

**A SYSTEMATIC REVIEW OF GROUP CONTINGENCIES TO DECREASE  
DISRUPTIVE BEHAVIORS IN ADOLESCENT  
AGE STUDENTS**

---

A Thesis  
Submitted to  
the Temple University Graduate Board

---

In Partial Fulfillment  
of the Requirements for the Degree  
MASTER OF SCIENCE IN EDUCATION

---

by  
Krista A. McArdle, B.S.  
August 2023

Thesis Approvals:

Jason Travers, BCBA-D, Thesis Advisor, Applied Behavior Analysis  
Matthew Tincani, Ph.D., BCBA-D, Applied Behavior Analysis  
Art Dowdy, Ph.D., BCBA-D, Applied Behavior Analysis

## **ABSTRACT**

Group contingency interventions are supported by research as an effective way to improve classroom behavior, but studies often have focused on the effects of group contingencies in elementary school settings. This systematic review of the current literature was conducted to investigate the author-reported outcomes and social validity outcomes of group contingency interventions used to decrease disruptive behaviors and increase academic engagement during class instruction for adolescent age students. The methodological rigor of included studies was evaluated. Results of the review of the literature revealed that group contingency interventions decrease challenging disruptive behaviors and increase desired academic engagement for adolescent students, and that studies of this intervention had high methodological rigor. Social validity outcomes indicate group contingency interventions are accepted by both teachers and students. Implications for researchers and professionals are discussed in light of some limitations.

# TABLE OF CONTENTS

	Page
ABSTRACT.....	ii
LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
Chapter	
1. INTRODUCTION.....	1
2. METHOD.....	9
Search Strategy.....	9
Screening and Inclusion Criteria.....	9
Coding and Reliability.....	10
Quality Indicator Coding.....	12
Methodological Quality Indicators.....	12
Determining Methodologically Sound Studies.....	15
Coding Reliability.....	15
3. RESULTS.....	17
Participants.....	17
Setting and Teacher Participants.....	21
Dependent Variables.....	22
Independent Variable, Reinforcement, and Dosage.....	23
Outcomes.....	27
Social Validity Outcomes.....	27
Methodological Quality Indicator Coding.....	30
4. DISCUSSION.....	36
Research Question One.....	36
Research Question Two.....	37
Research Question Three.....	38
Research Question Four.....	38
General Conclusions.....	39
Limitations.....	40
Implications for Research and Practice.....	41

Conclusion .....	44
REFERENCES CITED.....	45

## LIST OF TABLES

Table	Page
1. 1: Study details regarding setting, student participants, and teacher participants .....	18
2. 2: Independent variables, dependent variables and outcomes .....	24
3. 3: Social validity collection methods and outcomes.....	29

## LIST OF FIGURES

Figure	Page
1. 1: Quality Indicator coding results by each indicator and study.....	31

# CHAPTER 1

## INTRODUCTION

Educators are faced with the challenges of creating classroom environments that foster academic engagement while preventing challenging behaviors from occurring (Christ & Christ, 2006). Establishing supportive learning environments for students academically and behaviorally promotes a healthy environment for students to flourish. Academic engagement, both active and passive, are critical to student learning. Students who are demonstrating lower levels of academic engagement are more likely to miss learning opportunities and as a result can face more learning difficulties compared to their peers. Lower levels of academic engagement in students have demonstrated an increase in off task and disruptive behaviors. These disruptive behaviors often result in disciplinary action, most commonly office visits which result in loss of instructional time for educators and students. This loss of instructional time can have detrimental effects on student learning.

Disruptive behaviors can be defined in many ways and can encompass a variety of topographies. Some examples of disruptive behaviors can include but are not limited to vocal and verbal disruptions such as yelling or talking while others are talking, physically aggressive behaviors such as hitting, property disruption and destruction and inappropriate use of technology such as cell phone or tablet use. Disruptive behaviors in classrooms and schools impact the learning process, reduce instruction time, and make it difficult for students to succeed academically (Luiselli, Putnam & Sunderland, 2002). Disruptive behaviors are not limited to a certain age group of students, they can be seen across all ages and grades.

In a study conducted by Solomon and Whaler (1973), the authors discussed how disruptive behaviors can be reinforced by peers in the classroom, which interrupts instruction and impacts the learning processes. In the study the authors evaluated sixth grade students who were considered by the teacher to be disruptive. The disruptive behavior was defined as out of seat behavior, talking, and playing at inappropriate times. Five students who displayed these behaviors frequently were selected as target students. Control students were then selected based on two criteria, the peers had to be of high social reinforcement value to the target students and had to be likely to cooperate with the adult-directed training program. The authors of the study found that the control peers attended to the disruptive behavior displayed by the target students and almost always ignored the prosocial behaviors they displayed. They also found that with adult-directed training and a shift in contingencies some control peers were able to ignore the target students' disruptive behaviors. When the control students were ignoring the disruptive behaviors the target students were less likely to engage in those behaviors and were more likely to engage in prosocial behaviors. The findings in this study can suggest that while challenging behaviors like disruption can be reinforced by peers, with proper contingencies in place peers can be taught to ignore these behaviors and increase prosocial behaviors in their peers. Similarly, Carden Smith and Fowler (1984) looked at peer reinforcement for increasing or maintaining disruptive or maladaptive behaviors of their classmates. In two experiments the authors wanted to determine if both teacher and peer monitored interventions were successful in decreasing disruptive behaviors and increasing prosocial behaviors, they also wanted to determine if peers could administer the peer monitoring intervention without prior adult implementation. The authors found that peers in the kindergarten class were successful in administering the intervention without prior adult implementation and both the adult and peer monitoring



interventions were successful in decreasing disruptive behaviors and increasing prosocial behaviors. Both the above studies demonstrate the effect that peers can have on each other in the classroom when proper contingencies are in place. Peers can influence each other to increase prosocial behaviors and decrease disruptive behaviors.

The aforementioned studies show, disruptive behaviors in the classroom may be maintained by peers but can also be managed by peers when contingencies and interventions are in place. It is imperative to manage disruptive behaviors because adolescent students who display high levels of disruptive behavior are subject to suspensions, and other disciplinary actions which could impact learning because of the lost academic time (Christ & Christ, 2007). Many interventions are used in classrooms to reduce the amount of student disruptions during academic activities. Reducing disruptive behaviors will increase academic learning time, which results in more opportunities to grow academically and socially in a positive learning environment.

The impact that disruptive behaviors have on learning suggests that there is a need for interventions to decrease disruptive behaviors in the classroom and school setting to increase the probability of academic success. There are many different behavior change interventions available with proven success at decreasing a variety of target behaviors. Among those successful interventions is the group contingency. Stage and Quiroz (1997) conducted a meta-analysis comparing individual and group contingencies and determined that group contingencies were more effective than individual contingencies to decrease disruptive behaviors.

According to Litow and Pumroy (1975), group contingencies are the operant techniques applied to the group management of the classroom. A common consequence is contingent upon the behavior of the entire group of people or select individuals of the group. There are many reasons for using group contingencies in applied settings. Group contingencies are beneficial,

cost effective, efficient and practical because the peer group controls the behavior of the classroom (Litow & Pumroy, 1975). Other advantages to using group contingencies include saving time in the administration of consequences because the same consequence is administered to the whole group and group contingencies can be used in situations where an individual contingency may be impractical (Cooper, Heron & Heward, 2007). Litow and Pumroy (1975), list and describe each of the different group contingencies, dependent, independent, and interdependent.

The dependent group contingency requires the same response contingency to be in effect for all group members but is applied only to the performance of select members of the group and the consequence effects the whole group (Litow & Pumroy, 1975). This means that only the selected group members or member must meet criteria for the whole group to receive the consequence. Dependent group contingencies may be useful in decreasing challenging behaviors and increasing prosocial behaviors.

Another group contingency discussed is the independent group contingency. Independent group contingencies are in effect for the whole group but only looks at the performance of the individual. This means that each individual in the group receives reinforcement based on their individual performance in meeting the criteria for the group. The same target behaviors, consequences and contingencies are in effect for all members of the group.

Interdependent group contingencies combine some aspects from dependent and independent group contingencies. Interdependent contingencies are contingencies in effect for the whole group and is applied to the level of the group performance, the group meet the set criterion. This means the group's behavior reinforced when they meet the set criterion. Skinner et al. (1996) discussed the advantages of an interdependent group contingency. The authors

discussed several advantages to interdependent contingencies including shared responsibility for meeting the criterion and delivery of reinforcing consequences without penalizing specific students or their group. The authors also state that interdependent group contingencies can target all individuals within a group, which is efficient for impacting the behavior of all students in the class. Studies have shown that interdependent group contingencies are more effective than dependent and independent (Litow & Pumroy, 1975).

An extension of the interdependent group contingency that is used frequently in classrooms is The Good Behavior Game. The Good Behavior Game (GBG) is used to describe an interdependent group contingency in which a group is divided into two or more teams and the team at the end with the fewest points against them will receive the reward (Cooper, Heron, & Heward, 2019). This game is very commonly used in school and classroom settings as a behavior management technique to decrease disruptive behaviors. The GBG is an evidence-based interdependent group contingency intervention that has been used to effectively reduce disruptive behavior in a variety of settings (Cooper et al., 2019). The GBG involves arranging rules, expectations, and coming up with a set of criteria for the class. Once the class is broken up into teams, points are delivered to the teams that do not meet expectations and the set criterion. The teams with the fewest points will receive the reward. The GBG was used in an early study by Barrish, Saunders and Wolff (1969). In the study the authors were using the GBG to decrease disruptive behaviors such as out of seat and calling out in a fourth-grade classroom. The authors found that the game was successful in decreasing those disruptive behaviors. A more positive variation of the GBG is the Caught Being Good Game (CBGG; Wahl et al., 2016; Wright & McCurdy, 2012). The CBGG awards points to teams for following the classroom rules and the team who meets or exceeds the set criterion receives the reward.

Group contingency interventions have demonstrated success with a wide range of problem behaviors and within diverse populations. The majority of published studies however have been conducted with preschool and elementary age students (Schanding & Sterling-Turner, 2010). Group contingency interventions have been implemented within younger age groups to reduce challenging behaviors such as bedtime compliance (Robinson & Sheridan, 2000), homework completion (Madaus, Kehle, Madaus, & Bray, 2003), inappropriate vocalizations (Davies & Witte, 2000), on-task behaviors (Heering & Wilder, 2006), and social interactions (Kohler et al., 1995). The most reported target behavior for group contingencies within the elementary age population was disruptive behavior (De Martini-Scully, Bray, & Kehle, 2000; Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Lohrmann & Talerico, 2004; Musser, Bray, Kehle, & Jenson, 2001; Hoag, 2007).

Kelshaw-Levering, Sterling-Turner, Henry and Skinner (2000) studied an interdependent group contingency to decrease disruptive behavior in a second-grade classroom. The authors randomized the contingency in two separate intervention phases. In the first, phase the reinforcer was randomized and in the second intervention phase all components were randomized (R-ALL). The authors found that both intervention phases were successful in decreasing disruptive behavior and the R-ALL phase resulting in more stabilized lower percentages. Hoag (2007) found that preschoolers disruptive behaviors could be decreased using the interdependent group contingency in conjunction with mystery motivators. Hoag discusses that the interdependent group contingency paired with mystery motivators was successful in decreasing disruptive behavior.

Group contingency interventions have frequently been investigated in elementary schools, but relatively few researchers have investigated its effects for adolescent students in

secondary schools (Schanding & Sterling-Turner, 2010). Previous reviews on group contingencies from Little et al. (2014) and Maggin et al. (2017) found the majority of studies of group contingencies occurred in elementary and preschool contexts. Maggin et al. (2017) stated only some more recent years have focused on the adolescent population. These findings were consistent with Jenson (1978), who discussed why behavior modification interventions may be less commonly used with the adolescent population. One reason could be that older students may be more motivated by reinforcers not found in school such as money, cars, and clothing. Another possible reason is due to changing schedules and classes, interventions may not be as feasible for educators of adolescents.

There are studies using group contingencies with the adolescent population that have been successful. Target behaviors include increasing social skills behaviors (Hansen and Lignugaris-Kraft, 2005), and academic performance such as peer tutoring (Malone and McLaughlin, 1997). Hansen Lignugaris-Kraft (2005) used a dependent group contingency to increase positive social interactions among nine middle school boys in a self-contained classroom. Before the intervention began the boys received social skills training. During the intervention, student's positive social interactions increased, and negative interactions decreased. Malone and McLaughlin (1997) compared the effects of reciprocal peer tutoring combined with a group contingency to a regular vocabulary program in a middle school class. The students who participated in the peer tutoring plus group contingency had higher quiz scores than those who had not participated.

Evidence-based practice (EBP) is a movement that aims to recognize, distribute, and encourage the utilization of proven research supported methods (Chafouleas et al., 2012). Recently, there has been a movement for more publications to disseminate evidenced-based

behaviors management interventions for the classroom (Chafouleas et al., 2012). Evidenced-based behavior management strategies are not new to the evidence-based practice movement, however. Stage and Quiroz (1997) conducted a meta-analysis related to behavior and school-based interventions and found that group contingencies were some of the most effective.

While many studies have been conducted on the efficacy and acceptability of group contingencies, most of those studies have not been extended to the adolescent population of students until recently. Four questions guided the research:

What were the author-reported outcomes of single case experimental studies investigating group contingencies on disruptive behavior in adolescent students in classrooms?

What are the social validity outcomes of teachers in single case experimental studies that investigated group contingency interventions for disruptive behavior in adolescent student classrooms?

What are the social validity outcomes for adolescent students in single case experimental studies that investigated group contingency interventions for disruptive behavior in adolescent student classrooms?

What is the methodological rigor of single case experimental studies that investigated group contingency interventions for disruptive behavior in adolescent student classrooms?

## CHAPTER 2

### METHOD

#### Search Strategy

Several search methods were used to identify the final list of studies for this review. The first step was to search electronic databases for titles, abstracts, search terms, and key words in ERIC, Academic Search Complete, and APAPsyc Articles. The following search terms were used: “Group Contingency” AND “adolescents” OR teen\* OR “young adult” OR “Good Behavior Game”. Results for each search were compiled into a database for a total of 296 articles. Duplicate articles were deleted for a total of 124 articles from the electronic database search. Next the reference sections of two previous reviews of the literature were conducted (Little et al, 2014; Maggin et al, 2017). This yielded two studies that were not identified by the electronic database search. Last the author completed a hand search of *The Journal of Applied Behavior Analysis*. This journal was chosen for a hand search because the author frequently refers to this journal professionally and because when looking at previous reviews there were articles included in it. The hand search produced one article that was not previously identified.

#### Screening and Inclusion Criteria

First, articles were screened by title and abstract. Articles were excluded if authors clearly did not use a group contingency or include participants who were between ages 11-18 years. Articles that reported pharmacological research, and punishment were excluded. Items that were clearly not peer reviewed articles were also excluded. A total of 37 articles remained after screening and were subjected to full examination according to the inclusion criteria.

Included studies had to be peer reviewed research studies of a group contingency intervention. Second, the participants of the study had to be of adolescent age. Adolescent age was defined as students in grades 6-12 and/or ages 11-18. Third, the research had to be

conducted during classroom instruction in a school. Classrooms included were general education, special education, alternative school classrooms, resource rooms, self-contained classrooms and inclusion classrooms. Fourth, studies were included if authors aimed to decrease disruptive behaviors, increase appropriate academic engagement behaviors, or both. Fifth, studies included students with or without disabilities who demonstrated problem behavior in the classroom. Studies that were excluded from the review were excluded based on student age group, setting, or the target behavior for decrease or increase. The full text assessment resulted in the exclusion of 27 articles resulting in a final list of 13 articles for this review.

### **Coding and Reliability**

A coding manual was developed by the author of this study, and each article was coded for the presence or absence of different details necessary to answer the research questions. When developing the coding matrix, the author created a preliminary list of codes with corresponding definitions, examples, and non-examples. The author added additional codes to this list as needed. Each study was coded for the different characteristics of their participants such as student and teacher age and gender, number of students, years of experience of teacher, socioeconomic background, and the exceptionality of special education students. The settings in which the studies took place were coded based on classroom type such as general education, alternative school setting, special education rooms, resource rooms, self-contained classrooms and inclusion rooms. The dependent variables for each study were coded based on whether authors measured disruptive behavior, academic engagement, or both. If one or both were present, then type of disruptive behavior and/or academic engagement also was coded. Specifically, disruptive behavior was coded based on original author description and included talking out, yelling, using profanity, out of seat, non-compliance, cell phone/technology use,



playing with objects, talking with peers, or other. When other was coded, the author specified the behavior using terms of the original authors in the coding database. Academic engagement included codes for task completion, task accuracy, task latency, and task duration and active and passive participation.

Studies were coded for the presence of the type of group contingency used. Coding was based on original author identification of the IV being independent, dependent, interdependent or a variation of the interdependent group contingency (e.g., Good Behavior Game; Caught Being Good Game). Components of these interventions were coded for the presence or absence of other procedures or interventions being used as part of a treatment package. For example, Christ and Christ (2006) paired an interdependent group contingency with an automated feedback board to provide students with non-contingent feedback throughout the class period. For this study, I coded reinforcer type, length of intervention time during a class period, the frequency of how often the intervention was implemented per day/week, and if core elements of group contingency interventions were followed to protocol as described by Litow and Pumroy (1975). Similarly, Flowers et. al (2014) used prizes and tokens to add to a class jar as reinforcement for following the rules of the Good Behavior Game. The intervention took place across two classrooms for 30 minutes per period.

Outcomes of studies were coded for whether the authors reported a positive, negative, mixed or no effect for each participant (class or group of students). If there was a functional relation present according to the authors, then this was coded as “yes” or “no”. Outcomes for the dependent variables were also coded for a decrease, increase or no change in participant behavior. Lastly social validity outcomes were coded if social validity data were collected and what they used to collect social validity data (ex. interviews, questionnaire, discussion).

## **Quality Indicator Coding**

Gersten et al. (2005), and Horner et al. (2005) introduced standards of identifying evidenced-based practice (EBP) in special education using single case research designs and group comparison designs, respectively. However, Horner et al.'s quality indicators have been significantly improved on. For example, the Council for Exceptional Children (CEC, 2014) introduced updated standards for identifying EBP by evaluating special education research, including single case experimental research. The purpose of identifying the quality indicators essential for methodologically sound intervention studies is to ensure researchers can determine which studies have the methodological features to warrant confidence in their findings (CEC, 2014). Accordingly, articles were appraised to determine the methodological rigor of group contingencies research. The CEC determined that all indicators must be met for a study to be considered methodologically sound. This conservative approach is aimed to ensure that only the most trustworthy studies are considered when classifying the evidence base of practice (CEC, 2014). Defined below are the eight quality indicators identified by the CEC that were coded in this review.

### **Methodological Quality Indicators**

***Context and Setting.*** To meet the criteria for this indicator, the investigators needed to describe features of the context or setting that are relevant to the review (CEC, 2014). This indicator was met if at least one setting or context feature such as school or classroom type was described.

***Participants.*** There are two parts to this indicator. To meet 2.1, investigators needed to describe participant demographics relevant to this review. 2.1 was met if at least one demographic component such as age or gender of students or teachers was reported. To meet 2.2,

the investigators needed to describe the disability of the participants or risk status and method of determination. If no participants were reported as having a disability, then this indicator was coded as not applicable.

***Intervention Agent.*** There were two components of this indicator. To meet 3.1, the investigators needed to describe the role of the intervention agent such as teacher or researcher. To meet 3.2, the investigators needed to report on information on how the intervention agent was trained and how they were checked for understanding.

***Description of Practice.*** Two components were included in this indicator. To meet 4.1, investigators needed to describe and provide enough details to ensure the study could be replicated. In 4.2, the investigators needed to include a description of the materials needed to implement the intervention or provide the references containing this information.

***Implementation fidelity.*** Three components were included in this indicator. For 5.1, investigators needed to assess and report on procedural fidelity using direct and reliable measures. For 5.2, investigators needed to assess and report procedural fidelity related to the exposure of treatment conditions, to be considered met the investigators had to report the length of time of the intervention or how long the intervention was in place.

***Internal validity.*** This indicator has six components that are specific to single subject research designs. To meet 6.1, investigators had to control and systematically manipulate the independent variable and measure treatment fidelity of the intervention. For 6.2, investigators needed to describe baseline for single subject studies or control/comparison conditions for group comparison studies. This would include curriculum, instruction, and interventions. To meet 6.3, control/comparison or baseline condition participants either have no or limited access to the

treatment intervention. 6.5 is for single subject designs only, investigators needed to demonstrate that their design provided at least three demonstrations of experimental effects at least three different times. 6.6 is for single subject research designs and investigators needed to include at least three data points in each baseline condition unless justified by the investigators. This was not applicable to designs that do not require baseline data to be collected (e.g., alternating treatment designs). To meet 6.7, investigators used an accepted single subject research design with procedural integrity.

***Outcome measures/Dependent variables.*** This indicator included six components, five of these are applied to both single subject research designs and group designs. To meet 7.1, investigators needed to discuss social significance of the goals, social appropriateness of the procedures and/or social importance of the effects and/or the measured and reported social validity. For 7.2, investigators needed to define and describe the dependent variable and use a valid measurement system. To meet 7.3, investigators needed to report the effects of the intervention across all outcomes and measures. For 7.4, investigators needed to measure outcomes with appropriate frequency and timing. This means having a minimum of three data points per phase, and at least four repetitions of alternating sequence. For 7.5, investigators needed to prove evidence of inter observer agreement (IOA) by meeting minimal standards. To meet 7.6, investigators reported either adequate validity coefficients or outcomes adequately represented content measured.

***Data Analysis.*** To meet 8.2, single subject research designs needed to clearly present outcome data by using graphs that allowed for visual analysis such as examining level, trend and stability across conditions. To meet 8.3, studies needed to report effect sizes or provide data in which they can be calculated.

## **Determining Methodologically Sound Studies**

According to the CEC (2014), methodologically sound studies are defined by meeting all of the quality indicators across components. Some researchers believe this approach may be too conservative and could present limitations. In 2009, Lane et al. tested these standards and suggested that the rigorous standards could cause the unintended consequence of having too few evidenced based practices for special education to use. As a result, Lane et al. (2009) developed a modified criterion which defined methodologically sound studies as meeting at least 80% out of 100% of all eight quality indicators. Since then, other authors have adopted the modified standards used by Lane et. al. (2009). In 2020, Common et al. used this modified criterion to evaluate the evidence-based practice of student opportunities to respond. Common et al. described using the modification as a weighted criterion based on the logic that rigor exists on a continuum rather than a dichotomy (Common et al., 2020). A database was created to code the articles included in this study for the quality indicators. For the purposes of this review the author adhered to the original guidelines for determining rigor set by the CEC (2014) looking to see which studies met 100% of indicators thus having good methodological rigor.

## **Coding Reliability**

Coding reliability was determined by an outside rater who was trained by the author of this study. The rater has a master's degree in special education and has previous experience with coding articles for a systematic review. To train the outside rater the author of this study created a definitions sheet for the rater to refer to that included important key terms and explanations of procedures included in the articles of this review. The author then modeled the coding procedure for the rater and explained the rationale for each coding decision using an article not included in the review. The researcher and rater completed practice coding using a different article not

included in the review and discussed results. If the rater and author's coding reached 90% agreement or higher, then training was completed. This same procedure was conducted for both the quality indicator coding and for the descriptive coding of the articles. Once training was completed, three articles were randomly selected from the 13 included studies for the rater to code for reliability purposes. Reliability was collected for 25% of articles.

Reliability was determined by comparing each code done by the rater and the investigator and counting the agreements and disagreements. Cooper et al. (2019) describes a formula to calculate the coefficient of agreements. Calculation of agreements included comparing responses to each item of the coding schemes and dividing by the total number of opportunities for agreement. This was calculated for both descriptive and quality indicator coding schemes. Any disagreements between the author and the rater were discussed and resolved. Results of the descriptive coding reliability were 100% for all studies rated. Results of the quality indicator coding was 100% for one of the articles selected and 95% and 90% for the other two randomly selected articles. Disagreements were discussed between the author and the rater until 100% agreement was reached for all articles.

## **CHAPTER 3**

### **RESULTS**

Results are divided into two sections, results of the descriptive evaluation of articles and results of the methodological quality of articles. Descriptive evaluation discusses results found for key details coded for each article. Quality indicator coding discusses the results found for articles for each of the quality indicators.

#### **Participants**

Table one shows the results of the descriptive coding for all student participants and setting. Participants were coded for the number of female and male students in participating classrooms. There were 162 reported male students (34%) and 128 reported female students (27%) out of a total 488 reported students. 59% of articles reported on the number and gender of students participating in their studies (Bohan et al., 2020; Chafouleas et al., 2011; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001; Williamson et al., 2009). 39% of studies did not include information on the number of male and female student participants (Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020). One study (Flower et al., 2014) did not provide any information on the number of student participants, their age or gender. Of the included studies, two occurred in 7<sup>th</sup> grade classrooms (Bohan et al., 2020; Dart et al., 2016) and two occurred in 8<sup>th</sup> grade classrooms (Chafouleas et al., 2011; Dart et al., 2016). Five occurred in 9<sup>th</sup> grade classrooms (Christ & Christ, 2006; Flower et al., 2014; Ford et al., 2020; Kleinman & Saigh, 2011; Mitchell et al., 2015), five in 10<sup>th</sup> grade classrooms (Christ & Christ, 2006; Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015; Williamson et al., 2009), three occurred in 11<sup>th</sup> grade classrooms (Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015), and two occurred in 12<sup>th</sup> grade classrooms (Lum et al., 2019; Mitchell et al., 2015). Ten studies (77%)

Table 1

Study details regarding setting, student participants, and teacher participants.

	Setting	Student Participants					Teacher Participants		
		Number and Gender	Grade Level	Average Age (Range in Years)	Percent of Participants with Disabilities	Participants Socio-economic Status	High-Need School	Number and Gender	Average Years of Experience (Range)
Bohan et al. (2020)	General Ed Classroom	10 Female; 10 Male	Middle School	12.6 (12-14)	0%	Unreported		1 Female	5 years
Chafouleas et al. (2011)	General Ed Classroom	25 Female; 32 Male	Middle School		0%	81% white; 0% black; 13% Hispanic; 6% biracial; 2% Asian	No	2 Female	1 year
Christ & Christ (2006)	General Ed Classroom / Inclusion		High School		0%		No	2 Female; 1 Male	3-15 years
Dart et al. (2016)	General Ed Classroom		Middle School		0%		Yes	2 Female; 1 Male	2.6 years (1-6)
Flower et al. (2014)	Resource Room		High School	15.1	100%		Yes	1 Female	3 years
Ford et al. (2020)	General Ed Classroom		High School		0%	18% white; 82% black; <1% Hispanic	Yes	2 Female; 1 Male	4.6 years (1-14)



Table 1. (continued)

Jones et al. (2019)	Alternative Classroom	3 Female; 3 Male	High School	14-18	50%	66% white; 33% black	Yes	1 Female	5 years
Kleinman & Saigh (2011)	General Ed Classroom	11 Female; 15 Male	High School	15.39	0%	0% white; 23% black; 73% Hispanic; 0% biracial; 0% Asian	Yes	1 Male	
Lum et al. (2019)	General Ed Classroom	28 Female; 44 Male	High School		10%	64% white; 34% black; >1% Hispanic; 0% biracial; 0% Asian	Yes	3 Female	3.3 years (1-6)
Mitchell et al. (2015)	General Ed Classroom	34 Female; 34 Male	High School	15.1 (14-17)	0%	<1% white; 89% black; 3% Hispanic; 5% biracial; 0% Asian	Yes	2 Female; 1 Male	5 years (2-8)
Schanding & Sterling-Turner (2010)	General Ed Classroom	16 Female; 14 Male		14.88 (14-17)	3%			1 Female	3 years
Theodore et al. (2004)	Special Education Room	0 Female; 5 Male			100%	100% white		1 Male	
Williamson et al. (2009)	Special Education Room	1 Female; 5 Male	High School		100%	100% black		1 Female	5 years

reported the grade level of students (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Williamson et al., 2009), and 3 articles (23%) did not report the grade level of students (Jones et al., 2019; Schanding & Sterling-Turner, 2010; Theodore et al., 2001). Of the 13 articles in this review, 6 (46%) reported on the age or age range of students in the class (Bohan et al., 2020; Flower et al., 2014; Jones et al., 2019; Kleinman & Saigh, 2011; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010). Seven studies (54%) did not report on the ages or age range of students (Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Ford et al., 2020; Lum et al., 2019; Theodore et al., 2001; Williamson et al., 2009). Only four (31%) studies reported on both age and grade level of student participants (Bohan et al., 2020; Flower et al., 2014; Kleinman & Saigh, 2011; Mitchell et al., 2015).

Authors of seven (54%) studies included students with disabilities and reported the exceptionality of those students (Christ & Christ, 2006; Flower et al., 2014; Jones et al., 2019; Lum et al., 2019; Schanding & Sterling-Turner, 2010; Theodore et al., 2001; Williamson et al., 2009). A total of 22 (5%) reported participants across all studies included students with disabilities. One study (Flower et al., 2014), reported that 100% of students in the class had a disability, but the number of students was not reported. Authors of two (15%) studies reported students had “Emotional Disturbance” (Theodore et al., 2004; Williamson et al., 2009). Authors of three (23%) studies reported students diagnosed with “Specific Learning Disability” (Flower et al., 2014; Lum et al., 2019; Williamson et al., 2009). Authors of three (23%) studies reported students diagnosed with “ADHD” (Flower et al., 2014; Schanding & Sterling-Turner, 2010; Williamson et al., 2009). Authors of one (8%) study reported students diagnosed with an

“Intellectual Disability” (Flower et al., 2014). Authors of one study (8%) reported students diagnosed with an “Emotional/Behavior Disorder” (Jones et al., 2019), and two studies (15%) reported students diagnosed with “Other Health Impairments” (Lum et al., 2019; Williamson et al., 2009). All authors who reported having student participants with disabilities described the exceptionality of those students. Socioeconomic backgrounds of student participants varied. 70% of the total number of student participants races were reported. Of the reported races of students 35% were Black, 28% white, 5% Hispanic, 1% Biracial, and less than 1% were Asian.

### **Setting and Teacher Participants**

A total of 24 teachers participated in the included studies. Authors of 11 studies reported on the years of experience of teacher participants (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Williamson et al., 2009). All study authors reported the number and gender of teacher participants. Of the 24 teacher participants, six (25%) were male and fourteen (75%) were female. The number of years of experience varied. The total number for all studies of reported years of experience was 69. There were two studies that did not report the years of experience (Kleinman & Saigh, 2011; Theodore et al., 2001).

There were various types of classrooms used as settings in which studies conducted their intervention. Nine (69%) studies occurred in General Education classroom (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010). Two (15%) studies took place in a Special Education classroom (Theodore et al., 2001; Williamson et al., 2009). One study (8%) used an alternative classroom (Jones et al., 2019). Two studies (15%)

occurred in a resource room (Flower et al., 2014; Williamson et al., 2009). One study (8%) occurred in an inclusive classroom (Christ & Christ, 2006) and one study (8%) occurred in a self-contained classroom setting (Theodore et al., 2001).

### **Dependent Variables**

Table 2 shows the descriptive coding results for the dependent and independent variables and outcomes. The dependent variables disruptive behavior and academically engaged behavior were presented in multiple ways across studies. Disruptive behaviors targeted by study authors included talk outs, using profanity, non-compliance, using cellphones or technology, out of seat, playing with objects, talking with peers, being unprepared and other. Academically engaged behaviors targeted by study authors included task completion, task accuracy, task latency, task duration, and active and passive engagement. Disruptive talk out behavior was targeted in nine (69%) studies (Bohan et al., 2020; Christ & Christ, 2006; Dart et al., 2016; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001). Profanity was targeted in one (8%) study (Theodore et al., 2001). Non-compliance was targeted in one (8%) (Theodore et al., 2001). Using cellphones or other technology was targeted in two (15%) studies (Jones et al., 2019; Theodore et al., 2001). Out of seat behavior was targeted in eight (62%) studies (Bohan et al., 2020; Christ & Christ, 2006; Dart et al., 2016; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010). Playing with objects was targeted in three (23%) studies (Bohan et al., 2020; Ford et al., 2020; and Lum et al., 2019). Talking with peers was targeted in three (23%) studies (Bohan et al., 2020; Christ & Christ, 2006; Theodore et al., 2001). Being unprepared was targeted in one (8%) study (Christ & Christ, 2006).

Authors of five studies also measured other target behaviors including off-task behavior, touching peers, and physical aggression (Dart et al., 2016; Flower et al., 2014; Kleinman & Saigh, 2011; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010). Task completion also was targeted in five (38%) studies (Christ & Christ, 2006; Dart et al., 2016; Ford et al., 2020; Lum et al., 2019; Williamson et al., 2009). Active participation was targeted in six (46%) studies (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Lum et al., 2019; Williamson et al., 2009). Passive participation was present targeted in six (46%) studies (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Lum et al., 2019; Williamson et al., 2009). Task accuracy, task latency, and task duration were not targeted by authors of any included studies.

### **Independent Variable, Reinforcement, and Dosage**

Independent Variables were types of group contingencies and included independent group contingency, dependent group contingency, interdependent group contingency, Good Behavior Game, and Caught Being Good Game. The Independent group contingency was targeted in one (15%) study (Lum et al., 2019). The Dependent group contingency was targeted in two (15%) studies (Theodore et al., 2001; Williamson et al., 2009). The Interdependent group contingency was targeted in five (38%) studies (Chafouleas et al., 2011; Christ & Christ, 2006; Jones et al., 2019; Schanding & Sterling-Turner, 2010; Theodore et al., 2004). The Good Behavior Game was targeted in four (31%) studies (Flower et al., 2014; Ford et al., 2020; Kleinman & Saigh, 2011; Mitchell et al., 2015). The Caught Being Good Game was targeted in one (8%) study (Bohan et al., 2020). Some studies used supplemental components in addition to

Table 2

## Independent variable, dependent variables and outcomes

	Independent Variable						Dependent Variables and Outcomes			
	Type	Supp. Comp.	Session Length	Session Freq.	Reinforcers	SR <sup>+</sup> Schedule	Disruptive Behavior Targeted	Target Behavior Outcome	Academic Engagement	Engagement Outcome
Bohan et al. (2020)	CBGG		30 min	5 days per week	"prize"	Weekly	Talk outs, out of seat, talking to peers, playing with objects	Positive (decreased)	Active and passive Participation	Positive (increased)
Chafouleas et al. (2011)	Inter-dependent	Yes (self-mgmt)			prize, edible	Weekly			Active and passive Participation	Positive (increased)
Christ & Christ (2006)	Inter-dependent	Yes (automated feedback)	34		free time	Daily	Talk outs, out of seat, talking to peers, unprepared	Positive (decreased)	task completion, active and passive participation	Positive (increased)
Dart et al. (2016)	Independent		15		free time, edible	Daily	Talk outs, out of seat, other	Positive (decreased)	task completion, active and passive participation	Positive (increased)
Flower et al. (2014)	GBG				prize	Daily	Other (off task)	Positive (decreased)		
Ford et al. (2020)	GBG				edible, homework pass	Daily	Talk outs, out of seat, playing with objects	Positive (decreased)	Task completion, other	Positive (increased)

Table 2. (continued)

Jones et al. (2019)	Inter-dependent		90	3 days per week	free time	Daily	Cell phone/technology use	Positive (decreased)		
Kleinman & Saigh (2011)	GBG			5 days per week	prize, free time, edible	Daily	Talk outs, out of seat, other (physical aggression)	Positive (decreased)		
Lum et al. (2019)	Independent		20		edible, homework pass	Daily	Talk outs, out of seat, playing with objects	Positive (decreased)	task completion, active and passive participation	Positive (increased)
Mitchell et al. (2015)	GBG		20	2-4 days per week	prize, free time, edible, homework pass	Daily	Talk outs, out of seat, other (off task)	Positive (decreased)		
Schanding & Sterling-Turner (2010)	Inter-dependent		15		prize, free time, edible, homework pass	Daily	Talk outs, out of seat, other (off task)	Positive (decreased)		
Theodore et al. (2004)	Dependent/ Inter-dependent	Yes (dependent and interdependent)	45	5 days per week	prize, edible, homework pass	Daily	Talk outs, profanity, non compliance, cell phone/technology use, talking with peers	Positive (decreased)		
Williamson et al. (2009)	Dependent		25		free time, homework pass	Daily			Task completion	Positive (increased)

the group contingency intervention, two studies (15%) used supplemental components (Chafouleas et al., 2011; Christ & Christ, 2006) and one study used two group contingencies (Theodore et al., 2001).

All study authors used all key components of group contingencies in their studies. Reinforcement type and schedule varied across studies. Prizes were used as reinforcement in seven (54%) studies (Bohan et al., 2020; Chafouleas et al., 2011; Flower et al., 2014; Kleinman & Saigh, 2011; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001). Free time was used in seven (54%) studies (Christ & Christ, 2006; Dart et al., 2016; Jones et al., 2019; Kleinman & Saigh, 2011; Mitchell et al., 2015; Schanding & Sterling-Turner 2010; Williamson et al., 2009). Edibles were used in eight (62%) studies (Chafouleas et al., 2011; Dart et al., 2016; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001). Homework passes were used in six (46%) studies (Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001; Williamson et al., 2009). Reinforcement schedules were either immediate (daily) or delayed (weekly). Daily reinforcement was provided in 11 (85%) studies (Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh., 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001; Williamson et al., 2009), and weekly reinforcement was provided in four (31%) studies (Bohan et al., 2020; Chafouleas et al., 2011; Flower et al., 2014; Kleinman & Saigh, 2011). There were two studies that provided daily and weekly reinforcement (Flower et al., 2014; Kleinman & Saigh, 2011).

There were four (31%) studies that did not include session length (Chafouleas et al., 2011; Flower et al., 2014; Ford et al., 2020; Kleinman & Saigh, 2011). When reported, session



lengths varied depending on type of classroom and period lengths. Some studies had 20-minute-long sessions (Lum et al., 2019; Mitchell et al., 2015), while others had 90-minute-long sessions (Jones et al., 2019). The average session length was 32 minutes. There were five (38%) studies who reported how often throughout the week the intervention was implemented (Bohan et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Mitchell et al., 2015; Theodore et al., 2001). There were seven (54%) studies who did not report on session frequency throughout the week (Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014). The average number of days per week the intervention was implemented was three days per week.

### **Outcomes**

The original author reported intervention effects on students and teachers as well as changes in the dependent variables for students were coded. Authors of 11 studies (85%) reported the intervention had a positive effect on student behavior (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh., 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding & Sterling-Turner, 2010; Theodore et al., 2001) and authors of two studies (15%) reported a mixed effect on student behavior (Dart et al., 2016; Williamson et al., 2009). Of the 11 studies that targeted decreases in disruptive behaviors, all reported reduction in disruption following intervention. Authors of seven studies targeted increasing academically engaged behaviors, and all reported these behaviors increased following intervention.

### **Social Validity Outcomes**

Table three shows the descriptive coding results for student and teacher social validity and their outcomes. Social validity data for students and teachers were coded. There were seven studies (54%) that collected student social validity data (Bohan et al., 2020; Flower et al., 2014;

Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001) and 10 studies (77%) that collected teacher social validity data (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2004). Of the studies that collected social validity, six studies (46%) solicited teachers and students (Bohan et al., 2020; Flower et al., 2014; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001). There were two studies (15%) that did not collect social validity data (Schanding & Sterling-Turner, 2010; Williamson et al., 2009). The methods used to collect student social validity included interviews and survey/questionnaires. Authors of one study (8%) conducted interviews (Flower et al., 2014) and six studies (46%) used survey/questionnaires for students (Bohan et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001). To collect teacher social validity, authors of nine studies (69%) used survey/questionnaires (Bohan et al., 2020; Chafouleas et al., 2011; Christ & Christ, 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001) and two (15%) used maintenance data to determine social validity (Flower et al., 2014; Kleinman & Saigh, 2011). Authors of four (57%) studies had student social validity results that indicated the intervention was valuable (Flower et al., 2014; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001) and three (42%) had mixed results (Bohan et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011). Teacher outcomes showed that 8 (80%) studies included social validity measures where participants rated their experience as positive (Bohan et al., 2020; Christ & Christ, 2006; Flower et al., 2014; Ford et al., 2020; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001) and 2 (20%) had mixed ratings (Chafouleas et al., 2011; Dart et al., 2016).

Table 3

Social Validity collection methods and outcomes.

	Student Participants			Teacher Participants		
	Reported	Instrument	Outcomes	Reported	Instrument	Outcomes
Bohan et al. (2020)	Yes	survey/ questionnaire	Mixed	Yes	survey/ questionnaire	Positive
Chafouleas et al. (2011)	No			Yes	survey/ questionnaire	Mixed
Christ & Christ (2006)	No			Yes	survey/ questionnaire	Positive
Dart et al. (2016)	No			Yes	survey/ questionnaire	Mixed
Flower et al. (2014)	Yes	Interview	Positive	Yes	survey/ questionnaire	Positive
Ford et al. (2020)	No			Yes	survey/ questionnaire	Positive
Jones et al. (2019)	Yes	survey/ questionnaire	Mixed	No		
Kleinman & Saigh (2011)	Yes	survey/ questionnaire	Mixed	Yes	Other (Maintenance)	Positive
Lum et al. (2019)	Yes	survey/ questionnaire	Positive	Yes	survey/ questionnaire	Positive
Mitchell et al. (2015)	Yes	survey/ questionnaire	Positive	Yes	survey/ questionnaire	Positive
Schanding & Sterling-Turner (2010)						
Theodore et al. (2004)	Yes	survey/ questionnaire	Positive	Yes	survey/ questionnaire	Positive
Williamson et al. (2009)	No			No		

Note. SV = Social validity. Mixed = indicates positive and negative social validity results were reported.

## **Methodological Quality Indicator Coding**

**Context and Setting.** Figure 1 shows the results of the quality indicator coding. The CEC (2014) requires studies to provide details about the settings in which studies are taking place. Studies must describe critical features of the context or setting such as the classroom type (eg. general education, inclusion, special education), school type (eg. public, private, alternative), geographical location, or physical description of the location. All 13 (100%) of the studies included in the review met this indicator by describing at least one detail about the setting in which their research took place.

**Participants.** The CEC (2014) requires participants of studies to be described relevant to the article such as age, gender, ethnicity, socioeconomic status, and language. If participants had a disability (specific learning disability, autism spectrum disorder, behavior problem), the disability needs to be described and the method for determining disability status needs to be described. All 13 (100%) of the studies in this review met the first component of this quality indicator by describing key details about the participants included in their study. This included details about the demographics of the student and teacher participants. Authors of six (46%) studies included participants with disabilities (Flower et al., 2014; Jones et al., 2019; Lum et al., 2019; Schanding et al., 2010; Theodore et al., 2001; Williamson et al., 2009). Of those six studies, all met the criteria for describing disability and method of determining disability status.

**Intervention Agent.** This indicator had two components. First, the role of the intervention agent needs to be described (i.e., teacher, paraprofessional, researcher) and second, when relevant race/ethnicity, and education/licensure background should be reported. All 13 (100%) of the articles included in the review met this criterion by stating who carried out the intervention and what their experience/background was. The second component of this indicator described

Figure 1

Quality Indicator coding results by each indicator and study.

	1.1	2.1	2.2	3.1	3.2	4.1	4.2	5.1	5.2	5.3	6.1	6.2	6.3	6.5	6.6	6.7	7.1	7.2	7.3	7.4	7.5	8.2	
Bohan et al. (2020)			X																				
Chafouleas et al. (2011)			X																				
Christ & Christ (2006)			X																				
Dart et al. (2016)			X																				
Flower et al. (2014)																							
Ford et al. (2020)			X																				
Jones et al. (2019)																							
Kleinman & Saigh (2011)			X																				
Lum et al. (2019)																							
Mitchell et al. (2015)			X																				
Schanding & Sterling-Turner (2010)																							
Theodore et al. (2004)																							
Williamson et al. (2009)																							

how the intervention agent was trained, and the procedures used to train the intervention agent if they were someone other than the researcher/author of the study. Only seven (54%) studies included information about how the intervention agents were trained or if training was provided (Bohan et al., 2020; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015; Schanding et al., 2010).

***Description of Practice.*** This indicator has two components concerning the replicability of the intervention. For the first component the study must describe detailed intervention procedures and the intervention agents' actions. 12 studies 92% adequately described intervention procedures to allow for replication (Bohan et al., 2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding et al., 2010; Theodore et al., 2001; Williamson et al., 2009). The second component describes all materials needed for the intervention with enough detail to allow for replication. Authors of 10 studies (76%) provided adequate details regarding materials needed to complete the intervention (Bohan et al., 2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015).

***Implementation Fidelity.*** This indicator consists of three components regarding treatment integrity data collection. The first component states that the study must collect treatment integrity using a direct and reliable measure. Authors of seven (54%) studies reported using a direct and reliable measure to collect treatment integrity data (Bohan et al., 2020; Chafouleas et al., 2011; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015). Components two and three describe the dosage in which treatment integrity data were collected. Authors of five of studies met component two (Chafouleas et al., 2011; Dart et al., 2016; Flower

et al., 2014; Lum et al., 2019; Mitchell et al., 2015) and six met component three (Chafouleas et al., 2011; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Lum et al., 2019; Mitchell et al., 2015).

***Internal Validity.*** The CEC (2014) requires that studies demonstrate high levels of internal validity and experimental control to be considered a good evidenced based practice. There are 6 components of this indicator relating to single subject research. Component 1 states that the researcher must control and systematically manipulate the independent variable. Authors of 12 (92%) studies met this criterion (Bohan et al., 2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding et al., 2010; Theodore et al., 2001; Williamson et al., 2009). The second component requires description of baseline data, specifically whether it was defined, how long it was, and the duration and frequency in which it was collected. 12 (92%) studies sufficiently met this component (Bohan et al., 2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001; Williamson et al., 2009). Component three states the participants have no or limited access to the treatment intervention during the control or baseline conditions. 12 (92%) studies met this component (Bohan et al., 2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001; Williamson et al., 2009). The next component states that the studies design must provide at least three demonstrations of experimental effects at different times. All 13 (100%) studies met this criterion. The sixth component states that for each baseline phase there must be at least three data points, 12 (92%) studies met this criteria (Bohan et al.,

2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Schanding et al., 2010; Theodore et al., 2001; Williamson et al., 2009). The last component's criteria states that the study's design must control for threats to internal validity. 100% of studies met this criterion by properly executing reversals, and multiple baselines.

***Outcome Measures/Dependent Variables.*** This quality indicator describes the outcome measures to gauge the effect on the study outcomes. The first component states that outcomes must be socially important, they must be linked to an improved quality of life or an important learning or developmental outcome. All studies met this criterion. The second component states the study must clearly define and describe measurement of the dependent variable, 100% of the studies met this criterion. For 7.3 studies must describe and report the effects of the intervention on all measures of the outcomes targeted, even if it is not a positive effect. All 13 (100%) studies met this criterion by describing their results, even if it was not a positive effect. For 7.4 investigators needed to measure outcomes with appropriate frequency and timing. This means having a minimum of three data points per phase, and at least four repetitions of alternating sequence, 12 (92%) studies met this criterion (Bohan et al., 2020; Chafouleas et al., 2011; Christ et al., 2006; Dart et al., 2016; Flower et al., 2014; Ford et al., 2020; Jones et al., 2019; Kleinman & Saigh, 2011; Lum et al., 2019; Mitchell et al., 2015; Theodore et al., 2001; Williamson et al., 2009). For 7.5, investigators needed to provide evidence of interobserver agreement (IOA) by meeting minimal standards. Authors of all studies met this criterion by collecting and reporting IOA data.

***Data Analysis.*** To meet 8.2, single subject research designs needed to clearly present outcome data by using graphs that allowed for visual analysis such as examining level, trend and



stability across conditions. All 13 articles met this component by including graphs of data using a single subject design. There were four studies that met all quality indicators (Dart et al., 2016; Flower et al., 2014; Lum et al., 2019; Mitchell et al., 2015). Bohan et al. (2020) met 19 (91%) quality indicators. They did not report treatment integrity procedures. Chafouleas et al. (2011) met 20 (95%) quality indicators. They did not report the training of the intervention agent. Christ and Christ (2006) met 17 (82%) of quality indicators. They did not report the training of the intervention agent or report on treatment integrity procedures. Dart et al. (2016) met 21 (100%) of applicable quality indicators. Flower et al. (2014) met 22 (100%) of quality indicators. Ford et al. (2020) met 18 (82%) quality indicators. They did not provide enough details for replication and did not report treatment integrity procedures. Jones et al. (2019) met 18 (82%) quality indicators. They did not report on the training of the intervention agent or on treatment integrity. Kleinman & Saigh (2011) met 16 (76%) quality indicators. They did not report on the intervention agent or how they were trained. Treatment integrity was also not reported. Lum et al. met 21 (100%) quality indicators. Mitchell et al. (2015) met 21 (100%) quality indicators. Schanding and Sterling-Turner (2010) met 15 (68%) of quality indicators. They did not provide enough description for replication, treatment integrity was not described or reported, and baseline conditions were not described. Theodore et al. (2001) met 17 (77%) of quality indicators. They did not report the training of the intervention agent and treatment integrity was not reported on. Williamson et al. (2009) met 17 (77%) of quality indicators. They did not describe the training procedures of the intervention agent and did not report treatment integrity.

## **CHAPTER 4**

### **DISCUSSION**

Group Contingencies typically involve establishing clear rules and delivering or withholding reinforcers based on targeted behavior (Cooper, Heron & Heward., 2019). Authors of previous reviews of group contingency interventions (Little et al., 2015; Maggin et al., 2017) found most studies included elementary school students. It was not until recent years, 2010 and later, that research began expanding to include more studies with adolescent age students in high school and middle school (Maggin et al., 2017). The lack of research on group contingencies with adolescent populations prompted this review to determine if group contingencies are an effective and socially acceptable intervention to use for this population to address the common school issue of disruptive behaviors and academic engagement.

#### **Research Question One**

The results of this review revealed that all group contingency interventions effectively reduced disruptive behaviors and improved academic engagement based on original author reports. The original author-reported outcomes of group contingency interventions on disruptive and academically engaged student behavior suggest that group contingencies reliably decrease disruptive behavior and increase academic engagement. Findings in the included studies suggest that using any of the group contingency interventions could result in positive effects for students and teachers. For example, Williamson et al. (2009) used a dependent group contingency to increase academic engagement. They reported positive effects indicated by an increase in passive and active student engagement. Also, Dart et al. (2016) implemented an independent group contingency to decrease disruptive behaviors like talk outs and out of seat and increase academic engagement (e.g., task completion and active and passive participation). The authors of this

study reported decreases in disruption and increases in academic engagement. Also, Schanding & Sterling-Turner (2010), implemented an interdependent group contingency to decrease disruptive behaviors such as talk outs and out-of-seat behavior in high school students. Researchers reported decreases in disruptive behaviors after the implementation of the interdependent group contingency.

### **Research Question Two**

Social Validity outcomes of studies in this review revealed that most teachers and students rated the interventions positively. Most educators rated the group contingency they implemented as having a positive effect and rated the intervention as favorable. Two studies reported maintenance data collection after the completion of the study as indicators of social validity (Kleinman et al., 2011 & Flowers et al., 2014). In the Flowers et al., (2014) study researchers found that the teacher participant did not continue to use the intervention after the completion of the study. The teacher gave positive ratings of the intervention and made comments that they would continue to use it, but at the three week follow up maintenance session, it was discontinued. When maintenance data were collected researchers reported student disruption was back to baseline levels. It was unclear why the teacher discontinued the intervention. The teacher in Kleinman et al., (2011) continued using the GBG when maintenance data were collected three weeks after study completion. The teacher rated the intervention highly when social validity data were initially collected, and student disruption levels remained at lower levels during maintenance observations. Two teacher participants in Mitchell et al., (2014) generalized their use of the GBG to other class periods not targeted during the study. The teachers stated that they believed the intervention would be useful for students in their other class periods as well. Overall teachers rated the use of group contingency interventions as favorable.

### **Research Question Three**

Of the studies that collected student social validity data, three reported mixed results and four reported positive outcomes. Bohan et al. (2020) discussed student mixed outcomes in their social validity findings. Although most students rated the CBGG highly in ways indicative of high social validity, some students discussed possible unfairness of the game if one team member was not following the rules. Student participants also expressed concerns that some students may not have been motivated by the putative reinforcers that were selected as a prize for the game. For example, Jones et al. (2019) reported some students had an unfavorable view of the intervention which aimed to decrease disruptive cell phone usage. Students expressed frustrations over not having access to their phones during class, but students did report an increase in focus during the contingency phases and they reported that the free time reinforcer received at the end of the period was powerful. Overall students rated group contingency interventions highly, showing acceptability of the interventions.

### **Research Question Four**

The findings of the quality indicator appraisal revealed that group contingencies may be considered an evidence-based practice intervention to implement for students. Results of the appraisal revealed that all included studies met the CEC standards with reservations. More recent studies included in this review met more of the quality indicator standards, perhaps because treatment integrity and social validity data are more commonly reported when compared to prior decades. Overall, the studies in this review may be considered to have good methodological rigor determined by the standards of the CEC and there is empirical support for using group contingencies to target adolescent disruptive and academically engaged behavior.

## **General Conclusions**

Collectively the findings from this review suggest that group contingencies reliably increase desired student academic engagement and decrease disruptive student behaviors. Student ages and grade levels ranged throughout the middle and high school age range. There were decreases in disruptive behavior and increases in academic engagement across all grade levels. This may suggest that group contingencies can have broad, positive effects on disruptive behavior and academic engagement for adolescent students. Studies took place in various settings and with a diverse group of students, including students with disabilities. This may mean that group contingencies have broad positive effects and can be implemented in specialized or general education settings. The socioeconomic diversity of participants suggest that group contingencies can be effective regardless of socioeconomic background. Many researchers did not include demographic details in their reports on effects of group contingency interventions.

Importantly, none of the included studies used function-based approaches to behavior interventions and supports. Determining function is unnecessary for group contingency interventions and are a relatively feasible and cost-effective way to improve student behavior. Very few materials are typically needed, and group contingencies can be implemented successfully by novice and veteran educators alike. Conversely, function-based interventions may require more time, money and resources typically are involved. For example, functional behavior assessment requires multiple observations, teacher and parent interviews, and records review, which may occur over the course of days or weeks. Functional behavior assessment and intervention also requires advanced training. However, group contingency interventions can be taught to implementors in a short amount of time and immediately applies by novice and veteran teachers while producing positive behavior change without the investment associated with

function-based interventions. Group contingency interventions may be superior to function-based intervention in some circumstances and should be considered if targeted behavior is not dangerous to the student or peers.

### **Limitations**

There are a few limitations that merit mention. First, the search terms used to locate relevant studies may be a limitation. The author tried different combinations of relevant terms such as “adolescent”, “group contingency”, “good behavior game”, “teen” and “young adult”. The terms were chosen because these terms may provide the most relevant articles to select from. Using only these search terms may have caused the author to miss some articles that may have met the inclusion criteria. For example, studies that used group contingencies (e.g., class-wide token economy), but did not characterize or describe their intervention as a group contingency may have not been located via electronic search. Second when conducting a hand search of journals only the *Journal of Applied Behavior Analysis* was searched. Relevant articles may have been missed because they were within other important journals. Third there was no reliability data collected on the search and screening of articles. Only the author of the study conducted the search using the terms selected, so it is unclear whether the search can be independently replicated. Screening of the articles retrieved from the search was only conducted by the author, so it is possible that some studies that met the inclusion criteria were discarded. Fourth the descriptive coding scheme was developed by the author. Other researchers who use different coding criteria or definitions may obtain different results. Fifth the author may have made decisions during coding that reflect implicit biases associated with various contingencies. One way this could have been addressed is by having a second researcher code all included studies to identify potential coding errors.

## **Implications for Research and Practice**

Group contingencies in classrooms can effectively improve for a wide range of behaviors and for various adolescent student populations. The different types of group contingencies afford professionals the capacity to decide which group contingency best fits the needs of their classrooms and students. Although all group contingency interventions have shown positive effects for students and teachers in previous research, professionals should choose a contingency that is feasible, will address the behavior(s) of concern, and that students will find enjoyable.

Reinforcement was important for the students. Group contingency interventions allow students to access reinforcement on different schedules and a variety of reinforcers can be used determined by student preference. Group contingencies are used widely in classrooms because they are simple to implement, and this review found this to be true amongst educators of adolescent students. Teachers interested in implementing a group contingency intervention in their classroom with adolescent age students should first ensure student buy-in. For example, adolescent age students may respond better to the intervention being called a competition rather than a game, but research has not yet established this. Describing group contingencies as a competition rather than a game also may sound more mature to adolescent students and more relatable. When awarding points, the teacher must decide whether to award points immediately (e.g., by keeping score on the board throughout implementation) or delay revealing the points awarded at the end of the session. Studies in this review suggest that delaying awarding points until the end of the session may be similarly effective as awarding them immediately and could cause less disruption throughout the class period. Future researchers may investigate whether and to what degree these variations influence outcomes and when one procedure is potentially better.

Choosing age-appropriate reinforcement is an important aspect to group contingencies. Adolescent age students will be motivated by different things than elementary age students. Older students may value reinforcers such as free time, homework passes, technology using, or a pizza party. When determining appropriate reinforcement for students, teachers must take into consideration the ease in which they can provide the reinforcement and what is feasible for them. This may mean providing a delay in reinforcement to students. Studies in this review found that secondary school teachers may prefer providing weekly rather than daily reinforcement to students so as not to disrupt the rest of the lesson and as to not run out of time due to changing class periods. Delaying reinforcement to weekly rather than daily may not have broad effects on student performance if the reinforcement provided is preferred. Future researchers might investigate the differential effects associated with immediate and delayed reinforcement.

Group contingencies can be implemented with different components such as self-monitoring. Peer-driven contingencies can be used in conjunction individualized interventions. Gaining peer approval and avoiding peer disapproval could have impacts on group contingency results. Peer contingencies in combination with teacher rule-governed contingencies may promote student self-monitoring of their own behavior. Group contingencies are good interventions to pair with other components, self-monitoring being one of them (Jones et al., 2019 & Christ & Christ, 2006).

There was only one study in this review (Jones et al., 2019) that aimed to decrease the amount of disruptive cell phone usage in the classroom. With the growing use of technology in schools and ubiquitous access to mobile devices, professionals may find themselves needing more support to decrease disruptive uses. For example, Thomas and Orthober (2011) found that over 80% of high school age students text daily and about half stated they text hourly. Kuznekoff



& Titsworth (2013) estimate that 64% of teenagers text during class, even if cell phone restrictions are in place. If students are texting or using their phones inappropriately during class, they are less likely to engage in academic activities (Kuznekoff & Titsworth, 2013).

Accordingly, future researchers should investigate whether and to what extent include group contingencies decrease disruptive technology use. Some researchers have begun to investigate the effects cell phones have on student learning, teen technology addictions, and mental and physical health impacts (Khoshgoftar et al., 2019; Kuznekoff & Titsworth, 2013; Stephens & Allen, 2013), but very little is known about interventions to address these issues. Notably, Jones et al. (2019) found interdependent group contingency resulted in decreased student cell phone, and that students reported feeling more focused during academic instruction when the group contingency was applied. Additional studies on the effect of group contingencies on disruptions related to mobile device technology could clarify whether and under what circumstances the intervention is effective for addressing this problem.

Although group contingencies can be considered an evidence-based practice for reducing disruptive behavior and increasing academic engagement in adolescent students, it remains unclear for whom and under what conditions specific students benefit (or do not benefit) from group contingencies. Finding this information may increase the efficiency in which these interventions are used. Future researchers should include details about participants, especially those from traditionally marginalized groups, in order to foster advancement in this area. Similarly, researchers should consider conducting meta-analyses of group contingencies to understand whether and to what extent student and teacher demographics, setting type, and type of group influence responding. This information could help professionals make decisions about whether to use group contingencies and how to best support their specific students. Such

guidance also could help professionals avoid wasting time and resources on interventions unlikely to produce the desired results. Overall, the findings of this review provide general support for group contingencies to decrease disruptive behaviors and increase academic engagement for adolescent students in secondary settings.

## **Conclusion**

The purpose of this review was to determine if group contingency interventions were appropriate for the adolescent population to decrease disruption and increase academic engagement in the classroom. To answer the subsequent questions a systematic review of the current literature was completed. This review expanded on the findings of previous reviews conducted by Little et al. (2015) and Maggin et al. (2017) by focusing on the effects group contingencies have on the adolescent population, which until recent years was not as included in the literature. The findings of this review suggest that group contingency interventions should be considered by professionals for use in secondary classrooms to decrease disruptive student behavior and increase student academic engagement. Outcomes of studies in this review suggest that group contingency interventions have broad positive effects at decreasing disruption and increasing academic engagement for the adolescent population. Social validity outcomes suggest overall students and teachers are in favor of the intervention. Results of the quality indicator appraisal suggest that group contingencies have good methodological rigor. Group contingencies are simple, cost-effective interventions that professionals can quickly implement to address behavior issues in the classroom. They are a viable option for professionals who are looking for an effective intervention to decrease disruption and increase academic engagement for adolescent age students.

## REFERENCES CITED

- Bohan, C., Smyth, S., & McDowell, C. (2021). An evaluation of the caught being good game with an adolescent student population. *Journal of Positive Behavior Interventions*, 23(1), 3-64. <https://doi.org/https://doi-org.libproxy.temple.edu/10.1177/1098300720928455>
- Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good Behavior Game: Effects of Individual Contingencies for Group Consequences on Disruptive Behavior in a Classroom. *Journal of Applied Behavior Analysis*, 2(2), 119–124. <https://doi.org/10.1901/jaba.1969.2-119>
- Carden Smith, L. K., & Fowler, S. A. (1984). Positive peer pressure: The effects of peer monitoring on children's disruptive behavior. *Journal of Applied Behavior Analysis*, 17(2), 213-227.
- Chafouleas, S. M., Hagermoser Sanetti, L. M., Jaffery, R., & Fallon, L. M. (2011). An evaluation of a class wide intervention package involving self-management and a group contingency on classroom behavior of middle school students. *Journal of Behavioral Education*, 21, 34-57. <https://doi.org/10.1007/s10864-011-9135-8>
- Christ, T. J., & Christ, J. A. (2006). Application of an interdependent group contingency mediated by an automated feedback device: An intervention across three high school classrooms. *School Psychology Review*, 35, 78–90.
- Common, E. A., Lane, K. L., Cantwell, E. D., Brunsting, N. C., Oakes, W. P., Germer, K. A., & Bross, L. A. (2020). Teacher-delivered strategies to increase students' opportunities to respond: A systematic methodological review. *Behavioral Disorders*, 45(2), 67–84.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). *Applied behavior analysis* (3rd ed.). Columbus, OH: Merrill Prentice Hall.
- Council for Exceptional Children: Standards for Evidence-Based Practices in Special Education. (2014). *TEACHING Exceptional Children*, 46(6), 206–212. <https://doi-org.libproxy.temple.edu/10.1177/0040059914531389>
- Dart, E. H., Radley, K. C., Battaglia, A. A., Dadakhodjaeva, K., Bates, K. E., & Wright, S. J. (2016). The classroom password: A class-wide intervention to increase academic engagement. *Psychology in the Schools*, 53, 416–431.
- Davies, S., & Witte, R. (2000). Self-management and peer-monitoring within a group contingency to decrease uncontrolled verbalizations of children with attention deficit/hyperactivity disorder. *Psychology in the Schools*, 37, 135–147
- De Martini-Scully, D., Bray, M., & Kehle, T. (2000). A packaged intervention to reduce disruptive behaviors in general education students. *Psychology in the Schools*, 37, 149–156.

- Flower, A., Mckenna, J. W., Bunuan, R. L., Muething, C. S., & Vega, R. (2014). Effects of the good behavior game on challenging behaviors in school settings. *Review of Educational Research, 84*(4), 546–571.
- Ford, W. B., Radley, K. C., Tingstrom, D. H., & Dufrene, B. A. (2020). Efficacy of a no-team version of the good behavior game in high school classrooms. *Journal of Positive Behavior Interventions, 22*(3), 131-191. <https://doi.org/https://doi.org.libproxy.temple.edu/10.1177/1098300719890059>
- Gersten, R., Fuchs, L. S., Innocenti, M. S., Compton, D., Coyn, M., & C. G. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Council for Exceptional Children, 71*(2). <https://doi.org/https://doi.org.libproxy.temple.edu/10.1177/001440290507100202>
- Hansen, S. D., & Lignugaris/Kraft, B. (2005). Effects of a dependent group contingency on the verbal interactions of middle school students with emotional disturbance. *Behavioral Disorders, 30*, 170–184.
- Heering, P. W., & Wilder, D. A. (2006). The use of dependent group contingencies to increase on-task behavior in two general education classrooms. *Education and Treatment of Children, 29*, 459–468.
- Hoag, J. (2007). Effects of mystery motivators vs. expected reinforcers on disruptive behavior in preschoolers. Dissertation Abstracts International: Section B. *The Sciences and Engineering, 67*(10-B), 6041.
- 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*(2), 165–179. <https://doi.org.libproxy.temple.edu/10.1177/001440290507100203>
- Jenson, W. R. (1978). Behavior modification in secondary schools: Review. *Journal of Research and Development in Education, 11*(4), 53–63.
- Jones, M. E., Allday, R. A., & Givens, A. (2018). Reducing adolescent cell phone usage using an interdependent group contingency. *Journal of Applied Behavior Analysis, 52*(2), 386-393.
- Khoshgoftar, Mohadeseh & Mazaheri, Maryam & Tarahi, Mohammad. (2022). The Effect of Educational Intervention Based on Health Belief Model to Decrease and Prevent Mobile Phone Addiction among Female High School Students in Iran. *International Journal of Pediatrics, 7*(10), 10175-10185.
- Kleinman, K. E., & Saigh, P. A. (2011). The effects of the good behavior game on the conduct of regular education New York City high school students. *Behavior Modification, 35*(1), 95–105.

- Kelshaw-Levering, K., Sterling-Turner, H. E., Henry, J. R., & Skinner, C. H. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37*, 523–533.
- Kohler, F. W., Strain, P. S., Hoyson, M., Davis, L., Donina, W. M., & Rapp, N. (1995). Using a group-oriented contingency to increase social interactions between children with autism and their peers: A preliminary analysis of corollary supportive behaviors. *Behavior Modification, 19*, 10–32.
- Kuznekoff, J. H., & Titsworth, S. (2013). The impact of mobile phone usage on student learning. *Communication Education, 62*, 233-252. <https://doi.org/10.1080/03634523.2013.767917>
- Lane, K., Kalberg, J. R., & Menzies, H. M. (2009). *Developing schoolwide programs to prevent and manage problem behaviors: A step-by-step approach*. New York: Guilford Press.
- 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. <https://doi.org/10.1177/0145445514554393>
- Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8*, 341–347. <https://doi.org/10.1901/jaba.1975.8-341>
- Lohrmann, S., & Talerico, J. (2004). Anchor the boat: A classwide intervention to reduce problem behavior. *Journal of Positive Behavior Interventions, 6*, 113–120.
- Lum, J. D., Radley, K. C., Tingstrom, D. H., Dufrene, B. A., Olmi, D. J., & Wright, S. J. (2019). Tootling with a randomized independent group contingency to improve high school classwide behavior. *Journal of Positive Behavior Interventions, 21*(2), 67-127. <https://doi.org/https://doi-org.libproxy.temple.edu/10.1177/1098300718792663>
- Luiselli, J. K., Putnam, R. F., & Sunderland, M. (2002). Longitudinal evaluation of behavior support intervention in a public middle school. *Journal of Positive Behavior Interventions, 4*(3), 130-194. <https://doi.org/https://doi-org.libproxy.temple.edu/10.1177/10983007020040030701>
- Madaus, M. R., Kehle, T. J., Madaus, J., & Bray, M. A. (2003). Mystery motivator as an intervention to promote homework completion and accuracy. *School Psychology International, 24*, 369–377.
- Maggin, D. M., Pustejovsky, J. E., & Johnson, A. H. (2017). A meta-analysis of school-based group contingency interventions for students with challenging behaviors: An update. *Remedial and Special Education, 38*, 353-370. <https://doi.10.1177/0741932517716900>.

- Malone, R. A., & McLaughlin, T. F. (1997). The effects of reciprocal peer tutoring with a group contingency on quiz performance in vocabulary with seventh- and eighth-grade students. *Behavioral Interventions, 12*, 27–40.
- Mitchell, R. R., Tingstrom, D. H., Dufrene, B. A., Ford, W. B., & Sterling, H. E. (2015). The effects of the good behavior game with general-education high school students. *School Psychology Review, 44*, 191207. <https://doi.org/10.17105/spr-14-0063>.
- Musser, E., Bray, M., Kehle, T., & Jenson, W. (2001). Reducing disruptive behaviors in students with serious emotional disturbance. *School Psychology Review, 30*, 294–304.
- Robinson, K., & Sheridan, S. (2000). Using the mystery motivator to improve child bedtime compliance. *Child & Family Behavior Therapy, 22*, 29–49.
- Schanding, G., & Sterling-Turner, H. E. (2010). Use of the mystery motivator for a high school class. *Journal of Applied School Psychology, 26*, 38-53.  
doi:10.1080/15377900903379448
- Solomon, R. W., & Whaler, R. G. (1973). Peer reinforcement control of classroom problem behavior. *Journal of Applied Behavior Analysis, 6*(1), 49-56.
- Stage, S. A., & Quiroz, D. R. (1997). A Meta-Analysis of Interventions to Decrease Disruptive Classroom Behavior in Public Education Settings. *School Psychology Review, 26*(3), 333-368. <https://doi.org/https://doi-org.libproxy.temple.edu/10.1080/02796015.1997.1208587>.
- Thomas, K., & Orthober, C. (2011). Using text messaging in the secondary classroom. *American Secondary Education, 39*, 55-76.
- Theodore, L. A., Bray, M. A., Kehle, T. J., & Jensen, W. R. (2001). Randomization of group contingencies and reinforcers to reduce classroom disruptive behavior. *Journal of School Psychology, 39*, 267–277. doi:10.1016/S0022-4405(01)00068-1
- Wahl, E., Hawkins, R. O., Haydon, T., Marsicano, R., & Morrison, J. Q. (2016). Comparing versions of the good behavior game: can a positive spin enhance effectiveness? *Behavior Modification, 40*(4). <https://doi.org/10.1177/0145445516644220>
- Williamson, B. D., Campbell-Whatley, G. D., & Lo, Y. (2009). Using a random dependent group contingency to increase on-task behaviors of high school students with high incidence disabilities. *Psychology in the Schools, 46*, 1074-1083.
- Wright, R. A., & McCurdy, B. L. (2012). Class-wide positive behavior support and group contingencies: examining a positive variation of the good behavior game. *Journal of Positive Behavior Interventions, 14*(3), 173–180.  
<https://doi.org/10.1177/1098300711421008>