X N/A
The observer does not provide any feedback to teachers.	✓ X N/A
Steps completed	/
Percentage of Steps completed	

APPENDIX M – Treatment Integrity Checklist for MSI

Teacher Name: _____ Date: _____

Observer: _____ Phase: _____

Intervention Integrity	
1. Interval timer is turned on and programmed to 90 second intervals	Y N NA
2. Teacher introduced the game and randomly selected two names from the name bag	Y N NA
3. Game rules and behavioral expectations were described and modeled	Y N NA
4. Teacher frequently reminded the class of the ongoing game (2-4 times)	Y N NA
5. Teacher completed data collection sheet	Y N NA
6. Teacher announced when the game was over and makes a reward decision	Y N NA
7. If criterion was met, revealed Mystery student(s) and explained why they earned a reward	Y N NA
8. If criterion was met, provided rewards.	Y N NA
9. If criterion is not met, did not reveal Mystery student(s), and explains why reward was not provided	Y N NA
10. Reminds class of tomorrow's game	Y N NA
11. Teacher completed C-DBR	Y N NA
Steps completed	/
Percentage of steps completed	
Teacher require retraining?	

REFERENCES

Andershed, A.-K., & Andershed, H. (2015). Risk and Protective Factors among Preschool Children: Integrating Research and Practice. *Journal of Evidence-Informed Social Work,* 12(4), 412–424. https://doi-org.lynx.lib.usm.edu/10.1080/15433714.2013.866062

Barnett, W. S. (2008). Preschool education and its lasting effects: Research and policy implications. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Feb. 2014.

Bear, G. G, Cavalier, A. R., & Manning, M. A. (2002). Best practices in school discipline. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 977–991). Bethesda, MD: National Association of School Psychologists.

Carter, A. S., Wagmiller, R. J., Gray, S. A. O., McCarthy, K. J., Horwitz, S. M., & Briggs-Gowan, M. J. (2010). Prevalence of DSM-IV disorder in a representative, healthy birth cohort at school entry: Sociodemographic risks and social adaptation. *Journal of the American Academy of Child and Adolescent Psychiatry, 49,* 686–698.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Conners, C. K. (1997). Manual for the Conners' Rating Scales-Revised. Toronto, Ontario, Canada: Multi-Health Systems.

Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). Applied Behavior Analysis (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Dadakhodjaeva, K., Radley, K. C., Tingstrom, D. H., Dufrene, B. A., & Dart, E. H. (2019). Effects of Daily and Reduced Frequency Implementation of the Good Behavior Game in Kindergarten Classrooms. *Behavior Modification*, 0145445519826528. doi:10.1177/0145445519826528

Egger, H. L., & Angold, A. (2006). Common Emotional and Behavioral Disorders in Preschool Children: Presentation, Nosology, and Epidemiology. *Journal of Child Psychology and Psychiatry*, 47(3–4), 313–337.

Ehrlich, S. B., Gwynne, J., Stitziel Pareja, A., Allensworth, E. M., Moore, P., Jagesic, S., & Sorice, E. (2014). Preschool attendance in Chicago Public Schools: Relationships with learning outcomes and reasons for absences: Research Summary. Chicago, IL: University of Chicago Consortium on Chicago School Research. Retrieved from https://eric.ed.gov/?id=ED553158

Elliott, S. N., & Treuting, M. V. B. (1991). The behavior intervention rating scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology*.

Eyberg, S.M. (1988). Parent–Child Interaction Therapy: Integration of traditional and behavioral concerns. *Child and Family Behavior Therapy,* 10, 33– 46.

Filcheck, H. A., McNeil, C. B., Greco, L. A. & Bernard, R. S. (2004). Using a whole class token economy and coaching of teacher skills in a preschool classroom to manage disruptive behavior. *Psychology in the Schools*, 4, 351–361

Filcheck, H. A. (2004). Evaluation of a whole-class token economy to manage disruptive behavior in preschool classrooms (Doctoral dissertation). Retrieved from PsycINFO. (Order No. AAI3132942)

Filcheck, H. A., & McNeil, C. B. (2004). The use of token economies in preschool classrooms: Practical and philosophical concerns. *Journal of Early and Intensive Behavior Intervention*, 1, 94–104.

Gilliam, W. S. (2005). Prekindergartners left behind: Expulsion rates in state prekindergarten systems. *Early Education.* New Haven, CT: Yale University Child Study Center.

Hedges, L., & Olkin, I. (1985). Statistical models for meta-analysis. San Diego, CA: Academic Press

Herman, S. H., & Tramontana, J. (1971). Instructions and group versus individual reinforcement in modifying disruptive group behavior. *Journal of Applied Behavior Analysis*, 4, 113–119.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71, 163–179.

Hunt, B. M. (2013). Using the good behavior game to decrease disruptive behavior while increasing academic engagement with a head start population (Doctoral dissertation). Retrieved from PsycINFO. (Order No. AAI3530731)

Johnson, T. C., Stoner, G., & Green, S. K. (1996). Demonstrating the experimenting society model with class-wide behavior management interventions. *School Psychology Review*, 25(2), 199–214.

Kelshaw-Levering, K., Sterling-Turner, H., Henry, J., & Skinner, C. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the*, 37(6).

Kratochwill, T. R., & Bergan, J. R. (1990). Behavioral consultation in applied settings: An individual guide. *New York, NY: Plenum Press.*

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case design technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kratochwill, T., Horner, R., Levin, J., Machalicek, W, & Ferron, J., & Johnson, A., (2021). Single-case design standards: An update and proposed upgrades. *Journal of School Psychology.* 89. 91-105. 10.1016/j.jsp.2021.10.006.

Ling, S. M., & Barnett, D. W. (2013). Increasing Preschool Student Engagement During Group Learning Activities Using a Group Contingency. *Topics in Early Childhood Special Education,* 33(3), 186–196.

Little, S. G., Akin-Little, A., & O'Neill, K. (2015). Group contingency interventions with children—1980-2010: A meta-analysis. *Behavior Modification,* 39, 322–341.

Love, J. M., & Society for Research on Educational Effectiveness, (2010). Effects of early head start prior to kindergarten entry: The importance of early experience. Society for research on educational effectiveness.

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology,* 50, 625–654.

Maggin, D., Pustejovsky, J. & Johnson, A. (2017). A Meta-Analysis of School-Based Group Contingency Interventions for Students with Challenging Behavior: An

Update. *Remedial and Special Education.* 38. 074193251771690. 10.1177/0741932517716900.

Maus, M. (2007). Independent group contingencies for reducing disruptive behavior in preschoolers with PDD-NOS (Doctoral dissertation). Retrieved from PsycINFO.

Miltenberger R. G., Flessner C., Gatheridge B., Johnson B., Satterlund M., & Egemo K. (2004); Evaluation of behavioral skills training to prevent gun play in children. *Journal of Applied Behavior Analysis.* 37:513–516.

Murphy, K. a., Theodore, L. a., Aloiso, D., Alric-Edwards, J. M., & Hughes, T. L. (2007). Interdependent group contingency and mystery motivators to reduce preschool disruptive behavior. *Psychology in the Schools,* 44(1), 53–63.

Odom, S. L., Chandler, L. K., Ostrosky, M., McConnell, S. R., & Reaney, S. R. (1992). Fading teacher prompts from peer-initiation interventions for young children with disabilities. *Journal of Applied Behavior Analysis*, 25, 307–317.

Pasqua, J (2019), Improving Individual and Group-Wide Behaviors in Head Start Classrooms with Mystery Student Intervention: Assessment of Effectiveness, Generalization, and Maintenance. *Dissertations.* 1660.

Pasqua, J. L., Dufrene, B. A., LaBrot, Z. C., Radley, K., Dart, E. H., & Lown, E. (2021). Evaluating the independent group contingency: "Mystery Student" on improving behaviors in head start classrooms. *Psychology in the Schools*, 1–19. https://doi.org/10.1002/pits.22540

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B, (2011). Combining nonoverlap and trend for single case research: Tau-U. *Behavior Therapy*, 42, 284–299

Pokorski, E. A., Barton, E. E., & Ledford, J. R. (2017). A review of the use of group contingencies in preschool settings. *Topics in Early Childhood Special Education,* 36(4), 230–241.

Prykanowski, D. A., Martinez, J. R., Reichow, B., Conroy, M. A., & Huang, K. (2018). Brief Report: Measurement of Young Children's Engagement and Problem Behavior in Early Childhood Settings. *Behavioral Disorders,* 44(1), 53-62.

Reitman, D., Murphy, M. A., Hupp, S. D. A., & O'Callaghan, P. M. (2004). Behavior change and perceptions of change: Evaluating the effectiveness of a token economy. *Child & Family Behavior Therapy,* 26, 17–36.

Reynolds AJ, Temple JA, Robertson DL, Mann EA. (2001). Long-term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest: A 15-Year Follow-up of Low-Income Children in Public Schools. *Journal of American Medical Association* 285(18):2339–2346. doi:10.1001/jama.285.18.2339

Riley-Tillman, T. C., Methe, S. A., & Weegar, K. (2009). Examining the use of direct behavior rating on formative assessment of class-wide engagement: A case study. *Assessment for Effective Intervention,* 34(4), 224–230.

Rusch, F. R., & Kazdin, A. E. (1981). Toward a methodology of withdrawal designs for the assessment of response maintenance. *Journal of Applied Behavior Analysis,* 14(2), 131–140. doi:10.1901/jaba.1981.14-131

Schweinhart, L. J. (2013). Long-term follow-up of a preschool experiment. *Journal of Experimental Criminology,* 9(4), 389–409. https://doi-org.lynx.lib.usm.edu/10.1007/s11292-013-9190-3

Simonsen, R. J. (1991). Primary, secondary, and tertiary prevention in early intervention. *Journal of Early Intervention,* 15, 124–134

Siebert, C. J. (2005). Promoting preservice teacher's success in classroom management by leveraging a local union's resources: A professional development school initiative. *Education,* 125, 385–392.

Sparrow, S. S., Balia, D. A., & Cicchetti, D. V. (1985). Vineland Adaptive Behavior Scales: Classroom edition. Circle Pines, MN: American Guidance Service.

Swiezy, N. B., Matson, J. L., & Box, P. (1992). The good behavior game: A token reinforcement system for preschoolers. *Child & Family Behavior Therapy,* 14, 21–32.

Temple, J. A., Reynolds, A. J., & Miedel, W. T. (2000). Can Early Intervention Prevent High School Dropout?: Evidence from the Chicago Child-Parent Centers. *Urban Education*, 35(1), 31–56. https://doi.org/10.1177/0042085900351003

Theodore, L., Bray, M., & Kehle, T. (2004). A Comparative Study of Group Contingencies and Randomized Reinforcers to Reduce Disruptive Classroom Behavior. *School Psychology Quarterly*, 19(3), 253–271.

U.S. Department of Education, Office for Civil Rights. (2014). 2013–14 Civil rights data collection. Washington, DC: Author. Retrieved February 1, 2016, from https://crdc. grads360.org/#communities/pdc/documents/5117.

U.S. Department of Education, National Center for Education Statistics. (2021). *The Condition of Education 2021* (NCES 2021-144).

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development,* 93(4), 403–411. https://doi.org/10.1002/jcad.12038

Walker, H.M., & Shinn, M.R. (2010). Systemic, evidence-based approaches for promoting positive student outcomes within a multitier framework: Moving from efficacy to effectiveness. In: M.R. Shinn & H.M. Walker (Eds.), *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 1-26). Bethesda, MD: National Association of School Psychologists.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development,* 84 6, 2112-30.

What Works Clearinghouse (2020). What Works Clearinghouse Procedures and Standards Handbook, Version 4.1. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks.

Whitted, K. (2011). Understanding how social and emotional skill deficits contribute to school failure. *Preventing School Failure,* 55(1), 10–16.

Zeng S, Corr CP, O'Grady C, Guan Y. (2019) Adverse childhood experiences and preschool suspension expulsion: A population study. *Child Abuse Neglect 97:104149.* doi: 10.1016/j.chiabu.2019.104149. Epub 2019 Aug 29. PMID: 31473382

Zinsser, K. M., Zulauf, C. A., Das, N. V., & Silver, C. H. (2019). Utilizing social-emotional learning supports to address teacher stress and preschool expulsion. *Journal of Applied Developmental Psychology*, 61, 33–42. doi: 10.1016/j.appdev.2017.11.006