THE EFFECTS, GENERALIZATION, AND INCIDENTAL BENEFITS OF CLASS-WIDE FUNCTION-RELATED INTERVENTION

By

© 2020

Lina G. Kitson Ed.S., University of Kansas, 2017 B.A., University of Cincinnati, 2013

Submitted to the graduate degree program in School Psychology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Christopher Niileksela

Patricia Lowe

Matthew Reynolds

David Hansen

Kathleen Zimmerman

Howard Wills

Date Defended: 30 March 2020

The dissertation committee for Lina G. Kitson certifies that this is the approved version of the following dissertation:

THE EFFECTS, GENERALIZATION, AND INCIDENTAL BENEFITS OF CLASS-WIDE FUNCTION-RELATED INTERVENTION

Chair: Christopher Niileksela

Date Approved: 6 April 2020

ABSTRACT

Problem behaviors are prevalent in schools and may result in negative effects for the students exhibiting those behaviors, teachers, and other students. Increasing on-task behaviors has been shown to minimize disruptive behaviors. The current study utilized an interdependent group contingency-based intervention, Class-Wide Function-related Intervention Teams (CW-FIT), to increase on-task behavior in elementary classrooms. A concurrent multiple baseline design was used to examine the effects of CW-FIT during whole-group instruction for six at-risk students in three classrooms (kindergarten, first, and third grade). The study examined whether CW-FIT would: 1) increase on-task behaviors of students, 2), increase teacher praise and reduce teacher reprimands, 3) result in improved on-task behavior, praise, and reprimands during times of the day when CW-FIT was not implemented, and 4) increase attendance and classroom engagement, and decrease office discipline referrals. Results showed CW-FIT improved on-task behavior for five out of six at-risk students. The sixth student already demonstrated high on-task behavior before CW-FIT was implemented. Praise rates increased for all teachers, although praise rates decreased over time for two teachers. Few teacher reprimands were observed, so it was not possible to evaluate effects on CW-FIT on reprimands. Insufficient data were available to make strong conclusions on whether the effects of CW-FIT generalized to other parts of the day; however, there was some evidence that on-task behavior improved for at least one student from each classroom when CW-FIT was not being implemented. No significant differences were found in attendance rates before and after CW-FIT was implemented. Self-reported ratings of classroom engagement for all at-risk students improved after CW-FIT. No office discipline referrals were reported across classrooms. CW-FIT was well accepted by students and teachers. Overall, CW-FIT produced an increase in on-task behavior, teacher praise, improved student engagement in the classroom, and the effects may generalize to other parts of the day.

ACKNOWLEDGMENTS

When they say it takes a village to finish your doctorate they are not kidding! First I would like to thank my dissertation chair, Dr. Niileksela, who was not only an incredible supporter and mentor throughout this process, but also one of the most amazing professors that I have had the privilege to learn from. He has made this process enjoyable, never made me feel inadequate, and always treated me more like a peer than a student. For that I will be forever grateful. I would also like to thank my committee who provided me with invaluable feedback and ideas that made this project possible. Thank you Dr. Wills for giving me access to this incredible intervention and for your guidance. Dr. Zimmerman, I cannot thank you enough for your expertise and advice. Thank you Dr. Reynolds and Dr. Hansen for giving me such valuable feedback during my proposal and defense. And of course thank you Dr. Lowe for your mentorship and support throughout my time as a doctoral student. I would also like to thank my wonderful teachers who took part in my study. Thank you for going along this journey with me and trusting in me. And of course none of this would have been possible without the support from my family and friends who cheered me on. Last but most certainly not least, I would like to thank my love, my heart, my husband. You have been my rock throughout this crazy process. You believed in me when I didn't, you never doubted me, and without your love and support I could not have done this. You inspire me to be a better person and to persevere when I want to give up.

Table of Contents

Chapter I: Introduction	1
Problem Statement	1
Group Contingency Interventions	2
Class-Wide Function-Related Intervention Teams	3
Previous Research with CW-FIT	3
Gaps in the CW-FIT Literature	4
Current Study	6
Relevance to School Psychology	7
Chapter II: Review of Literature	9
Negative Correlates and Outcomes of Behavior Problems	9
Ineffective Strategies to Manage Disruptive Behaviors1	1
Effective Strategies to Manage Behaviors1	4
Multi-Tiered Supports	5
Progress Monitoring	1
Academic progress monitoring	2
Behavioral progress monitoring	3
Evaluating behavior change	5
Targeting On-Task Behavior	6
Dependent group contingency	7
Independent group contingency	8
Interdependent group contingency	8
Interventions using Interdependent Group Contingencies	9
Good behavior game	9
PAX good behavior game	0
Class-wide function-related intervention teams	1
Theoretical Orientation	3
Review of Studies that Utilized CW-FIT	5
Gaps in the Literature on CW-FIT	8
Current Study and Research Questions	2
Chapter III: Methods	4

44
44
44
45
45
46
47
48
48
48
49
49
49
49
49
50
51
51
53
54
54 55
54 55 55
54 55 55 57
54 55 55 57 58
54 55 55 57 58 59
54 55 55 57 58 59 60
54 55 55 57 58 59 60 60
54 55 55 57 58 59 60 60 60
54 55 55 57 58 59 60 60 60 60
54 55 55 57 58 59 60 60 60 60 61 61
•

Procedural Fidelity	
Interobserver Agreement	
Research Question 1: Replicating Effects of CW-FIT	
On-task behavior	
On-task behavior maintenance	
Praise and reprimands	
Research Question 2: Generalization of CW-FIT	
On-task behavior generalization	
Generalization praise and reprimand	
Research Question 3: CW-FIT and Other Positive Outcomes	
Office discipline referrals	
Attendance	
Social Validity	
Peer comparisons	
Teacher questionnaire.	
Student questions	
Chapter V: Discussion	
On-Task Behavior	
Praise and Reprimands	
Teacher praise	
Teacher reprimands	
Generalization of CW-FIT Effects	
On-task behaviors	
Praise and reprimand	
Attendance/Student Satisfaction/Office Discipline Referrals	
Attendance	
Office discipline referrals	
Student classroom satisfaction	
Social Validity	
Limitations	
Recommendations for Future Research	
Implications for Practice	

List of Appendices

Appendix A. Approval Letter from the University of Kansas Institutional Board	123
Appendix B. Modified Version of the EvsD Scale	127
Appendix C. CW-FIT Procedural Fidelity Checklist	131
Appendix D. Consumer Satisfaction Survey (Social Validity)	133
Appendix E. Student Consumer Satisfaction Questions (Social Validity)	138

List of Tables

Table 1. Student Demographic Information	46
Table 2. Intervention Fidelity Components	63
Table 3. Proportion of student absences before and after the intervention was implemented	
Student satisfaction in the classroom	86
Table 4. Percentage of at-risk students that rated each item as True (3) or Very True (4)	86

List of Figures

Figure 1. On-task behaviors for each student across teachers and phases during math class 66
Figure 2. Whole class praise and reprimand rates across teachers and phases during math class
Figure 3. On-task behaviors for each student across teachers and phases during generalization
Figure 4. Generalization class-wide praise and reprimand rates across teachers and phases83

Chapter I: Introduction

Problem Statement

Problem behaviors have numerous negative outcomes. Students who exhibit problem behaviors are more likely to experience peer rejection, mental health difficulties, poorer academic outcomes, and have lower attachment to school (Dunlap et al., 2006; Merrell & Gueldner, 2010). Conversely, past research has found that students who demonstrate on-task behaviors such as eye contact with the teacher, orienting to task, and working quietly are less likely to engage in disruptive behaviors, tend to earn better grades, and therefore have more positive behavioral and academic outcomes (Rhode, Jenson, & Reavis, 1993; Van Houten & Nau, 1980).

Problem behaviors also impact those around the students that engage in the unwanted behaviors, such as teachers and other students. Surrounding students report difficulty focusing on academic content due to the noisy environment that is often created by disruptive students and believe that more learning would occur if those students were removed from the classroom (Public Agenda, 2004). Additionally, teachers report losing as many as 50 school days each year of instruction due to managing disruptive behaviors (Bru, 2009). When teachers spend a substantial amount of time handling these behaviors, they have poorer job satisfaction and lower self-efficacy, which may lead them to leave the profession (Kokkinos, 2007; Public Agenda, 2004).

Some teachers use methods such as punishment and exclusion from the classroom to manage disruptive behaviors (Lannie & Murakami, 2007; Rose, 1988; Sprick, Borgmeier, & Nolet, 2002). However, these methods tend to be ineffective at producing long-term behavior change (Costenbader & Markson, 1998) and negatively impact teacher-student relationships (Jamieson & Thomas, 1974; Lewis, 2001). In contrast, strategies such as relationship building,

1

positive reinforcement, and clear expectations have been found to be more effective and positive ways to teach appropriate behaviors. Due to diverse needs of students, a continuum of supports such as Multi-Tiered System of Support (MTSS) has been developed. In this type of framework students are provided increasingly intense levels of support based on their needs using interventions designed to address problematic behaviors and academic difficulties (McIntosh & Goodman, 2016).

Determining what specific disruptive behaviors should be targeted may be challenging. Past findings show that increasing student on-task behavior decreases disruptive behaviors as students have fewer opportunities to engage in the unwanted behaviors (Hawken & Horner, 2003). Interventions that targeted decreasing off-task behaviors found a decrease in behaviors such as property destruction (McComas, Goddard, & Hoch, 2002), refusal behaviors, and throwing objects (Peterson et al., 2005).

Therefore, given the knowledge that teachers spend a lot of time handling behavior problems along with the importance of targeting off-task behaviors, it is not only critical to provide teachers with strategies that target decreasing off-task behaviors but also strategies that are time-efficient and easy to implement. One cost effective and highly acceptable set of classroom management strategies that have been used to decrease off-task behaviors and improve academic performance are group contingencies (Little, Akin-Little, & O'Neill, 2015).

Group Contingency Interventions

There are three types of group contingencies: dependent group contingency, independent group contingency, and interdependent group contingency (Litoe & Pumroy, 1975). Amongst those types, the interdependent group contingency has been found to be the most commonly used (Little et al., 2015). Various interventions have used the interdependent group contingency as

their structure such as the Good Behavior Game (GBG, Barrish, Saunders, & Wolf, 1969), PAX Good Behavior Game (PAX GBG, Embry, 2002), and Class-wide function-related intervention teams (CW-FIT; Wills et al., 2010).

Class-Wide Function-Related Intervention Teams

While all three previously mentioned interventions have been effective in decreasing disruptive behaviors, the current study will focus on CW-FIT. This intervention has three components which include: 1) teaching of appropriate skills, 2) extinction (i.e., removal of reinforcement) of unwanted behaviors, and 3) rewarding student teams with points for engaging in appropriate behaviors. CW-FIT aims at teaching students prosocial behaviors within the classroom environment while using positive reinforcement (Wills et al., 2010). During this intervention the interdependent group contingency addresses attention, which is a common function of unwanted behaviors, such as: 1) how to get the teacher's attention correctly, 2) how to follow directions the first time, and 3) how to ignore inappropriate peer behavior. Students are divided into teams, and after a specified interval of time if all students in a team are displaying appropriate behaviors, the team will receive a point. All teams are aiming to earn a predetermined number of points to receive a reward at the end of the game (Wills et al., 2010).

Previous Research with CW-FIT

CW-FIT has been implemented across grade levels from preschool to seventh grade. In addition to core classes (e.g., reading, mathematics), it has been implemented in other contexts, such as physical education classes, special education classrooms, and music classes. Results from these studies have shown an increase in on-task behaviors and teacher praise with a concurrent decrease in disruptive behaviors and teacher reprimands. These effects were found class-wide and in individual students. Other behaviors such as compliance, hand raising, remaining seated when expected, and talking out that were taught as part of the CW-FIT protocol also improved. Specifically, CW-FIT resulted in increased compliance, hand raising, remaining seated, and a decrease in talking out (Conklin, Kamps, & Wills, 2017). Incidental benefits of CW-FIT have also been examined, such as work completion and grades; however, findings did not show a meaningful effect on these factors (Conklin, 2010). Finally, the effects of CW-FIT across the day were examined in which the intervention was implemented three times per day. Results showed positive effects across all implementation sessions. Nevertheless, there are still some questions about CW-FIT that have not been answered in the literature.

Gaps in the CW-FIT Literature

Other student outcomes that are important indicators of success have not been examined comprehensively with CW-FIT. For instance, student attendance is an important component in positive academic and behavioral outcomes. Loss of instruction time due to missed school days leads to poorer grades (Devadoss & Foltz, 1996), an increased risk of drop out, and students with poor attendance are more likely to engage in problem behaviors (Sheldon, 2007). In past literature, group contingencies have not been utilized to increase student attendance, however, factors such as positive reinforcement, clear rules, and structure have been linked to an increase in student attendance. As these components are included in the CW-FIT intervention, it is possible to hypothesize that attendance could be positively impacted by the implementation of CW-FIT.

Office discipline referrals have additionally not been studied using group contingency interventions. However, past findings show that when expectations are taught, students are pre-corrected and reminded of the expected behaviors, and appropriate behaviors are reinforced,

there appears to be a decrease in office discipline referrals (Weber, 2011). Additionally, interventions that have focused on disruptive behaviors have demonstrated a decrease in office discipline referrals (Hawken, Macleaod, & Linda 2007). Therefore, based on these findings it is possible to hypothesize that an intervention that includes components such as teaching expectations, pre-correcting behaviors, and reinforcing positive behaviors, as well as targeting off-task behaviors may reduce office referrals.

Similar to attendance and office discipline referrals, student engagement has not been targeted using group contingency interventions. However, previous research has shown that student engagement is a predictor of academic success. Studies also show that school factors such as choices, clear goals, and student involvement are present in schools with higher student engagement. Therefore, as CW-FIT includes these components as part of the intervention, it is possible to also hypothesize that student engagement in the classroom will improve (Fredricks et al., 2004).

Finally, another important component that has not been addressed in the literature is generalization, where any behavior change that took place in the setting the intervention was implemented also occurred in other settings and over time (Stokes & Baer, 1977). Generalization is an important component in the transference of skills and the future success of using those skills and other skills that may be related to it. Even if not systematically planned and implemented, it is possible that the effects of the intervention may generalize to other settings. While CW-FIT has been implemented multiple times during the day, previous research has not examined the generalization of an intervention during times of the day that the intervention is not implemented. As CW-FIT is intended to be implemented during times of the day that are most challenging for students, it is important to determine if the skills taught and reinforced during the

intervention session generalize to other parts of the day. To address these gaps, the following research questions will be answered:

- Can the effects of CW-FIT that have been found in previous research be replicated in new classrooms? Specifically, it is hypothesized that during the implementation of CW-FIT there will be: a) an increase student time on-task, b) an increase teacher praise statements, and c) a decrease teacher reprimand statements.
- 2. Do the effects of CW-FIT found in previous research generalize to the portions of the day the intervention is not used? Specifically, it is hypothesized that during times when CW-FIT is not implemented there will be: a) an increase in student time on-task, b) an increase teacher praise statements, and c) a decrease teacher reprimand statements.
- 3. Does the implementation of CW-FIT produce other positive outcomes for students not investigated in previous research? Specifically, it is hypothesized that after CW-FIT is implemented there will be: a) a reduction in office referrals, b) an increase in attendance, and c) be an improvement in classroom engagement.

Current Study

The current study included three teachers and three students from each of their classrooms from at Title I public elementary school. Initially, teachers were trained on the implementation of CW-FIT. Next, each teacher rank ordered all students in the classroom based on number of disruptive and off-task behaviors. Only students that are fully included in the classroom were ranked. Two students who demonstrated the most disruptive behaviors and an average comparison peer were chosen in each classroom. While all students in each classroom engaged in the CW-FIT intervention, on-task observations, attendance, student engagement, and office discipline referrals were collected from the individual students. Along with individual student data, teacher praise and reprimands were also be collected. A concurrent multiple baseline design was used for the study and each teacher implemented the intervention for seven weeks. Given the effectiveness of CW-FIT as shown in previous findings, the current study aimed to replicate the positive effects of CW-FIT in the classroom setting. Furthermore, this study provided new contributions to the literature by examining the generalization of the intervention to other settings as well as other outcomes such as attendance, student engagement, and office discipline referrals.

Relevance to School Psychology

This paper describes an intervention that is not only easily implemented by teachers in the classroom but has also shown to positively impact student and teacher behavior. While teachers have vast knowledge about classroom management, various individuals within the school have specialized training in behavior and classroom management that could aid teachers when the methods that they have attempted have not produced desired results. One of those individuals is the school psychologist. The role of school psychologists has changed over the years (Curtis, Grier, & Hunley, 2004; Skalski et al. 2015). In today's educational system, school psychologists often collaborate and consult with teachers to determine what interventions and strategies can be used in the classroom to improve student learning. While problem-solving individual student needs is important, it is also vital to utilize class-wide strategies to engage all students. School psychologists may encounter teachers who would like to improve their general classroom management, and CW-FIT may be one intervention package that they could recommend and help teachers implement. School psychologists have training in prevention and early intervention methods and can determine if a specific program like CW-FIT would be appropriate for a specific student population. School psychologists are also trained in data-based decision making and would be able to help teachers develop data collection methods to determine if the intervention is having the desired effect on the target behaviors. Fidelity is an additional component important to increasing the chance of an intervention producing desired results. School psychologists have knowledge in creating and determining whether an intervention is being implemented with fidelity and provide feedback if needed. Finally, as CW-FIT includes an opportunity for increased support, school psychologists can aid teachers in analyzing the data to determine if increased supports are necessary. Overall, school psychologists need to have a strong repertoire of options to help teachers at the individual, student, and classroom levels. An evidence-based, cost-efficient, feasible, and effective intervention that improves on-task behaviors like CW-FIT may be especially useful for school psychologists to recommend to teachers.

Chapter II: Review of Literature

It is estimated that 3-6% of students in the United States display markedly disruptive behaviors (Kauffman, 2001). More recently, the prevalence of early elementary students exhibiting problem behaviors has been found to be between 7-10% and as high as 20% through later elementary and secondary schooling (Beaman, Wheldall, & Kemp, 2007). Problem behaviors are defined as "any repeated pattern of behavior that interferes with or is at risk of interfering with optimal learning or engaging in prosocial interactions with peers and adults" (Perry, Dallas Allen, Brennan, & Bradley, 2010). Teachers continue to report negative behaviors such as bullying, disrespectful behaviors, verbal abuse, and disorder in the classroom on a daily or weekly basis, and therefore must spend a significant amount of time handling disrespectful and off-task behaviors. Such behaviors were reported to be most challenging for teachers and resulted in the most stress on them professionally (Scott, Park, Swain-Bradway, & Landers, 2007).

Negative Correlates and Outcomes of Behavior Problems

Students with behavioral problems often experience various negative outcomes, such as peer rejection, low academic performance, and mental health concerns that can last into adulthood (Dunlap et al., 2006; Merrell & Gueldner, 2010). These students tend to have reduced access to instruction due to behavioral disruptions, removal from the learning environment, or preoccupations with social or emotional difficulties, and therefore tend to have lower academic achievement (Spira, Bracken, & Fischel 2005). Additionally, students with behavioral problems often have less attachment to school, as well as higher levels of truancy (Battin-Pearson et al., 2000) and dropout (Vitaro, Brendgen, & Larose, 2005; Vitaro, Larocque, Janosz, & Tremblay, 2001). Conversely, student behaviors such as eye contact with a teacher, working quietly, and

orienting to a task are positively related to academic functioning and negatively related to disruptive behaviors (Rhode et al., 1993; Van Houten & Nau, 1980). These behaviors have been labeled as on-task behaviors (Rhode et al., 1993). Specifically, students who engage in on-task behaviors are able to complete a greater quantity of academic tasks in a more accurate manner and earn better grades. Additionally, compared to students who engage in more off-task behaviors, students who are on-task more frequently engage in less disruptive behavior, such as talking out of turn, making noises, wandering around the room, property destruction, and physical aggression. Therefore, students who are on-task more often experience higher levels of positive behavioral and academic outcomes (Rhode et al., 1993; Van Houten & Nau, 1980).

Disruptive behaviors not only affect the academic progress of the students who are engaging in the behavior, these behaviors also affect the academic progress of other students in the classroom. Bru (2009) found that 60% of students agree that they would have learned more if their classroom environment was quieter, and 40% indicated that noisy classrooms decreased their ability to focus on schoolwork. Additionally, in a 2004 survey, 70% of students said that they were distracted by disruptive students and 53% believe that they would learn more if these students were removed from the classroom (Public Agenda, 2004). Student disruptive behaviors add to a noisy learning environment and make it difficult for other students to concentrate on learning tasks.

Another negative consequence of disruptive behavior is that it can decrease teacher job satisfaction. Failure to decrease student disruptive behaviors may reduce teacher self-efficacy (Langdon, 1996), which may lead to burnout and the decision to leave the profession (Kokkinos, 2007; Public Agenda, 2004). According to Scates (2005), 56% of teachers reported losing as much as 25 minutes out of a 90-minute block due to disruptive behaviors, which equates to approximately 50 school days of lost instruction in an academic year. When there are disruptive students in the classroom, teachers are forced to spend much of their time and energy managing student behaviors, which reduces the amount of time spent on instruction and supporting other students in the classroom (Bru, 2009). Beaman et al. (2007) found that approximately 50% of teachers across all grade levels report spending more time dealing with student behavior than they believe they should.

Overall, the presence of disruptive behaviors in the classroom has a negative impact on both students and teachers. Specifically, students who engage in the disruptive behaviors have more negative academic and behavioral outcomes and other students in the classroom have more difficulty focusing when disruptive behaviors are present. Additionally, teachers spend valuable instructional time to manage these behaviors and experience a reduction of self-efficacy which may be related to a decrease in job satisfaction.

Ineffective Strategies to Manage Disruptive Behaviors

When students are disruptive in the classroom, one of the methods teachers may use to handle these behaviors are negative strategies, such as exclusion from the classroom (e.g., sending students into the hallway or office) or some form of punishment (e.g., taking away recess or other privileges). Although exclusion and punishment may result in an immediate ceasing of problem behavior (Cooper, Heron, & Heward, 2007), these methods are generally ineffective at producing long-term reduction in problem behavior (Costenbader & Markson, 1998). Despite a lack of long-term effects, they remain the most common methods of responding to behavior problems in schools (Lane & Murakami, 2007; Rose, 1988; Sprick et al., 2002).

Punishment is defined as a response to a behavior that is intended to decrease the behavior in the future. It is important to note that punishment is not defined by the actions of the

person delivering the punishment, it is only qualified as punishment when the behavior actually decreases in the future. There are two types of punishment: positive and negative. Positive punishment occurs when an aversive stimulus is applied to an individual in order to decrease the behavior (Cooper et al., 2007). For example, a form of positive punishment in the classroom may occur when a teacher assigns a student more assignments because they forgot to turn in their homework as a method for reducing the likelihood of the student forgetting to turn in their homework in the future. In contrast, negative punishment occurs when something desired is taken away from a person in order to reduce the behavior (Cooper et al., 2007). For example, a form of negative punishment may take place when a child has been aggressive to a sibling the parent may take away the child's phone as a method for reducing the likelihood they will be aggressive toward a sibling in the future.

There are many known negative effects of punishment. One such effect, termed *recovery from punishment*, occurs when punishment is discontinued and the behavior that was being punished not only reemerges later, it also increases and exceeds the level that it was at prior to the punishment. Punishment can also evoke emotional responses in the person receiving the punishment and leads to an increased rate of the unwanted behavior when the punishment is not administered. Another consequence of punishment may be escape or avoidance of punishment (Cooper et al., 2007). For example, a student who is punished for poor grades in math may begin to avoid going to math class altogether, or a student may begin to lie or engage in behaviors more discreetly as a way to escape or avoid punishment. Despite the numerous negative consequences of punishment, this method continues to be used, likely because of the immediate discontinuation of the problem behavior which acts as a reinforcement for the punisher (Cooper

et al., 2007). Nevertheless, continued use of this method may result in the negative side effects listed above.

Specific forms of punishment such as exclusionary practices (e.g., seclusion time-out) and zero tolerance policies are additionally ineffective behavior change agents. These practices result in the removal of student from educational opportunities which diminishes academic success. Additionally, they do not improve more appropriate or desired behaviors and may actually increase behavioral problems, as it allows for increased free time and access to peers that do not model appropriate behaviors (Sugai & Horner, 2002). Harmful behavioral response strategies such as punishment and exclusion have been shown to negatively impact teacherstudent relationships (Jamieson & Thomas, 1974; Lewis, 2001). Student-teacher relationships have been found to be predictors of academic and behavioral outcomes as early as elementary school. In addition to exclusionary and zero tolerance policies, Hamre and Pianta (2001) found that negative reports about the student's behavior in the classroom from kindergarten teachers were predictors of poor student grades, lower standardized test scores, and poor work habits. Behavioral outcomes were also predicted by relational negativity (i.e., amount of conflict between student and teacher and student dependence on the teacher), especially for students with early behavioral problems.

In addition to the ineffective behavioral management strategies that are often used to discipline students with behavioral problems, there is a well-documented disproportionality in which students receive disciplinary referrals and suspensions, specifically related to student race and gender. According to the National Center for Education Statistics (NCES, 2003), African American and Latino students are more than twice as likely to experience school suspension compared to their Caucasian peers. Skiba et al. (2002) found that African American students

were more likely to receive office referrals for more subjective offenses such as being disrespectful, making excessive noise, threats, and loitering. However, their Caucasian peers were more likely to receive office referrals for more objective offenses such as smoking, eloping, vandalism, and inappropriate language. Additionally, Rausch and Skiba (2004) found that African American students received suspensions at rates four times higher than Caucasian students, and Hispanic students received suspensions at rates two times higher than Caucasian students. With regards to the disproportionality in gender, African American males were much more likely to receive discipline referrals and suspensions than their Caucasian male peers (Wallace et al., 2008). This difference was also found in African American and Caucasian females (Bryan et al., 2012). In addition to the differences in race, Bryan et al. (2012) found that overall, female students were far less likely to be referred for disruptive behavior than were male students.

Effective Strategies to Manage Behaviors

While there are a number of negative methods for handling disruptive behavior, there are also positive methods that can decrease disruptive behavior and improve appropriate behavior. One component of improving disruptive behaviors is building a positive student-teacher relationship, which may promote positive behaviors and reduce negative behaviors. Studies show that students who feel emotionally supported by their teacher are more likely to have higher levels of on-task behaviors and lower levels of disruptive behaviors (Bru, 2006; Fraser & Fisher, 1982; Hamre & Pianta, 2001). Research has also shown a link between teacher behavior management and student classroom behavior. Adequate behavior management strategies such as clear expectations, rules, and consequences, along with the use of praise and rewards reinforces appropriate classroom behavior and has shown a reduction in disruptive classroom behaviors

(Barber & Olsen, 2004; Doyle & Carter, 1987; Ferguson & Houghton, 1992; Sutherland, Wehby, & Copeland, 2000). In contrast, behavior management techniques such as reprimands, corrections, and commands may elicit more child disruptions (Nelson & Roberts, 2000; Wehby Symons, & Shores, 1995).

Both academic and social skills are important to overall school success and academic engagement. Academic skills are defined as the knowledge of content taught and can be observed and measured by work completion, classroom participation, test scores, and appropriate study habits (Caldarella et al., 2017). On the other hand, social skills, or the ability to effectively interact with others, are a set of verbal and non-verbal behaviors and cognitive skills which are necessary to achieve positive short- and long-term outcomes from social interactions (Caldarella et al., 2017; Spence et al., 1999). These skills are associated with positive relationships with peers and teachers as well as increased peer acceptance (Walker et al., 2004). The school setting is an important platform for intervention as it is universally accessible to children (Wilson & Lipsy, 2007); therefore, students that have behavior and/or academic difficulties can be taught the skills necessary to improve in those areas (Anderson et al., 2008). A coordinated, schoolwide, multitiered, early identification and interventions system such as Response to Intervention (RTI) and Positive Behavior Interventions and Supports (PBIS) can help facilitate the improvement of academic and social skills. Such systems are recommended as best practice to increase appropriate behaviors and decrease disruptive behaviors (O'Shaughnessy et al., 2003; Sugai & Horner, 2002).

Multi-Tiered Supports

Amongst individuals, there are a variety of needs that can worsen if interventions addressing these needs are not provided. If needs are identified and addressed early, it is possible that more significant problems will be prevented later on. This is one of the main goals of the public health model, which was developed to provide supports to all individuals, and not just those with significant needs (Whitcomb & Merrell, 2013). Within this model, individuals receive preventative and early intervention supports when lower level of needs are present. More intensive supports are provided to individuals with higher needs. Within this model there are typically three levels of supports: 1) primary, 2) secondary, and 3) tertiary. The primary level of support is provided to all individuals and is intended to prevent future needs. The secondary level of support is for those who are at-risk for developing problems in the future. This level of support is provided in order to prevent the current difficulties from intensifying. The tertiary level of support is provided to those with a higher risk of developing problems or have already developed significant problems. The goal of this level of support is to intervene and provide individualized supports to reduce the intensity of existing problems (Whitcomb & Merrell, 2013).

To meet the needs of all students in the school setting, the public health model was adapted to address academic and behavioral difficulties. The Multi-Tiered System of Support (MTSS) is defined as a model in which data-based problem solving is used to integrate behavioral and academic interventions into instruction and provide interventions based on student needs (Whitcomb & Merrell, 2013). MTSS is an integration of two service delivery models, Positive Behavior Supports (PBS), often also called Positive Behavior and Intervention Supports (PBIS), and Response to Intervention (RTI). Through the MTSS model, similar to RTI and PBIS, students receive academic, behavioral, and social emotional instruction on a tiered level depending on their needs. PBIS emphasizes a prevention-targeted approach that is focused on identifying potential student behavioral needs and implementing strategies in order to prevent future problem behaviors (Sugai et al., 2016). The PBIS framework includes four components. The first component is predicting potential problems, defining what these problems will look like, when these problems may occur, and under what circumstances they occur. The focus of the first component tends to be on common behavioral issues that are present in a school that may easily be addressed and prevented with the appropriate planning and support. In the second component, rules, routines, and physical restructuring occurs. These routines and physical arrangements include the development of antecedent (i.e., what occurs prior to the behavior) and consequence (i.e., what occurs after the behavior) strategies. Physical arrangement may include how desks are placed, where the adults are placed, and how the students are arranged depending on the activity which they are engaged in. Nevertheless, these routines and arrangements should be manageable and tailored to the teacher and setting that the students are in.

An important component of PBIS is teaching students the rules and routines to develop awareness of the expectations. The third component involves consistent implementation of rules, routines, and physical arrangement. In this component the teacher should create rules that are appropriate for the students' behaviors and what behaviors are expected for the student to be successful. Establishing routines and physical arrangement of the classroom increases the likelihood that the rules are followed. Another important part of the third component is consistent teacher feedback. The teacher must provide the students positive feedback, prompts, and reminders for the expected behavior to continue. Finally, the fourth component of PBIS is evaluating the student's progress to determine if the rules, routines, and physical arrangement are successful in reducing the problem behavior. To determine progress, goals for student's behavior should be set. Next, behavioral data are collected and outcomes are monitored to examine if the student is meeting the goals and if the problem behavior has decreased. Consistent monitoring of behavioral data is necessary to inform if the intervention is successful or if changes need to be made (Scott et al., 2007).

PBIS is a framework that can be implemented at a school-wide as well as at a class-wide level. School-wide PBIS promotes a positive, safe, and productive learning environment in which rules are taught and positively reinforced. Student progress is monitored and changes to the interventions and supports are made based on the data collected (Scott et al., 2007). The PBIS system includes positively stated behavioral expectations, teacher-directed instruction, and reinforcement systems that reward appropriate behavior (Sugai et al., 2000). In using this system, levels of support vary depending on student needs. This entails using behavioral procedures to decrease the unwanted behaviors and providing tertiary interventions that are specific to the function of the student's behavior (Scott et al., 2007).

PBIS at the classroom level functions similarly to the school-wide level and includes the features of prediction, prevention, consistency, and evaluation to decrease problem behaviors and improve academic and social success. While school-wide PBIS focuses more on establishing expectations for all students, class-wide PBIS allows for the considerations of the unique needs in the classroom. In the classroom, teachers are able to target a small group or individual students. For example, if a student has difficulties attending to task, the teacher may provide the student verbal prompts to pay attention. Using the PBIS model, the teacher instructs the student on how to attend to task and provides reinforcement and praise when the student is being successful. Other strategies may include pre-corrections (Scott et al., 2007), which are teacher-directed activities that occur before a student enters in a situation that has previously been

associated with a problem behavior. In other words, when teachers use precorrections they provide students reminders of appropriate behaviors prior to the student entering in a situation when problem behaviors were present in the past (De Pry, & Sugai, 2002). An additional strategy includes desk and student arrangement, in which teachers arrange the classroom or position students to minimize distractions and increase productivity (Scott et al., 2007). Visual supports are an additional strategy, which typically include images, drawings, or photographs to represent tasks, needs, goals and rewards (Cramer et al., 2011).

Academic Response to Intervention (RTI) is a similar approach to PBIS that is based on the notion of changing the intensity of academic interventions depending on student needs. Through the academic RTI model, student needs are assessed through screening. Those who demonstrate some academic difficulties are identified and provided with interventions to help improve academic skills, intervention progress is monitored, and changes are made to the intervention based on their progress (Jimerson, Burns, & VanDerHeyden, 2007). Although RTI has been used in the literature as both a behavioral and academic service delivery model, presently it is described most often as an academic-only service delivery model. The overarching goal of RTI is to improve school-wide and individual achievement through high-quality universal instruction and increasingly intensive tiered supports that are provided in response to the students' needs (McIntosh & Goodman, 2016). Through this model students do not have to wait to be identified as a student with a disability to begin receiving an intervention. Rather, schools are able to provide sound instructional methods to meet the student's individual needs in the general education setting.

The RTI model includes three tiers of intervention. The first tier, which is designed to meet the needs of about 80% of students, is the comprehensive and universal tier. This type of

instruction is available to all students in each grade level and is based on the assumption that providing strong, effective, classroom instruction will be sufficient for students to be successful. The second tier, which includes approximately 15% of students, is available to those who are not successful in the first tier and need more support to be make adequate progress in the classroom. Tier 2 interventions are provided in conjunction with tier 1 instruction and are often group-based interventions that target specific skills. The third tier includes a small number of students, approximately 5%, and is used for students who did not respond sufficiently to tier one and two interventions. These interventions may be more individualized to target specific skills that require remediation.

Within RTI, there are two paradigms that address how and what types of interventions are to be implemented, the standard protocol and the problem-solving model. When using the standard protocol, schools implement scientifically validated interventions and examine growth by progress monitoring. This type of method allows for consistency of implementation across staff. Nevertheless, there are some disadvantages to using this approach. For example, these types of interventions are often expensive. Another possible disadvantage to the standard protocol may be the inability to tailor the intervention to the student to address their specific concerns.

The problem-solving approach is similar to the standard protocol approach in which scientifically validated practices and progress monitoring are used. However, the problemsolving approach allows for individualization of interventions that are specific to the student needs. Specifically, students are matched with an intervention that is designed to address their identified difficulty. Nonetheless, the problems solving approach also has disadvantages. While students are making some progress on the intervention, the teacher may consider that the progress is insufficient, as they are not improving to a level the teacher deems appropriate. This may be a result of teacher training, tolerance, and resources. In other words, teachers may not have the adequate training or enough time to monitor the student's progress with fidelity. Additionally, the teacher may have a misunderstanding of the rate of progress that the student should be making in the intervention. Each of these factors may result in an inaccurate perception of the student's progress (Hale et al., 2006).

Although providing behavioral supports in PBIS and academic supports in RTI is important, integrating the two models to support academic and behavioral and social emotional needs into a one system is crucial (McIntosh, Bohanon, & Goodman, 2010). Considering the documented relationship between academic skills and problem behaviors (McIntosh & Goodman, 2016), evidence suggests that combining RTI and PBIS is associated with a greater increase in academic and behavioral improvement (Eagle et al., 2015). Moreover, combining the two approaches could provide a more seamless service delivery in the school rather than having two separate systems. Students with both behavioral and academic challenges may not receive the required support if both components are not well integrated. Additionally, as RTI and PBIS share common theoretical backgrounds, schools that have experience implementing either system can utilize the knowledge of their current system in combining them to implement MTSS (McIntosh & Goodman, 2016).

Progress Monitoring

An important component of RTI, PBIS, and MTSS is data-based decision making. This includes collecting data as the students are going through the interventions and making decisions about whether supports need to be changed or intensified based on the data (Brown-Chidsey & Steege, 2011). Progress monitoring is a form of data-based decision making and continuous

progress monitoring has been found to be an important component of PBIS, RTI, and MTSS (O'Shaughnessy et al., 2003; Wehby & Kern, 2014). Progress monitoring is described as a set of procedures that are used to determine if a student is making adequate progress as a result of the instructional program or intervention (Berkeley & Riccomini, 2011). This is useful to check on the student's progress and help make decisions about the continuation of the interventions or if changes need to be made to the supports that the student is receiving. Monitoring progress can also help teachers develop realistic and attainable goals, which may serve as a motivator for students to improve (O'Shaughnessy et al., 2003). Lembke and Sticher (2006) found that students with behavioral problems are more likely to be successful both academically and socially if their needs are identified and monitored.

Academic progress monitoring. To measure academic progress, two common progress monitoring tools are used: Curriculum-Based Assessment (CBA) and Curriculum-based Measurement (CBM). CBA are teacher-made tests that assess a student's skills development across the entire curriculum (i.e., general outcome measurement) or assess a more specific component of the curriculum (i.e., specific subskill mastery measurement). When developing a general CBA, a teacher may create a spelling test that samples from all the spelling words the students were expected to know over the course of the year. However, when creating a more specific CBA, a teacher may only choose specific types of words to assess the student's knowledge in a narrower area of skill (Hintze et al., 2006).

CBM are similar to CBA as they also measure student abilities, however there are some differences. CBM are standardized measurement procedures that are used to determine the student's general skills in broad areas such as reading, spelling, mathematics computation, and written expression. One of the differences between CBA and CBM is that CBA focuses more on mastery of short-term objectives, whereas CBM focuses on long-term broad objectives. Therefore, CBM allows for assessment of the student's retention and generalization of what they have learned. Lastly, one of the major differences between the two types of measures is the standardized nature of CBM. CBM have specific standardized procedures that are used when providing a CBM probe to a student. The standardized nature of CBM allows for comparison across time within a single student as well as comparison between students (Hintze et al., 2006). While this tool has been shown to be sensitive to student change over time, it is influenced by other variables which may impact its reliability. For example, reading probes can be impacted by the individual administering the probe, location of the probe's administration, and the type of curriculum used while the probe was developed (Hintze & Christ, 2004).

Behavioral progress monitoring. While academic skills are more concrete and standardized, behavior concerns are broader and need to be well-defined. Due to the many behaviors that students can engage in as well environmental or internal variables that may influence behavior, a variety of methods may be used to measure behavioral changes. Different dimensions of behavior can be measured, such as the number of times it occurs, how long it occurs, or the time it takes someone to engage in a behavior after an antecedent stimulus. The following behavioral methods of data collection are often easy to use and effective for measuring behavior directly, and these methods do not rely on teacher perceptions or recollection of behavior. There are many possible methods that can be used to measure behavior, one being event recording which is the count of how many times the behavior occurs during the observation period. Another means of measuring behavior is timing the duration of the is inappropriate. As the behavior discontinues, the timer is not reset and the total duration of

the out of seat behavior is accumulated at the end of the observation period. Another dimension of behavior that can be measured is response latency. In this method the duration between the onset of a stimulus to the initiation of a response is timed. A similar measure to latency is interresponse time (IRT), which is the time between the end of target behavior and the beginning of another instance of the target behavior (Cooper et al., 2007).

An additional way of measuring behavior is time sampling. There are several variations of time sampling that may be used depending on the type of behavior. With these methods, an observation period is divided into time segments (e.g., 10-seconds, 15-seconds). For whole interval recording the behavior must occur throughout the entire interval to be counted as having occurred. This type of time sampling usually underestimates the occurrence of the behavior during the observation period as the expectation is that the behavior occurs during the entire interval. For partial interval recording the behavior can occur during any part of the interval in order for it to count as having occurred. In contrast to whole-interval time sampling, this type of time sampling often overestimates the occurrence of the behavior as the behavior may have occurred for a small portion of the interval. Finally, momentary time sampling is another method to record behavior. Using this momentary time sampling the observation period is divided into intervals and at the end of each interval the observer looks up at the student and records if the target behavior is occurring. While an advantage to this method is that the observer does not have to consistently attend to the observed individual, some of the behavior that is observed may be occurring when the observer is not watching the target individual. However, if the intervals are short enough, these methods will more likely approximate the actual amount of behaviors occurring (Cooper et al., 2007).

Evaluating behavior change. While understanding how to measure academic and behavioral progress is important, it is also important to evaluate if meaningful change has been made due to the implementation of the intervention or program. One important component of determining if change is made due to a program is obtaining a baseline of the student's current skills or behaviors. The baseline functions as a control phase, where the amount of behavior occurring without the intervention in place is monitored. Establishing a baseline also aids in determining an appropriate goal for the student. In other words, understanding how much or how often the behavior is occurring can help school staff to better understand what a reasonable and attainable goal might be for the intervention they are developing. Otherwise, a goal may be set that is too difficult or unreasonably high for the student to achieve (Cooper et al., 2007). When measuring academic progress, a baseline can be obtained by administering a CBM prior to starting the program that will target the deficient skill (Berkeley & Riccomini, 2011). For example, a student's reading fluency is measured prior to the start of a reading fluency intervention to determine if the intervention improves their ability to read fluently. According to Jenkins et al. (2005), at least three data points are recommended to establish a baseline using CBM. When behavioral progress is measured it is recommended that the baseline data points are stable (i.e., show little variability or trend over observation periods) prior to beginning the intervention in order to obtain an accurate picture of the behavior prior to implementing the program. When the data path is not stable, it may imply that the behavior that is being measured is changing and it may be difficult to attribute any future changes to the implementation of an intervention. Thus, if stability has not been achieved, additional data points should be collected until an upward or downward trend is no longer present (Cooper et al., 2007).

Targeting On-Task Behavior

Given the challenges that teachers face with students who have behavioral needs, in addition to the importance of early intervention and prevention, reducing challenging behaviors using in-class methods is crucial. A body of research has shown that increasing on-task behavior may reduce problem behaviors (Hawken & Horner, 2003; McComas at al., 2002). Part of the rationale behind this is that if students are engaging in on-task behavior, they have less opportunity to engage in disruptive behavior. For example, a reduction in behaviors such as talking out during class, talking back to teachers, aggression, and refusal to follow directions resulted from implementing an intervention that targeted increasing on-task behavior (Hawken & Horner, 2003). Another study found that an intervention that targeted decreasing student off-task behaviors such as drawing on the desk, destroying school materials, and talking to peers also resulted in a decrease in problem behaviors (McComas at al., 2002). Moreover, Peterson et al. (2005) examined the effects of an intervention intended to increase task engagement of two participants with inappropriate behaviors such as flopping on the ground, leaving when the student was asked to complete a task, throwing objects, and refusal behaviors. Their results showed that as student task engagement increased it corresponded with a decrease of the target behaviors. These findings suggest that targeting on-task behavior improves a variety of disruptive behaviors in the classroom. Therefore, interventions that target increasing student ontask behaviors are important in the reduction of challenging behaviors.

While there are various classroom management strategies to increase desired behaviors like remaining on-task, group contingencies are a set of classroom management strategies that have been shown to be successful in reducing unwanted behaviors. Group contingency interventions have been studied across a range of educational contexts, behaviors, and
populations (Theodore et al., 2004). Stage and Quiroz (1997) identified group contingencies as being the most effective management strategy available to educators.

Group contingency interventions are defined as a management system in which multiple students are expected to meet a specific set of behavioral expectations in order to earn a reward (Heering & Wilder, 2006). In other words, the reward is contingent on one or more students in the group meeting the behavioral expectations. According to Maggin et al. (2012) the effectiveness of the group contingency intervention lies in the use of peer influences to encourage others to engage in the appropriate behaviors and the ability of the interventionist to deliver positive and negative consequences. Group contingencies can also aid in managing the behavior of an individual or a group of students and may be used to target specific behaviors as well as continuously motivate students to attain the reward (Theodore et al., 2004).

These interventions are cost-effective, time efficient, easily implemented, and highly acceptable to teachers and students. They have been used to decrease noise level, improve disruptive and off-task behaviors, and improve academic performance (Little et al., 2015). According to Litoe and Pumroy (1975) there are three types of group contingencies: dependent group contingency, independent group contingency, and interdependent group contingency.

Dependent group contingency. In a dependent group contingency, the consequence or reward is contingent on the performance of one or more individuals that were selected by the teacher. The performance of non-selected group members is irrelevant in determining if the contingency is met. This type of group contingency is useful in decreasing externalizing behaviors in select individuals (Hansen, S. D. & Lignugaris, 2005). However, this type of intervention may be unfair to other students as they may lose an opportunity for reinforcement because of the actions of other students (Romeo, 1998). Students may also choose to sabotage

the class's ability to earn the reward. Another disadvantage is that if the students know who is responsible for them not earning the reward, the other students may tease or become angry with that student. Additionally, it makes it difficult to target the behaviors of all the individuals within a group and teachers may feel that it diminishes the responsibility for each student to engage in the appropriate behaviors that would earn them the reward (Brantley & Webster, 1993).

Independent group contingency. In an independent group contingency, each individual within a group receives reinforcement based on whether their individual performance satisfies the contingency set for the group. This type of group contingency system is more useful in targeting all individuals within a group rather than one or group of individuals (Litoe & Pumroy, 1975). There are, however, some disadvantages that are associated with independent group contingencies. For example, this type of group contingency does not incorporate cooperation between group members because access to the reinforcer depends only on individual performance. This is a disadvantage as promotion of cooperation amongst students may help in the development of social skills (Winn, 2006). Additionally, teachers must plan other activities for students who do not earn the rewards. Another weakness of this type of group contingency is that the teacher has to set an appropriate criterion for the reward. A failure to do so may result in few students earning the reinforcer, resulting in student frustration and a lack of motivation to earn the reward. Students may also become reliant on being extrinsically motivated by the reward rather than intrinsically motivated to engage in the appropriate behaviors (Winn, 2006).

Interdependent group contingency. In an interdependent group contingency, a group effort is required in order to meet the contingency specification and receive the reinforcement. In other words, all group members need to meet the expectations in order for any member of the group to access reinforcement (Litoe & Pumroy, 1975). This type of group contingency was

found to be the most commonly used amongst the three different group contingency types (Little et al., 2015). As the responsibility of receiving the reinforcement is shared amongst the group members, the likelihood of the reinforcement or lack or reinforcement being associated with a single individual is decreased (Skinner et al., 1996). Therefore, this type of system allows for the impact on behavior of all class members. While there are a number of advantages to this type of contingency system, there are some disadvantages. Due to the expectation that all group members are to engage in the behavior that is contingent on the reinforcement, students who meet the behavioral expectations may not receive reinforcement as a result of the behavior of the other group members (Litoe & Pumroy, 1975). Additionally, similar to the dependent group contingency, students may intentionally behave in such a way that prevents the other students from earning the reward. Also, students may become angry with the student that caused the group to lose the reward (Winn, 2006). Nevertheless, it is recommended that students that are not cooperating or do not meet the group expectations have their own group or contingency, which prevents the student from sabotaging the group (Skinner at al., 1996).

Interventions using Interdependent Group Contingencies

While all three group contingency types have been shown to decrease unwanted behavior and improve academic performance, the most commonly used group contingency type is the interdependent group contingency (Little et al., 2015). There are a number of programs that have utilized an interdependent group contingency to improve behavior in the classroom. Several of these are discussed below, including the Good Behavior Game, PAX Good Behavior Game, and Class-Wide Function-Related Intervention Teams.

Good behavior game. One of the most widely examined interdependent group contingency intervention is the Good Behavior Game (GBG, Barrish et al., 1969). In studies that

have implemented the GBG, teachers divided their class into teams comprising two to three students. Team names, class rules, and behavioral expectations were established, defined, and posted. During the intervention period, the teacher placed a mark under a team name when a member of the team violated a rule. Teams were able to win by either having the fewest marks or if they remained below a preestablished criterion number of marks (Barrish, et al.,1969; Tingstrom et al., 2006). According to a review of studies that utilized the GBG from 1969 to 2002, this intervention has been commonly used to decrease disruptive behaviors such as talking and out of seat behaviors, increase work completion, and decrease inattentive behaviors, name calling, cursing, and verbal/physical aggression in students first through sixth grades (Tingstrom et al., 2006).

PAX good behavior game. A variation of the GBG is the PAX Good Behavior Game (PAX GBG; Embry, 2002), which is a universal classroom-based preventative intervention designed to improve social-emotional competence and behavior. In this intervention, students create the rules and the vision for the class using lessons that are used as part of the intervention. "PAXIS" is a made-up word that is defined as productivity, peace, health, and happiness. An additional made-up word, "spleems", are the unwanted behaviors that get in the way of achieving PAXIS. This language is designed to foster positive debriefs with students, avoid negative attention, reduce the verbally inflected emotionality that accompanies a reprimand when a rule is broken by a student, and increase generalization (Embry, 2002; Embry & Biglan, 2008). Similar to the GBG, when the intervention is in place the teams work to maintain appropriate behavior in the classroom and points (i.e., "spleems") are given to teams when a member of the team displays inappropriate behaviors. Unlike the GBG, PAX-GBG includes components called "kernels" that are designed to improve compliance, reduce disruption, increase student

attendance, reduce transition times, and improve classroom management using tools like a timer and nonverbal cues (Domitrovich et al., 2010).

Class-wide function-related intervention teams. Class-wide function-related intervention teams (CW-FIT; Wills et al., 2010), an interdependent group contingency intervention, is another variation of the GBG. Unlike GBG and PAX GBG, CW-FIT focuses only on tracking positive behavior rather than negative behavior. CW-FIT is a function-related intervention, in that the group contingency addresses attention, which is a common function of problematic behavior (Kamps et al., 2011). CW-FIT's primary focus is to directly teach and reinforce student prosocial classroom behaviors through group contingencies within the classroom environment. According to Wills et al. (2010), this intervention can be used as part of the MTSS process as it includes appropriate skills that are expected in the classroom, positive reinforcement, and allows for increased supports that can be provided to students that are not responding to the general intervention. CW-FIT is used in the present study. Below is a description of the main components of CW-FIT.

Teaching. In this component the students are taught appropriate communication skills to use when obtaining attention from teacher and peers, gaining access to activities, and escaping demands in an appropriate manner. Students are taught how to get the teacher's attention, follow directions, and ignore inappropriate behaviors. As each skill is introduced, a poster that outlines the skill is placed in the classroom in a location that is visible to all students. Each lesson includes the review of the skill poster, modeling of examples and non-examples of the behavior, and the chance for students to practice the skill and receive feedback. Each lesson lasts approximately 10 to 15 minutes. These skills are taught one time and reviewed briefly at the beginning of every CW-FIT session (Wills et al., 2010).

Extinction. This component is intended to eliminate or minimize social reinforcement of unwanted behaviors. Throughout the intervention, teachers recognize and reinforce students that are appropriately gaining their attention and withholding attention from students who attempt to obtain their attention inappropriately (Wills et al., 2010).

Reward. During the implementation of CW-FIT, students earn points for their teams by engaging in the desired behaviors. A chart is posted in the classroom, which displays a list of the skills and the points that each team has earned. The teacher has an option of putting individual students on their own teams if their behavior is more significant or they are sabotaging the group. At the end of every CW-FIT session, the teacher tallies the points and teams that reached the predetermined amount of points receive a reward. As part of the intervention, the teacher is provided with potential reward options to be used during the intervention (Wills et al., 2010).

Additional tiers of support. Additional supports can be provided for students who are not responding to the primary components of CW-FIT. One option includes a self-management component. In this component, a student is provided their own chart that is similar to the class chart. As the teacher awards points to the teams, they will also award points to the students who are using the self-management tool if they are engaging in appropriate behaviors. An additional second tier includes help cards, which are utilized for students that avoid academic demands or seek teacher and/or peer help. While using this method, students are provided with cards that can be used when teacher or peer help is required. Students receive a predetermined number of cards and the number of cards is systematically decreased over several sessions when the student becomes more aware of when help is required. A third tier of the intervention is also available in which a functional assessment is used for students who do not respond favorably to the second tier. This assessment provides additional information to the teacher about the function, or

motivation behind the student's behavior in order to target and support them more individually (Wills et al., 2010).

Intervention. At the start of every intervention period, the teacher uses the posters to review and remind the students of the expected behaviors. At that time, the teacher also establishes a goal and the reward for the day. During the intervention period, the teacher sets a timer for every 2-3 minutes and as the timer stops, the teacher pauses to praise and award points to the teams in which all the students are engaging in the appropriate behaviors. The teacher will also provide reminders to teams that were not engaged in the appropriate behaviors. At the end of the intervention period, the teacher publicly tallies up the points and provides the reward to the teams that met their goal (Wills et al., 2010).

Theoretical Orientation

The theoretical foundations of CW-FIT are primarily in behaviorism, which is the scientific approach to observing and analyzing environmental events that may influence the occurrence of behavior (Cooper et al., 2007). Behaviorism posits that behavior is developed and maintained through the interaction between an individual and the environment (e.g., antecedents and consequences) rather than being caused by hypothetical or unobservable internal mental states (Gresham et al., 2001). The acquisition of behavior occurs through conditioning (Cooper et al., 2007). While there are several types of conditioning, operant conditioning is the primary mechanism used in CW-FIT for developing and maintaining behavior.

Operant conditioning is the process in which the behavior is developed and maintained due to the consequences that occur after the behavior. The consequences that occur after the behavior or relatively soon after the behavior changes the probability the behavior will occur under similar circumstances in the future. According to the concept of operant conditioning, there are two ways to change behavior, reinforcement and punishment. Reinforcement occurs when consequences increase the probability a behavior will occur in the future, and punishment occurs when consequences decrease the probability a behavior will occur in the future (Cooper et al., 2007). CW-FIT uses reinforcement to help develop and maintain positive classroom behaviors.

There are two types of reinforcement, positive reinforcement and negative reinforcement. Positive reinforcement occurs when a behavior is followed by an addition of a stimulus which results in the increase of the behavior's frequency in the future. In CW-FIT this type of a reinforcement is used in three different ways. First, the teacher is expected to verbally praise and highlight appropriate behaviors that the students are engaging in. This type of praise and prompt is designed to help students learn which behaviors are appropriate. Second, teachers provide points to teams of students, where all students must engage in appropriate behaviors to receive points (Cooper et al., 2007). Third, only the teams that reach the pre-determined number of points receive a reinforcer at the conclusion of the game. This reinforcer is intended to increase the frequency of the appropriate behaviors that students are engaging in (Cooper et al., 2007).

In addition to consequences, antecedents are also important in behavior change. Antecedents are stimuli or events that precede the behavior (Gresham et al., 2001). In the current study, CW-FIT served as a *discriminative stimulus*, which is an antecedent stimulus associated with the presence of reinforcement (Gresham et al., 2001), during which the students understood that when their team is engaging in appropriate behaviors, they will receive a reinforcement. CW-FIT was only used during a certain class, and the teacher was clear when it was being used. Thus, starting CW-FIT was an antecedent stimulus indicating to students that specific behaviors

that are defined by the game, would be reinforced through praise, points, and a specific reinforcer at the end of the class.

Review of Studies that Utilized CW-FIT

Research has explored the effects of CW-FIT on increasing on-task behaviors and teacher praise, as well as reducing disruptive behaviors and decreasing reprimands across grade levels from preschool to seventh grade (Caldarella et al., 2015; Conklin et al., 2017; Jolstead et al., 2017; Kamps et al., 2015; Kamps et al., 2011; Wills et al., 2018). Their findings showed that not only class-wide on-task behavior improved, but disruptive behaviors of individual students who were nominated as at-risk for emotional and behavior problems also decreased. For example, 10-minute observations showed a decrease from a mean frequency of 18.2 disruptive behaviors during baseline to a mean frequency of 5.2 during intervention (Kamps et al., 2011). Additionally, these studies showed a consistent reduction of teacher reprimands and increase in praise.

In addition to on-task behavior, teacher reprimands/praise, and disruptive behaviors, Conklin et al. (2017) examined behaviors that are taught within the CW-FIT protocol such as compliance, hand raising, remaining seated when expected, and talking out within kindergarten, second grade, and two seventh grade classrooms. Their findings showed an increase in compliance and hand raising as well as a decrease in out-of-seat behaviors and talking out when it is not appropriate across all grade levels. Furthermore, in a dissertation, Conklin (2010) examined the effects of classroom work completion and grades in two seventh grade classrooms as incidental effects of the intervention and found that assignment completion and grades were not impacted by the introduction of CW-FIT. Gender differences were explored in one study and were not found for student on-task behavior, teacher praise to group, teacher reprimands to individuals, and teacher reprimands to group during the baseline phase. However, as the intervention was implemented there were significant gender differences in student disruptions and teacher praise to individuals, where males had significantly more disruptions than females (Wills et al., 2016). These findings are consistent with other literature which suggests that females are less likely to have disruptive behaviors in the classroom than males (Martinez et al., 2016). Additionally, the results showed that teachers praised male students more frequently than female students (Wills at al., 2016).

While the aforementioned studies have examined the effects of CW-FIT during one part of the day, Wills et al. (2014) implemented the intervention in a first-grade classroom three times per day to examine the intervention effects across the day. Similar to the findings above, an increase in on-task behavior, teacher praise, as well as a decrease in reprimands and disruptive behavior was observed. These results occurred throughout the three periods when the intervention was implemented.

CW-FIT was found effective in core classes, and the effects of the intervention have also been examined in classes that have a different structure, pacing, and requirements such as physical education (PE, Hirsch et al., 2016) and music class (Caldarella et al., 2017). In both studies an increase in on-task behavior and teacher praise as well as a reduction in reprimands was found compared to the baseline condition (Caldarella et al., 2017; Hirsch et al., 2016). Nevertheless, on-task behavior was slightly lower in the PE class and only reached 80.0% or above during 5 out of 14 sessions. The authors hypothesized that the lower level of student engagement may have been due to the components of the intervention that were not included such as the students were only taught two of the expectations and the tier 2 and 3 interventions were excluded. Additionally, the authors noticed that the highest levels of student engagement were associated with the highest intervention fidelity. Therefore, the fidelity implementation may have affected the levels of student engagement (Hirsch et al., 2016).

Another difference was discovered in the teacher praise and reprimand ratio in one music classroom. The authors found that teacher praise to reprimand ratios increased from baseline to the intervention phase (from 1.7 to 4.5 per class). However, when the intervention was reintroduced after the reversal this ratio was lower (2.5 per class) than when the intervention was introduced initially. The authors explained that it is possible that the more frequent use of praise was less necessary as student behavior improved as a result of learning and using the social skills and working for group points and rewards (Caldarella et al., 2017). A similar difference was found by Hansen et al. (2017) in which they implemented CW-FIT in a French dual immersion program. While they also found an increase in on-task behavior during the intervention phases, they similarly found that teacher praise decreased when the intervention was re-introduced compared to the initial introduction of the intervention following a reversal in two of the classes.

In addition to general education classrooms, CW-FIT was also implemented in a special education classroom that serves students with disabilities that have social emotional and behavioral needs. In this study, the authors intended to implement a single subject withdrawal design, but due to the strong effects of the intervention, the teacher asked that a withdrawal phase not be implemented. A compromise was reached; brief withdrawals of one day would be implemented. Like previous studies, the authors found a large increase in on-task behavior, from 54.0% during baseline to 87.0% while the intervention was implemented. Overall, percent of intervals on-task was 55.0% across withdrawal phases. Additionally, teacher praise increased and reprimands decreased when the intervention was implemented (Weeden et al., 2016).

While some of the studies listed that they used tier 2 strategies (i.e., student engagement and/or help cards) as part of the intervention, to the author's knowledge only one study evaluated the effects of using the component of student engagement and help cards on student engagement. Wills et al. (2016) found that the students who were chosen to be included in the tier 2 intervention had mean on-task behaviors of 63.7% during the baseline phase, which increased to 80.8% during the CW-FIT intervention and increased to 86.0% when self-management was introduced. Similarly, on-task behaviors for students receiving the help cards intervention and to 85.8% when the help cards were added to the intervention.

In addition to these positive effects, teachers that have implemented the CW-FIT intervention have responded positively to the intervention and commented on the ease of implementation. Students also enjoyed participating in the intervention. Additionally, across the studies, intervention fidelity was rated above 80.0% (Caldarella et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2014), which is considered acceptable (Cooper et al., 2007). To measure fidelity, studies used a form to rate which components of the intervention were present, such as posters posted for students, reviewing of skills at the beginning of lesson, and providing points and praise (Caldarella et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Wills et al., 2014).

Gaps in the Literature on CW-FIT

CW-FIT has been replicated across different grades, content areas, and has been found to increase on-task behaviors, teacher praise, student compliance, hand raising, remaining seated when expected, as well as decrease in teacher reprimands and students talking out when it is inappropriate. Although one study has examined the effects of CW-FIT as it was implemented several times per day (Wills et al., 2014), to the author's knowledge no studies have explored the effects of the intervention on other parts of the day when the intervention is not actively implemented. Generalization is an important response to an intervention, in which the occurrence of the desired behavior proves durable over time, appears in a variety of settings in which the intervention is not being implemented, and may possibly spread to other related behaviors (Stokes & Baer, 1977). In other words, generalization is attained when the behavior change that occurred in the setting with a treatment condition is also present in another setting when the treatment is not applied. While successful behavior change in one setting due to an intervention is promising, it is not enough. Eliciting a behavioral change in one setting may not help the student use the skills that they learned in other settings. For example, if a student learned how to turn in her homework in math class but this skill did not generalize to science class, the student will require an additional intervention to increase the desired behavior. Therefore, it is important to generalize the skills taught in an intervention to allow for success in other settings and over time. In a review conducted by Stokes and Baer (1977) they provided guidelines to increase the possibility of generalization. First, they noted that students should practice the skills outside of the treatment and these skills should be reinforced by others in a non-treatment setting. Additionally, they reported that it is important to train peers to demonstrate expected behaviors. Other guidelines included decreasing the number of reinforcers over time for the expected behavior to elicit social reinforcement, increasing the unpredictability of the reinforcement, using stimuli that is found in other settings as part of the treatment, providing reinforcement to students when they report desired behavior and promote self-monitoring of desired behaviors, and lastly if generalization of behaviors occurs intermittent reinforcement should be used.

Additionally, factors such as attendance, office discipline referrals, and student engagement being incidentally impacted by the intervention have not been explored. Attendance appears to be an important component that impacts student achievement and behavior at school. Students who do not attend schools as frequently lose instruction time, perform more poorly on exams (Gottfried, 2010) and standardized achievement tests (Nichols, 2003), and have overall poorer grades (Devadoss & Foltz, 1996). Additionally, lower attendance rates are associated with increased risk of student drop out and students engaging in delinquent behaviors (Sheldon, 2007). While there are various factors that may influence attendance, student engagement (i.e., student feelings of belongingness and attachment and involvement in school-related tasks such as homework completion, classroom participation, and compliance with school rules) has also been linked to attendance and drop out (Appleton et al., 2008; Archambault et al., 2009). While no studies were found that have utilized an interdependent group contingency intervention to increase student attendance, previous findings do show that structure, positive reinforcement, fair consequences, and clear rules are associated with increased student motivation (Williams & Williams, 2011). As these components can be included as part of CW-FIT, it is possible to speculate that the use of these types of interventions may be related to increasing student attendance.

Furthermore, office discipline referrals are more common in students who are less engaged, have more absences, and exhibit more problematic behaviors (Weber, 2011). Office discipline referrals are defined as a staff member's response to student behavior, followed by a response by an administrator that aligns with the values and expectations defined within the school system (Irvin et al., 2004). However, office discipline referrals have not been targeted using group contingency interventions. Nevertheless, past findings have shown that when expectation are taught and reinforced, students are pre-corrected and reminded of the appropriate behaviors, and appropriate behavior rewarded, rate of office discipline referrals decrease (Taylor-Green et al., 1997). Additionally, interventions that target disruptive behaviors have shown a decrease in office discipline referrals. For example, Hawken et al. (2007) examined the effects of a check-in, check-out intervention in elementary-aged students. During the intervention students checked in with a staff member at the beginning and end of the day, received feedback for their behavior, were positively reinforced, and set daily goals. Results from this study showed a 25%- 51% percent reduction in office discipline referrals. Therefore, it can be hypothesized that as disruptive behaviors decrease, office discipline referrals will also decrease as part of implementing an intervention like CW-FIT.

Student engagement has also been found to be an important contributor to student success (Skinner et al., 2008). Fredricks et al. (2004) conducted a review of student engagement studies and indicated that researchers have divided student engagement into three components: Behavioral engagement, emotional engagement, and cognitive engagement. Behavioral engagement is defined as positive student conduct, behaviors that are associated with learning such as participating in discussions, putting forth effort, and participating in school-related activities. Emotional engagement is defined as student affective reactions (e.g., anxiety, interest). Finally, cognitive engagement is defined as the student's psychological investment that is required in mastering of content. Fredricks et al. (2004) has also reported that positive engagement is related to academic achievement in elementary, middle, and high school students. Additionally, student engagement has shown to be a predictor for drop out. Students who typically drop out tend to complete less homework, generally give less effort, do not participate in school activities, do not feel belonging to the school, and have more behavioral problems at school which result in more frequent discipline efforts. Past research has shown that schools that provide students choices, clear and consistent goals, and encourage student involvement in school policy have higher student engagement (Fredricks et al., 2004). Similar to school factors that have improved office discipline referrals, some of the factors that are present in schools that are related to positive student engagement are present as part of CW-FIT. Therefore, it can be hypothesized that classroom engagement may improve due to the implementation of CW-FIT.

Current Study and Research Questions

The purpose of the current study is to replicate the effects of CW-FIT in three elementary classrooms. Consistent with previous research, this study intends to examine individual on-task behavior, teacher praise, and teacher reprimands. As an extension of previous research, this study intends to determine if there are incidental benefits of the intervention, such as increased attendance, a reduction in office discipline referrals, and increased classroom engagement. Finally, as an additional extension of previous research, this study will explore the generalization of effects of CW-FIT to other parts of the day when the intervention is not implemented. The following research questions will be answered as part of the study:

- Can the effects of CW-FIT that have been found in previous research be replicated in new classrooms? Specifically, it is hypothesized that during the implementation of CW-FIT there will be: a) an increase student time on-task, b) an increase teacher praise statements, and c) a decrease teacher reprimand statements.
- 2. Do the effects of CW-FIT found in previous research generalize to the portions of the day the intervention is not used? Specifically, it is hypothesized that during times when CW-FIT is not implemented there will be: a) an increase in student time on-task, b) an increase teacher praise statements, and c) a decrease teacher reprimand statements.

3. Does the implementation of CW-FIT produce other positive outcomes for students not investigated in previous research? Specifically, it is hypothesized that after CW-FIT is implemented there will be: a) a reduction in office referrals, b) an increase in attendance, and c) be an improvement in classroom engagement.

Chapter III: Methods

Ethics and Recruitment

Approval ethics. The study was approved by the University of Kansas Institutional Review Board and permission to conduct research at the school district was obtained. Informed consent was obtained from teachers and parents of target students. Additionally, verbal assent was obtained from all students participating in the intervention. The approval by the University of Kansas Institutional Review Board for human subjects research is included in Appendix A.

Teacher recruitment. The teacher recruitment process occurred by the researcher directly contacting all teachers in the elementary school via email. The email included an explanation of the intervention, the purpose of the study, potential benefits, and what would be expected from them if they agree to participate. A consent form was provided to teachers that expressed interest in participating in the study.

Student recruitment. After the teachers consented to participate in the study, they were asked to rank order the students in their classroom based on the number and intensity of disruptive and off-task behaviors. Student ranking occurred after approximately three weeks of classroom instruction following the first day of school for students. Teachers were asked to rank students who were fully included in the classroom. Two students that had the most disruptive and off-task behaviors were chosen along with a typical student. The typical student was used for comparison purposes to help determine if the students with more disruptive and off-task behaviors improve to a level consistent with typical students in class. Only students who were atrisk for social emotional difficulties were individually monitored in addition to an average comparison peer as part of the study; however, every student in the classroom received the intervention. The teachers used a process similar to the *Systematic Screening for Behavior Disorders* (SSBD; Walker & Severson, 1990) to rank order the students based on their

externalizing behaviors. Specifically, the teachers ranked the highest risk student as number one, the next highest as number two, and so forth until all of the students in the classroom were ranked. Once ranked, two of the students ranked highest and one average student were targeted for data collection throughout the study. The students' parents were contacted in order to explain the study and determine if they were interested in their child participating in the study. A consent form was provided to the parents if they expressed interest in their child participating in the study. Nine consent forms were provided and all nine were returned. Following parent consent, the researcher acquired verbal assent from each student that took part in the intervention. Participation in the study was optional for all participants and they were able to withdraw at any time.

Participants and Settings

Participants in this study included three classroom teachers and three individual students from each of their classrooms. The study was conducted in a Title I public elementary school located in a midwestern suburban community. The elementary school has 481 students and includes preschool through fifth grades. The school has an English Language Learner (ELL) program and 65% of students receive free and reduced lunch. Race and ethnicity composition of the school as reported by parents was 1% Asian, 47% Hispanic, 7% African American, 41% Caucasian, and 4% indicated two or more races.

Teachers. Teacher 1 is a 40-year old Caucasian female third grade teacher with 17 years of teaching experience and a bachelor's degree. Teacher 2 is a 29-year old Caucasian female Kindergarten teacher with 6 years of teaching experience and a bachelor's degree. Teacher 3 is a 50-year old Caucasian female second grade teacher with 27 years of teaching experience and a bachelor's degree.

Students. All students in the classrooms participated in the intervention. Each teacher

had three individually targeted students, two at-risk students, and one peer comparison student.

Teacher 1 and 3's at-risk students were two males and the peer comparisons were female.

Teacher 2's at-risk students were one female and one male; the peer comparison was a female.

Six of the targeted students were Hispanic and the remaining were Caucasian. Three of the

targeted students were receiving ELL services. Table 1 further describes the student's

demographic information.

-		
Ta	ble	

Student Demographic Information						
	Teacher 1 (3 rd grade)	Teacher 2 (Kindergarten)	Teacher 3 (1 st grade)			
Student 1	Hispanic eight-year-old male. Spanish and English are spoken at home. This student has been attending school in the United States since the beginning of his kindergarten year.	Hispanic six-year-old male. Spanish is spoken at home. This student has been attending school in the United States since the beginning of his kindergarten year. He receives ELL services.	Caucasian eight-year- old male. English is the only language spoken at home.			
Student 2	Caucasian eight-year-old male. English is the only language spoken at home.	Hispanic five-year-old female. Spanish is spoken at home. This student has been attending school in the United States since the beginning of her kindergarten year. She receives ELL services.	Hispanic eight-year-old male. English and Spanish are spoken at home.			
Student 3	Caucasian eight-year-old female. English is the only language spoken at home.	Hispanic five-year-old female. Spanish is spoken at home. She has been attending school in the United States since the beginning of her kindergarten year. Receives ELL services.	Hispanic seven-year-old female. Spanish is spoken at home. Has been attending school in the United States since the beginning of her kindergarten year.			

Student Demogr	aph	ic Inform	ation
		1 ord	• `

The school has a positive behavioral support system in which students earn tickets for

appropriate behaviors and were provided regardless of the phase. No other specific positive

supports or other interventions were provided to the students during the duration of the study.

Some of the students in the classroom were receiving special education services, but none of the students who were identified as at-risk or peer comparison students were receiving special education services.

Research Design

For this study a concurrent multiple baseline design was used. This type of design was used to demonstrate a functional relation between the intervention and behavior change across different classrooms. Additionally, this type of design was used due to the ability to demonstrate a functional relation without requiring a withdrawal of the intervention. Teachers were able to continue implementing the intervention in all classes and at the same time being able to see if change in behavior was observed only when the intervention was implemented. In other words, this design was useful in determining if the behavior changed when and only when the independent variable (i.e., CW-FIT intervention) was applied to each classroom (Cooper et al., 2007).

More specifically, all teachers began with the baseline phase. During this phase the teachers taught their lessons without any intervention materials. On-task behavior data was used to determine when the intervention should be implemented. Baseline on-task behavior data was collected until the data were stable. In other words, teachers did not move from the baseline phase to the intervention phase if an upward or downward trend in on-task behavior was observed using visual analysis for all of the at-risk students. If the trend in on-task behavior appeared to be increasing or decreasing in the direction expected from the intervention, the teachers stayed in the baseline phase until the trend was no longer present and the on-task behavior appeared stable. It was planned that at least three on-task behavior data points would be collected during the baseline phase for Teacher 1. On-task behavior for both at-risk students was

examined to determine if a trend was present. After stability was reached in on-task behavior for both at-risk students during the baseline phase, Teacher 1 began implementing the intervention, the data was monitored for approximately one week to ensure stability and at least three on-task behavior data points were collected. As long as stability was reached for both at-risk students in Teacher 1's classroom during the intervention phase and Teacher 2 and 3's baseline on-task behavior data remained stable for both at-risk students in their classrooms, Teacher 2 began the intervention. As long as stability was reached in on-task behavior for both at-risk students in Teacher 2's classroom during the intervention phase and Teacher 3's baseline on-task behavior data remained stable for both at-risk students, Teacher 3 began the intervention.

Materials

Skill posters. Teachers were provided the following skill posters: 1) *How to Get the Teacher's Attention*, 2) *Follow Directions the First Time*, and 3) *Ignore Inappropriate Behaviors*. Each poster contained steps that students were expected to follow. The steps included in the *How to Get the Teacher's Attention* poster included the following: 1) Look at the teacher, 2) Raise your hand, 3) Wait for the teacher to call on you, and 4) Ask your question or give an answer. Steps included in the *Follow Directions the First Time* poster were: 1) Look at the teacher and listen, 2) Say OK in your head, 3) Do it now, and 4) Check back if needed. Lastly, the third poster *Ignore Inappropriate Behaviors* included the following steps: 1) Keep a nice face 2) Look away from the person, 3) Keep a quiet mouth, and 4) Follow directions and do your work. Posters were obtained from the CW-FIT teacher implementation manual.

Point chart. Teachers also received a point chart that was used during every CW-FIT session. The chart included the date of the session, what reward the students will receive if the goal is met, a location for the teachers to tally the points, total number of points per team, and the

goal amount of points. Additionally, it included a number that denotes each team and a space for the teacher to tally team points. The point chart was laminated and re-used by the teacher during every CW-FIT session. The point chart was modified from the point chart found in the CW-FIT teacher implementation manual. Dry erase markers were also provided to the teachers to write on the point chart.

Lesson scripts. A CW-FIT introduction script along with lesson scripts were provided to the teachers. The introduction script that was read to the students included the purpose of the intervention and overview of the components of the intervention such as lessons, teams, goals, and points. Lesson scripts were included for each skill taught (i.e., *How to Get Teacher's Attention, Follow Directions the First Time,* and *Ignore Inappropriate Behaviors*). The skill scripts included the definition of the expected behavior, rationale, role play, and review of the skill taught. Lesson scripts were obtained from the CW-FIT teacher implementation manual.

Reinforcer menu. A blank reinforcement menu was provided to the teachers. This was used to create a list of possible reinforcers for the students. Additionally, a sample reinforcement menu was provided with ideas for tangible reinforcers, reinforcers that are inexpensive, and nontangible reinforcers, and larger prizes. The sample reinforcement menu was obtained from the CW-FIT teacher implementation manual.

Timer. A timer that included a countdown function was provided to the teachers. The timer displayed both minutes and seconds. It also included a clip that the teachers could use to clip on to their clothes as well as a magnet, which allowed them to place the timer on the board.

Procedure

Teacher training. Teachers received training from the primary investigator prior to implementing the CW-FIT intervention. Two one-hour training sessions occurred. During the

first training session the teachers were presented with the purpose of the intervention and how the intervention is conducted. They were provided intervention scripts, posters, and timers. At that time, they were given the opportunity to practice and ask questions. During the second training session the teachers were provided a review of the intervention procedures, additional practice time, and another opportunity to ask questions. To ensure that the teachers understood the intervention procedures, at the end of the second training session the teachers conducted a mock CW-FIT session. The primary investigator used the fidelity form while the teacher was conducting the mock CW-FIT session to determine if all steps of the intervention were followed. If the teacher did not obtain 100% on the fidelity document, then feedback was provided to them and they were asked to re-do the mock intervention. All teachers were expected to obtain 100% on the fidelity document prior to implementing the intervention. Following the training sessions, the teachers were provided support from the investigator when requested. Teacher support included the primary investigator attending the first intervention session to provide assistance with the intervention if needed and additional feedback following sessions was provided to improve fidelity if needed.

If tier 2 strategies needed to be implemented with individual students, teachers would receive additional training. This training would be provided on an individual basis depending on the need for those strategies. In this training the teacher would have received information on the two options for tier 2 strategies (i.e., self-management and help cards). They would be provided an overview and purpose of the intervention and implementation materials such as point sheets for the students along with scripts.

Baseline. During the baseline phase the teachers were expected to teach their lessons as planned without using any intervention materials. Although the teachers received the CW-FIT

training prior to starting the baseline phase, they were asked to not implement any of the intervention steps (e.g., posters) at that time.

Identifying intervention period. Teachers were asked to choose between math and reading as the primary time in which the intervention was implemented based on the number off-task behaviors, student engagement, or which of the two required more structure. Teachers were asked to choose either math or reading because those subjects are taught on a daily basis. The time period with most teacher concerns was chosen for CW-FIT to be implemented. All three teachers identified math as being the time during the day that the intervention was implemented. Teachers implemented the intervention every day during the chosen time period. The intervention was not to be implemented when the teacher was absent or during a special event (e.g., classroom party). Each teacher implemented the intervention for a minimum of seven weeks.

CW-FIT Intervention Implementation. During the intervention the teacher taught the students the following appropriate classroom behaviors, including: 1) How to get the teacher's attention correctly, 2) Follow directions the first time, and 3) Ignore inappropriate peer behavior. These skills were taught by modeling examples and non-examples, student practice, and feedback from the teacher. Initially, the teacher taught each skill using a 15-minute class-wide lesson and after each skill was taught a poster was hung in the classroom to remind the students of the appropriate behaviors. During the initial introduction of each lesson, the primary investigator was present to assist and answer any questions the teacher may have. Following the initial instruction of each skill, the teacher reviewed each skill taught at the beginning of each CW-FIT session.

Prior to implementing the intervention, the teacher divided the class into teams comprising three to five students. The teacher differentiated teams using seating arrangements (i.e., seating students close to one another using desks or a seating arrangement on the floor). During the intervention the team members stayed in close proximity to each other and if one student from the team moved around the room during an activity, the rest of the team followed them. Next, a reinforcer menu was created by the teacher with the assistance of students in the class. A sample reinforcer menu was provided to the teacher by the researcher.

A goal number of points was determined daily by the teacher prior to starting CW-FIT. According to the CW-FIT teacher implementation manual, goal amount of points should be determined based on the session length and time interval. The manual states that the goal should be set to match to 75% to 85% of when the timer stops. For example, if the session length is 45 minutes with a time interval of three minutes, students would have the opportunity to earn as many as 15 points. In this case, an appropriate goal should be set to 12 points, which is 80% of the total points they could earn. The teachers used those guidelines to set their daily goal initially. However, as they implemented the intervention several times in their classroom they changed the goal based on the class's performance. The teachers were encouraged to increase the goal as groups began to consistently reach the goal to challenge the students. At the beginning of the intervention each teacher set the goal at 80%. The teachers increased the goal to an average of 93.3% by the end of the intervention.

The teacher used a timer in order to provide consistent attention to appropriate behaviors. When the timer went off, the teacher scanned each team, provided labeled praise for specific positive behaviors, and rewarded points to the teams in which every student was engaging in appropriate behaviors. As the points were earned, the teacher provided specific praise to the

group (e.g., "Team 4 gets a point because they are all looking at the teacher and following directions"). When a team did not receive a point, the teacher should have provided direct feedback to the group (e.g., "Team 2 does not get a point this time and needs to work on following directions the first time"). The point sheet was displayed so that it was visible to all students. Teachers were also encouraged to provide bonus points to teams at their own discretion. It was explained to teachers that bonus points should be used for teams that are either doing an exceptional job or for teams that have not earned many points and the teacher notices that the students on the team are following the rules before the timer goes off.

During the first day of implementation of the intervention, the teachers posted the first skill along with the team point chart in a location visible to all students. Next, they used a script to introduce CW-FIT to the students. The teachers taught the first lesson using a script, wrote team names and set a goal on the point chart, chose a reward, started the timer, and conducted the CW-FIT session. The teachers all chose to set the timer to go off every three minutes. When the timer went off, the teacher scanned the room and awarded a point each time to all teams that are following the rules. At the end of the session the teachers totaled the points and provided the reward for all teams who met the goal. During subsequent CW-FIT sessions, the teacher taught and/or reviewed skills before beginning the intervention.

Tier 2 strategies. Tier 2 strategies were available to students who were not responding to the class-wide intervention. The individually targeted students that continued to be below 80% on-task after the intervention was implemented with fidelity for three weeks would have received a Tier 2 intervention until they were consistently above 80% for a week. If students reached 80% or above for on-task behavior, this was an indication that adequate progress was being made

using the intervention. Teachers would have had two options to choose from for Tier 2 strategies, self-management intervention and help cards intervention.

Self-management cards. In the self-management intervention the student is engaging in goal setting, monitoring, and recording his/her behavior. During this intervention the student has a small point chart on their desk during the CW-FIT session. When the teacher timer goes off, the student's job is to decide if their behavior was appropriate, and if it was, the student would mark a point on their chart. If their behavior was not appropriate, they would not receive a point. At the end of the CW-FIT session, the teacher would collect the student's chart and give them a reward if they earned it. Students engaging in this intervention would remain on the same team, have the same point goal as the class goal, and earn the reward if their team earns the reward. However, a separate reward could be provided to the student if decided by the teacher.

Help cards. Another option for a tier 2 strategy was using help cards. This strategy is designed to teach students an appropriate alternative to becoming off-task when they are challenged by the assignment. This strategy consists of the student requesting individual assistance from their teacher or peer on the task they are expected to complete. The student has a pre-determined number of self-help cards for the CW-FIT session. The student's job is to determine when they need help and use the cards to signal the teacher or a peer to receive help. At the end of the CW-FIT session, if the cards were used appropriately a reward could be provided to the student.

Maintenance. Following the conclusion of the study the researcher conducted two follow-up observations to determine if the teacher continued to use the intervention and if the effects of the intervention were maintained. These were conducted after five and ten weeks following the conclusion of data collection in Teacher 3's classroom. At that time the researcher conducted an on-task observation of the target students and rate of praise and reprimands per minute was calculated. The researcher also asked the teachers if they continued to implement the intervention, and if they did not continue to implement the intervention they were asked to provide their reasoning for the decision.

Data Collection

In vivo data collection. Direct observational data were collected during this study. The study consisted of three phases, the baseline, intervention, maintenance phases. During each of these phases, data were collected during two times throughout the day, the intervention class period and generalization class period. During the intervention phase data were collected two times per week during the intervention class period and one time per week during the generalization period in each classroom. Each observation lasted 20 minutes and was completed by the primary investigator.

On-task behavior. On-task behavior data was collected in vivo during intervention and generalization class periods using direct observations. Data collected during a time when the intervention was not being conducted was used to determine if the effects of the intervention generalized to other parts of the day when the intervention was not implemented. Because all teachers chose math for implementing CW-FIT, the generalization observation period was during the reading instruction. During the baseline phase, data were collected during the class period that the teachers implemented the CW-FIT intervention (i.e., math) as well as during the generalization period (i.e., reading).

The researcher used momentary-time sampling to observe the identified students and looked up at the students every 20 seconds and recorded if they were on-task. On-task behavior was defined as the student engaging in the expected activity such as actively working on an assignment, looking up at the teacher or the board while they were instructing, looking at a student while they were talking during a classroom discussion, and following teacher directions. During transition between activities, on-task behavior was defined as following teacher expectations during that time (e.g., sit quietly, talk with your table group, gathering materials).

Prior to completing the observations the primary investigator identified a consistent sequence for observing the students. For example, at the end of each 20 second interval the observer looked at student 1, then student 2, then student 3. Those students remained in the same sequence for the duration of the study. Each observation lasted for 20 minutes. A paper and pencil form was used to record student behavior. If a student was engaging in on-task behavior a "+" was marked and if they were not a "-" was marked on the observation form. The researcher listened to a recording while conducting the observation to know when to observe each student. The recording included an audible beat for each second, and a voice was recorded which notified the observer which interval they would be starting, a five second count down to the upcoming interval, and the recording notified the observer when to look at each student. For example, five seconds before the one-minute interval began the observer heard the following: "one minute, five, four, three, two, one, student one, student two, student three". The observer listened to the recording using one headphone to continue to attend to teacher praise/reprimands as well as when points are being given to the students. When another observer was conducting the observation alongside the primary researcher to obtain reliability the second observer wore the other headphone which allowed consistency of when observations were occurring during each interval. Observations occurred three days per week. Two of these observations were conducted during the intervention (i.e., math) and one was conducted at a time the intervention was not implemented to examine generalization (i.e., reading). Observers completed the observations

independently and did not have visual access to each other's observation forms during the observation.

Praise and reprimands. Teacher praise and reprimands were also collected in vivo during each observation period by the primary researcher. Praise was defined as verbal statements to individual students, groups, or whole class indicating approval of behavior. The definition did not include the acknowledgement of a correct response to a question. An example of praise is "Great job for sitting quietly and listening, John!" or "I like that group number three are working on their assignments!" Reprimands were defined as verbal statements to individuals or groups used to comment negatively about student behavior. An example of a reprimand is "Quit standing out of your seat", or "I already told you to sit back down!" Rate of praise or reprimand statements per minute was calculated for each observation period. Rate of praise and reprimands were measured at the classroom level, not just those provided to the targeted students.

Archival data. Attendance records and office discipline referrals were collected as part of the study.

Attendance. Student attendance records were obtained. For the purposes of this study, attendance data was collected for students who were absent the whole day as well as for only part of the day. To assess the impact of the intervention on student attendance a test of proportion was used. In other words, the proportion of days that the student missed prior to the start of the intervention was compared to the proportion of days missed after the intervention began.

Office discipline referrals. Office discipline referrals were defined as infractions that resulted in the student being sent to the office. Office discipline referrals were tracked using the

forms that teachers and the administrator complete after an infraction was committed. Office referrals were calculated using rate per day.

Questionnaire data. Student engagement was obtained to determine how engaged and satisfied the students are with their classroom.

Student engagement. Student classroom engagement was measured using a modified version of the Engagement vs. Disaffection with Learning Student-Report (EvsD; Skinner et al., 2009). The original measure includes questions about student engagement which is divided into two components; emotional and behavioral engagement. The authors defined engagement as "students' active participation in academic activities in the classroom" (Skinner et al., 2009, p. 766). The behavioral engagement component includes two factors, engaged behavior and disaffected behavior. Questions within these factors examine student effort, attention, and persistence during initiation and execution of learning activities. The emotional engagement component also has two factors engaged emotion and disaffected emotion which focuses on states that indicate emotional involvement such as enthusiasm, interest and involvement. This measure includes 20 items, 10 of which target the emotional engagement and 10 that target behavioral engagement (Skinner et al., 2008). The original measure has been validated on students in third grade to tenth grade. Students responded to items using a Likert-type scale, ranging from 1 (Not at All True) to 4 (Very True). According to Skinner et al. (2009), evidence of construct validity was provided through confirmatory factor analyses which found that a fourfactor model had the best fit. Internal consistency reliabilities were found to be between 0.61 and 0.85, which are low to moderate. Test-retest reliability correlations were reported to be between 0.53 and 0.68 from the fall to spring semesters, which suggests a moderate level of stability.

The modified version of the EvsD scale was used for the current study. Only questions that were relevant to the classrooms in the current study were used. The modified questionnaire contained seven questions that focused primarily on engagement, with three addressing engaged behavior (e.g., "When I'm in class, I participate in class discussions") and four addressing engaged emotion (e.g., "When we work on something in class, I feel encouraged"). Items were rated based on a four-point Likert scale (1 = Not true at all; 2 = Somewhat True; 3 = True; 4 = Very True). Smiley faces were used to help students determine their response.

The questions were provided to the target students individually by the researcher one time during the baseline phase and three times throughout the intervention phase. Information obtained from the survey was used qualitatively, where the percentage of students responding in each category were examined each time the scale was administered. Specifically, the percentage of at-risk students who responded with a 3 or 4 on each item was calculated, suggesting general agreement with the item, and compared across time. The modified version of the EvsD scale that was used in the current study can be found in Appendix B.

Data Analysis

Visual analysis was used to interpret graphed baseline and intervention data. This method helped determine if meaningful change occurred in behavior across conditions due to the application or removal of the CW-FIT intervention. Visual analysis was used to examine the variability, level, and trend of on-task behavior and teacher praise/reprimand data (Cooper et al., 2007; Olive & Smith, 2005). Baseline and intervention data both within and between classrooms were compared to determine if behavior changed when and only when the intervention was in place and if the behavior changed immediately after the intervention was introduced.

Reliability

Interobserver agreement. To measure interobserver agreement (IOA), an additional observer observed alongside the primary investigator for 20 minutes. IOA was obtained for 33% of sessions, which was within Cooper at al.'s (2007) recommendation of collecting IOA for 25-33% of sessions per student. The observer recorded the number of teacher praise and reprimand statements as well as student on-task behaviors. IOA for the time on-task observation was calculated using point by point agreement, which is recommended for time sampling. IOA levels of 80% or higher are considered acceptable. With this method, the number of intervals in which the two observers agreed on is divided by the total number of observation intervals and multiplied by 100 to obtain the percentage of agreement (Cooper et al., 2007). IOA for the teacher reprimands/praise was calculated using total count IOA which is calculated by dividing the smaller of the counts by the larger count and multiplying by 100 (Cooper et al., 2007).

Cohen's kappa. To supplement point-by-point agreement, Cohen's kappa was also calculated in order to obtain a more conservative estimate of reliability. Kappa is a reliability measure of agreement between two observers that indicates the proportion of agreement beyond that expected by chance alone. Kappa is calculated by subtracting chance agreement from the observed agreement and dividing the difference by one minus the chance agreement (Sim & Wright, 2005). According to McHugh (2012) values of kappa are interpreted as follows: 0.00-0.20 represents a level of agreement of "*None*"; 0.21-0.39 represents a level of agreement of "*Minimal*"; 0.40-0.59 represents a level of agreement of "*Weak*"; 0.60-0.79 represents a level of agreement of "*Moderate*"; 0.80-0.90 represents a level of agreement of "*Strong*"; and a value of kappa above .90 represents a level of agreement of "*Almost Perfect*". The author noted that 0.80 of higher of agreement is deemed as acceptable.

Procedural Fidelity

A 16-item checklist was used to measure CW-FIT procedural fidelity in order to examine the extent to which the intervention components were implemented as designed. The items on the checklist were modified from the *Procedural Fidelity Checklist* in the CW-FIT Teacher Implementation Manual. The checklist was completed immediately after each observation of a CW-FIT session.

The items in the checklist were directly related to the procedures and components of the CW-FIT intervention (e.g., skills are displayed on posters, timer used during entire session, teacher provides specific praise for use of skills). The items were scored based on the presence (yes) or absence (no) of the component. Questions 9 through 12 were monitored on the observation sheet. Specifically, the use of the timer during the entire session was monitored by the observer marking a + or - if the timer was used throughout the entire session and if it was set at appropriate intervals. Points awarded to teams for use of skills was monitored using a + or - as well. After the timer went off, the observer marked if the teacher provided points to the teams that were following the rules. Additionally, 4:1 praise to reprimand ratio was monitored by tallying praise and reprimands. If a 4:1 ratio or higher was used, the teacher received a score of 1. However, if the ratio was lower than 4:1, the teacher received a 0 on the fidelity checklist. To obtain a fidelity score, the sum of teacher ratings was divided by the sum of possible points. The fidelity checklist is included in Appendix C.

Social Validity

On-task behavior data for peer comparison students in each classroom were collected and used to determine if the on-task behavior of at-risk peers reached a similar level as peers who were identified as not being at-risk by the teacher. This was considered a direct measure of social validity for the effects of the intervention, where a goal of the intervention was not only to improve on-task behavior, but also determine if the level to which on-task behavior improved was to a similar level for an average student in the class. Additionally, at the end of the study a consumer satisfaction survey was distributed to all teachers to assess social validity, which is the extent to which the intervention produced a positive, meaningful change and if the procedures were acceptable (Cooper et al., 2007). Teachers received a 24-item rating scale, which was a modified version of the *Behavior Intervention Rating Profile-15* (BIRP-15; Elliot & Von Brock Treuting, 1991). This scale assesses the intervention acceptability and effectiveness using a five-point Likert scale, where 1 = Strongly Disagree, 2 = Disagree, 3 = Slightly Disagree, 4 = Slightly Agree, and 5 = Agree (Elliot & Von Brock Treuting, 1991). For the current study, BIRP-15 questions were modified specifically for off-task behaviors targeted in the CW-FIT intervention (see Appendix D).

Additionally, at the conclusion of the study the primary investigator asked all students in each classroom six questions to obtain general student input on the intervention acceptability and effectiveness. The questions were created by the primary investigator and addressed student satisfaction with the intervention, possible improvements that could be made, and their perception of the intervention effectiveness (see Appendix E).
Chapter IV: Results

Procedural Fidelity

A checklist for intervention fidelity was used during each observation, which examined the extent to which the components of the intervention were implemented as designed. Overall, the average fidelity rating of 98.3% (SD = 6.3%) and ranged from 64.3% to 100% across all sessions. Thus, 98.3% of the intervention components were implemented across the duration of the intervention suggesting strong implementation of the intervention across classrooms. Due to fidelity data being collected during every observation session of CW-FIT, the observer was able to provide teachers direct feedback about improvements that they could make to increase fidelity. Therefore, overall fidelity was high, with some lower scores at the beginning which were discussed and remediated.

Table 2.

	Teacher 1	Teacher 2	Teacher 3	Overall
	%	%	%	%
	Implemented	Implemented	Implemented	Implemented
1. Getting attention skill poster.	100%	93.3%	93.3%	95.5%
2. Following directions skill poster.	100%	93.3%	100%	97.8%
3. Ignoring behaviors skill poster.	93.3%	100%	100%	97.8%
4. Team point chart displayed.	100%	100%	100%	100%
5. Daily goal posted.	100%	100%	93.3%	97.8%
6. Pre-corrects on getting attention.	100%	100%	93.3%	97.8%
7. Pre-corrects on following directions.	100%	100%	100%	100%
8. Pre-corrects on ignoring behavior.	100%	100%	100%	100%
9. Timer used.	100%	100%	100%	100%
10. Timer set appropriately.	100%	100%	100%	100%
11. Points awarded to teams.	100%	100%	100%	100%
12. 4:1 praise/reprimand ratio.	100%	100%	100%	100%
13. Specific praise.	100%	100%	100%	100%
14. Points tallied.	100%	93.3%	87.5%	93.6%
15. Winners announced.	100%	100%	87.5%	95.8%
16. Rewards delivered.	100%	100%	100%	100%

Intervention Fidelity Components.

If fidelity was not rated at 100%, the primary researcher discussed the components missed with the teacher and provided suggestions on how to improve intervention implementation. Out of 47 observations that were conducted, 4 did not produce 100% fidelity ratings. After each of those four observations the primary researcher met with the teacher and provided feedback.

Interobserver Agreement

Interobserver agreement was collected for 33% of the observations. Mean point by point agreement between the primary researcher and reliability partner for on-task behavior was 97.39% (SD = 0.02%) and ranged from 91.0% - 100%. To obtain a more conservative estimate of IOA, Kappa was also calculated to determine agreement between the two observers while accounting for chance agreement for on-task behavior. Mean Kappa for on-task behavior was .73 (SD = 0.17) and ranged from 0.45 - 1.00, which is considered a moderate level of agreement. Overall, observations could be considered appropriately reliable based on these two analyses.

Mean total count IOA for praise was 90.82% (SD = 6.32 %) and ranged from 80.0% - 100%. Mean total count IOA for reprimands was 100%. These values were considered acceptable (Cooper at al., 2007).

Research Question 1: Replicating Effects of CW-FIT

The first research question addresses the replication of the effects of CW-FIT. Specifically, can the effects of CW-FIT that have been found in previous research be replicated in new classrooms? It was hypothesized that the implementation of CW-FIT will: a) increase student time on-task, b) increase teacher praise statements during the implementation of CW-FIT, and c) decrease teacher reprimand statements during the implementation of CW-FIT? **On-task behavior.** On-task behavior for each student in the three classrooms during math class when CW-FIT was implemented is presented in Figure 1.

Follow up observations were conducted for all teachers 5 and 10 weeks post the end date of the intervention for Teacher 3. These observations occurred 8 and 13 weeks after the intervention ended for Teacher 1, and 7 and 12 weeks after the intervention ended for Teacher 2.

First, a description of each student's data is provided. Next, a visual analysis of each student's data as well as a comparison within and between classrooms is described.

Teacher 1. In Teacher 1's classroom, 3 baseline data points and 15 intervention data points were collected during math class when CW-FIT was implemented.

Student 1 (at-risk). Average baseline on-task behavior for Student 1 was 15.0% (SD = 1.7%) and ranged from 13.3% to 16.7%. Average intervention on-task behavior was 79.6% (SD = 8.1%) and ranged from 68.3% to 88.3%. Student 1's on-task behavior was higher during the intervention phase compared to the baseline phase. During the baseline phase, 0% of the data points were at or above the 80% criterion of on-task behavior, whereas 64.3% of the intervention phase data points were at or above the 80% criterion of on-task behavior. These data suggest an increase in on-task behavior during the intervention phase compared to the baseline phase.



Figure 1. On-task behaviors for each student across teachers and phases during math class.

Student 2 (at-risk). Average baseline on-task behavior for Student 2 was 66.7% (SD = 2.9%) and ranged from 63.3% to 68.3%. Average intervention on-task behavior was 89.3% (SD = 5.2%) and ranged from 78.3% to 95.0%. Similar to Student 1, Student 2's on-task behavior was higher during the intervention phase compared to the baseline phase. During the baseline phase, 0% of the data points were at or above 80% criterion of on-task behavior whereas 86.7% of the intervention phase data points were at or above the 80% criterion of on-task behavior. Similar to Student 1 from this classroom, this suggests an increase in on-task behavior during the intervention phase.

Teacher 1 visual analysis. During the baseline phase, on-task behavior for the at-risk students, Students 1 and 2, were overall low and below 80%. Nevertheless, the baseline behavior for all the at-risk students was stable during baseline data collection. When the intervention was introduced, Student 1 and 2's on-task behavior increased immediately. On-task behavior was stable across all at-risk students towards the end of the intervention (i.e., last five data points) and was at or above the 80% criterion of on-task behavior. While both at-risk students demonstrated an increase in on-task behavior when the intervention was introduced, Student 1's on-task behavior had a larger increase during the first intervention session. His ontask behavior continued to gradually increase for the duration of the intervention observations. Student 1 had some variability in on-task behavior, although his on-task behavior continued to increase gradually and became more consistent towards the end of the intervention phase. Student 2's on-task behavior also had some variability during the first two weeks of the intervention but continued to increase gradually for the remainder of the intervention. Overall, the on-task behavior for the two at-risk students was consistently above 80% by the end of the intervention phase.

Teacher 2. In Teacher 2's classroom, 5 baseline data points and 16 intervention data points were collected during math class when CW-FIT was implemented.

Student 1 (at-risk). Average baseline on-task behavior for Student 1 was 84.7% (SD = 6.1%) and ranged from 78.3% to 93.3%. Average intervention on-task behavior was 91.1% (SD = 5.4%) and ranged from 81.7% to 98.3%. While there was somewhat of an increase in on-task behavior during the intervention phase compared to the baseline phase, the change was not large. During the baseline phase, 80% of the data points were at or above the 80.0% criterion of on-task behavior. While there was an increase in how many of the data points were at or above the 80.0% criterion of on-task behavior. While there was an increase in how many of the data points were at or above the 80.0% criterion of on-task behavior, the majority of the baseline points already reached the criterion.

Student 2 (at-risk). Average baseline on-task behavior for Student 2 was 76.7% (SD = 3.7%) and ranged from 71.7% to 81.7%. Average intervention on-task behavior was 95.8% (SD = 2.6%) and ranged from 90% to 100%. Unlike Student 1, there was an increase in on-task behavior during the intervention phase compared to the baseline phase. During the baseline phase, 20% of the data points were at or above the 80% criterion of on-task behavior whereas 100% of the intervention phase data points were at or above the 80% criterion of on-task behavior. This suggests that on-task behavior increased during the intervention phase compared to the baseline phase.

Teacher 2 visual analysis. Student 1's on-task behavior remained stable across baseline and intervention phases. When the intervention was introduced his on-task behavior did not change in a meaningful way. No trend was observed in Student 1 data across baseline and intervention phases. Conversely, Student's 2's on-task behavior was much lower during the

baseline phase. While there appears to be small upward trend during the baseline phase for Student 2, her data was relatively stable. Her on-task behavior remained high for duration of the intervention with little to no variability. Although only one student in Teacher 2's classroom demonstrated an increase in on-task behavior. When the intervention was introduced in Teacher 1's classroom no change was observed in Teacher 2's classroom during the baseline phase. Thus, any changes in Teacher 2's classroom occurred when and only when the intervention was introduced in Teacher 2's classroom and there was no change in behavior until the intervention was introduced in the classroom.

Teacher 3. In Teacher 3's classroom, 7 baseline data points and 16 intervention data points were collected during math class when CW-FIT was implemented.

Student 1 (at-risk). Average baseline on-task behavior was 65.0% (SD = 2.7%) and ranged from 61.7% to 70.0%. Average intervention on-task behavior was 85.0% (SD = 6.5%) and ranged from 70.0% to 98.3%. Student 1's on-task behavior was higher during the intervention phase compared to the baseline phase. During the baseline phase, 0% of the data points were at or above the 80% criterion of on-task behavior, whereas 86.7% of the intervention phase data points were at or above the 80% criterion of on-task behavior. This data suggests that on-task behavior increased during the intervention phase compared to the baseline phase.

Student 2 (at-risk). Average baseline on-task behavior was 70.8% (SD = 6.2%) and ranged from 63.3% to 78.3%. Average intervention on-task behavior was 95.5% (SD = 4.4%) and ranged from 85.0% to 100%. Similar to Student 1, Student 2's on-task behavior was higher during the intervention phase compared to the baseline phase. During the baseline phase, 0% of the data points were at or above the 80% criterion of on-task behavior whereas 100% of the intervention data points were at or above the 80% criterion of on-task behavior. This further

suggests that on-task behavior increased during the intervention phase compared to the baseline phase.

Teacher 3 visual analysis. While there was a slight upward trend in Student 2's baseline data, the data were relatively stable. Student 1'sbaseline data were stable. No change in Teacher 3's baseline data was seen when the intervention was introduced in Teacher 1 or Teacher 2's classroom. Thus, only when the intervention was introduced in Teacher 3's classroom a change in on-task behavior was observed, specifically where Student 1 and 2's on-task intervention data began to consistently increase. Student 2's behaviors began to steadily increase following the introduction of the intervention and his on-task behavior remained high and continued to increase for the duration of the intervention.

Summary. Taken together, five out of six of the at-risk students appeared to demonstrate a meaningful increase in on-task behavior when CW-FIT was introduced. The sixth student already had high on-task behavior before the intervention, so there was less room for improvement. All of the at-risk students met or exceeded the 80% on-task behavior criterion set at the beginning of the study. Additionally, due to the concurrent multiple baseline design, it is possible to speculate that the behavior changes that occurred in the classrooms were likely only due to the introduction of CW-FIT as their behavior was stable prior to the introduction of the intervention. Thus, for five out of the six students, there appeared to be a functional relation between the intervention and the increase in on-task behavior.

Tier 2. Tier 2 strategies were not utilized because all students reached the 80% criterion of on-task behavior. Although plans were made to use Tier 2 strategies if students were not meeting the criterion, these were not necessary for any of the students in the study.

On-task behavior maintenance. All students were above the 80% criterion during the maintenance observations, and the at-risk students had an average of 90.6% at the first maintenance observation (between 5 and 7 weeks after the intervention ended) and 94.1% at the second maintenance observation (between 10 and 13 weeks after the intervention ended). Importantly, Teacher 2 did not continue the intervention after data collection was completed, but the effects of the intervention appeared remain in her classroom.

Praise and reprimands. Teacher praise and reprimand statements during each observation were calculated using rate (i.e., number of praise statement divided by the number of minutes). Teacher praise and reprimand data is presented in Figure 2.

Teacher 1. During the baseline phase, the rate of praise was much lower compared to when the intervention was introduced. As the intervention phase began, there was a large increase in praise rate for Teacher 1. However, as the intervention continued, there appears to be a downward trend in the praise rate. Nevertheless, while a downward trend is present, the rate of praise continued to be above baseline levels, although there is some variability is present in the teacher's praise. Teacher reprimands during baseline were low, nevertheless, as the intervention was introduced, the teacher only used one reprimand during one of the observations when she was observed throughout the entire intervention phase.



Figure 2. Whole class praise and reprimand rates across teachers and phases during math class when CW-FIT was implemented. Follow up observations were conducted for all teachers 5 and 10 weeks post the end date of the intervention for Teacher 3 (8 and 13 weeks for Teacher 1 and 7 and 12 weeks for Teacher 2).

Teacher 2. During the baseline phase, Teacher 2 had a higher rate of praise compared to Teacher 1, although the rates of praise were inconsistent. When the intervention was introduced, the teacher had an increase in the rate of praise for the first three observations. However, following those observations there was a slight downward trend which continued for the duration of the intervention. Overall, it is difficult to determine if implementation of the intervention resulted in increased praise as the praise rate during baseline was high and consistent with the intervention praise rate. Nevertheless, the teacher had a higher rate of praise as the intervention was introduced, despite a decline as the intervention continued. Therefore, it is possible that the introduction of CW-FIT initially impacted how much the teacher praised the students despite a slight decrease over time. Teacher 2 did not use any reprimands during the baseline or intervention phases.

Teacher 3. Teacher 3's praise rates were much lower during baseline compared to when the intervention was introduced. As the intervention was introduced there was a large increase in praise rate. Following that initial increase, there was a slight decrease, but unlike Teacher 1 and 2, Teacher 3's praise rate increased slightly throughout the intervention phase. While there were 3 data points that were lower than baseline, there appears to be an overall increase in praise rate as the intervention was introduced. These data suggest that the introduction of CW-FIT likely contributed to the increase of teacher praise. There were some reprimands present during the baseline phase, however after the intervention was introduced the reprimands were no longer used by Teacher 3.

Taken together, Teachers 1 and 2 demonstrated a slight increase in praise as the intervention was introduced, but this gradually declined over time. Teacher 1 had a larger increase in the rate of praise compared to Teacher 2 as the intervention was introduced. Unlike Teachers 1 and 2, Teacher 3 demonstrated an increase in praise rate as the intervention was introduced and this continued to increase slightly for the duration of the intervention. There were some reprimands present while observing Teacher 1 and 3; however, as the intervention was introduced few to no reprimands were observed. As there were very few reprimands present across phases for all the teachers it is difficult to determine if the intervention had an impact on reducing the rate of the reprimands. However, it is promising that after the intervention was introduced, almost no reprimands were used across teachers. Overall, it appears the intervention impacted how frequently the teachers praised the students for Teacher 1 and 3.

Praise and reprimand maintenance. Whole class raise and reprimand rates were maintained at 5 weeks and at 10 weeks. Importantly, Teacher 2 did not continue the intervention after data collection was completed, but the effects of the intervention appeared to remain in her classroom.

Research Question 2: Generalization of CW-FIT

The second research question addresses the generalization of CW-FIT. Specifically, do the effects of CW-FIT found in previous research generalize to the portions of the day the intervention is not used? It was hypothesized that CW-FIT will: a) increase on-task behavior when CW-FIT is not being implemented, b) increase teacher praise statements when CW-FIT is not being implemented, and c) decrease teacher reprimand statements when CW-FIT is not being implemented.

On-task behavior generalization. Generalization on-task behavior data for each student in the three classrooms is presented in Figure 3. CW-FIT was not implemented during this time.

Teacher 1. Only one data point was collected for all three students during the baseline phase, therefore it is difficult to determine if any of the changes in behavior are due to the introduction of the intervention or some other factor. Eight intervention data points were collected during reading class when CW-FIT was not implemented.

Student 1 (at-risk). Student 1's on-task behavior during the baseline generalization observation was 91.7%. Average intervention on-task behavior during generalization was 73.5% (SD = 10.9%) and ranged from 55.0% to 93.3%. Overall, the student's on-task behavior shows a downward trend and decreased from the baseline data point. Student 1's baseline on-task behavior exceeded the 80% criterion and during the intervention phase and only 12.5% of the intervention data points were at or above the 80% criterion of on-task behavior. This demonstrates that there was a decline in on-task behavior following the baseline phase. Nevertheless, it is difficult to determine why there may have been a decrease in on-task behavior, as the first intervention generalization data point was consistent with the baseline data point.



Figure 3. On-task behaviors for each student across teachers and phases during generalization. Follow up observations were conducted for all teachers 5 and 10 weeks post the end date of the intervention for Teacher 3 (8 and 13 weeks after the intervention ended for Teacher 1, and 7 and 12 weeks after the intervention ended for Teacher 2). Observations were conducted during reading class when CW-FIT was not implemented.

Student 2 (at-risk). Student 2's on-task behavior during the baseline generalization observation was 73.3%. Average intervention on-task behavior during generalization was 93.6% (SD = 10.5%) and ranged from 71.7% to 100%. Unlike Student 1, there appears to be an upward trend in on-task behavior. Student 2's on-task behavior improved over time, was consistently high, and overlapped with the peer comparison's on-task behavior. Student 2's baseline on-task behavior did not meet the 80% criterion and during the intervention phase 85.7% of the intervention data points were at or above the 80% criterion of on-task behavior. This further demonstrates that there was an increase in on-task behavior following the baseline phase. However, it is difficult to determine if the change in behavior was a result of the intervention as only one baseline data point was collected and the first intervention data point was consistent with the baseline data point.

Teacher 1 generalization visual analysis. Overall, it is difficult to determine if the intervention produced any behavioral change in a class period where the intervention was not being used due to the lack of baseline data. However, Student 2 appears to have an upward trend as the intervention continued. Nevertheless, the increase in on-task behavior did not occur directly after the intervention as introduced. Conversely, Student 1's on-task behavior seemed to have declined; however, similar to Student 2, change did not occur directly after the intervention was initiated it is difficult to determine why his on-task behavior decreased.

Teacher 2. Only two data points were collected during the baseline phase, thus it is difficult to determine if any change in on-task data was due to the introduction of the intervention due to insufficient amount of baseline generalization data points. Seven intervention data points were collected during reading class when CW-FIT was not implemented.

Student 1 (at-risk). Student 1's on-task behavior during the baseline phase was 81.7% and 78.3%. Average intervention on-task behavior during generalization was 88.1% (SD = 4.0%) and ranged from 83.3% to 93.3%. Overall, there appears to be an upward trend in on-task behavior, nevertheless it is difficult to determine if the change in behavior was due to the introduction of intervention as there were only two baseline data points. Additionally, one of the baseline points met the 80% criterion and during the intervention phase 100% of the data points were at or above the 80% criterion of on-task behavior. This demonstrates that while there was somewhat of an increase in on-task behavior, it is difficult to determine whether the baseline data points were good estimates of the student's behavior during the baseline phase, as there were only two data points.

Student 2 (at-risk). Student 2's on-task behavior during the baseline phase was 73.3% and 71.7%. Average intervention on-task behavior during generalization was 93.3% (SD = 6.0%) and ranged from 81.7% to 98.3%. Similar to Student 1, there is an upward trend during the intervention phase in on-task behavior. During the baseline phase, the data points did not meet the 80% criterion and during the intervention phase 100% of the data points were at or above the 80% criterion of on-task behavior. This demonstrates that there was an increase in on-task behavior change in behavior following the baseline phase. However, it is difficult to determine if behavior change in behavior occurred due to the introduction of intervention as there were not enough baseline data

points and it is difficult to know if the two baseline data points were a good estimate of the student's behavior.

Teacher 2 generalization visual analysis. It is difficult to determine if generalization occurred due to the insufficient number of baseline points. Nevertheless, Students 1 and 2 had some change in behavior after the baseline data points. Specifically, their on-task behavior increased after the intervention was introduced. Although there was insufficient number of data points during baseline, there is some evidence to suggest that some behavior change occurred only when the intervention was introduced in Teacher 2's classroom.

Teacher 3. Three data points were collected during the baseline phase. Data were collected during reading class when CW-FIT was not implemented. However, there was a large amount of time that elapsed between the last baseline data point to the time when the intervention was implemented, and no baseline data was able to be collected after the intervention was introduced in Teacher 2's classroom. Therefore, it is difficult to determine if student on-task behavior remained on baseline levels during the gap in the data. Seven intervention data points were collected.

Student 1 (at-risk). Student 1's average on-task behavior was 62.2% (SD = 1.0%) and ranged from 61.7% to 63.3%. Average intervention on-task behavior during generalization was 70.8% (SD = 3.2%) and ranged from 65.0% to 75.0%. Student 1's data was relatively consistent across phases, with a slight upward trend after the intervention was introduced, but his behavior did not appear to demonstrate a meaningful change as a result of introducing the intervention. The student's on-task behavior continued to be below 80% across both phases.

Student 2 (at-risk). Student 2's average on-task behavior was 62.8% (SD = 2.0%) and ranged from 61.7% to 65.0%. Average intervention on-task behavior during generalization was

79

90.4% (SD = 2.8%) and ranged from 86.7% to 93.3%. Overall, Student 2's baseline was stable with no apparent trend. However, unlike Student 1, Student 2's on-task behavior increased when the intervention was introduced. Therefore, it is possible that the intervention generalized and impacted his on-task behavior. None of the baseline data points met the 80% criterion compared to 100% of the intervention data points were at or above the 80% criterion of on-task behavior. This demonstrates that there was an increase in on-task behavior during the intervention phase. Nevertheless, due to the gap between baseline and intervention data points, it is possible that an increase in on-task behavior occurred prior to the introduction of the intervention.

Teacher 3 generalization visual analysis. Student 1 on-task behavior was fairly consistent across the baseline and intervention phases and did not seem to be impacted by the introduction of the intervention. Student 1's on-task behavior had a slight upward trend when the intervention was introduced. However, his on-task behavior remained below 80% without any meaningful change across phases. Student 2's on-task behavior was low during the baseline phase, however as the intervention was introduced, his on-task behavior increased and it appears that the intervention possibly had an impact on his behavior. Additionally, change in behavior only occurred when the intervention was introduced in Teacher 3's classroom. Furthermore, as the intervention was introduced in Teacher 3's classroom, behavior did not change in the other classrooms. Nevertheless, due to the gap between the last baseline data point to the first intervention point it is difficult to determine if the baseline data points were good estimates of the student's behavior during the baseline phase. Therefore, any change in behavior is difficult to interpret.

Summary. Overall, it is difficult to determine whether the effects of CW-FIT generalized to a part of the day when the intervention was not implemented as the results are inconsistent.

Additionally, it is difficult to determine if generalization occurred due to the insufficient number of baseline data points for Teachers 1 and 2 and the large gap between the last baseline data point and the first intervention data point for Teacher 3. Nevertheless, at least one student from each teacher showed an upward trend during the intervention phase. However, Student 1 in Teacher 2's classroom demonstrated a downward trend indicating that his on-task behavior decreased as the intervention continued. Taken together, it is possible that some generalization occurred, however this question is difficult to answer using the data that was collected in the study.

On-task behavior generalization maintenance. Two of the at-risk students were above the 80% criterion during the maintenance observations, and the at-risk students had an average of 76.1% at 5 weeks and 86.5% at 10 weeks. On-task behavior maintained for three of the at-risk students (Student 2 from Teacher 2's classroom and Student 1 and 2 from Teacher 3's classroom). Importantly, Teacher 2 did not continue the intervention after data collection was completed, but Teachers 1 and 3 did continue the intervention.

Generalization praise and reprimand. Generalization class-wide teacher praise and reprimand data is presented in Figure 4. These data were collected during reading class. CW-FIT was not implemented during this time.

Teacher 1. As only one baseline data point was collected, it is difficult to determine the differences in praise and reprimands before the intervention was introduced compared to after it was introduced. However, there does appear to be an initial increase in praise during the first part of the intervention, then a general decrease over time. There is one data point that is lower than the rest, however following that data point rates of praise continued to rise. Throughout the baseline and intervention phases only one reprimand was observed, therefore the absence of reprimands was consistent across phases.

Teacher 2. Teacher 2's praise rate was inconsistent over time. There appears to be a large decline in praise following the baseline data points, which remains consistent for three data points. The decrease in praise seems to correspond with an increase in reprimands. Rate of praise then increase to the levels of which they were at during baseline. Overall, reprimands were not present during the baseline observations, however they occurred more frequently as the intervention was occurring. After reprimands increased they remained fairly consistent with a decrease during one data day and then an increase, which remained consistent for the duration of the intervention.

Teacher 3. There was very little change in the rate of praise across phases for Teacher 3. Regarding reprimands, Teacher 3 was only observed reprimanding the class on one of the observations during the baseline phase. No other reprimands were observed following that time.



Figure 4. Generalization class-wide praise and reprimand rates across teachers and phases. Follow up observations were conducted for all teachers 5 and 10 weeks post the end date of the

intervention for Teacher 3 (8 and 13 weeks for Teacher 1 and 7 and 12 weeks for Teacher 2). Observations were conducted during reading class when CW-FIT was not implemented.

Summary. Overall, class-wide praise and reprimands during the generalization observations were inconsistent across teachers. Only one of the teachers, Teacher 1, demonstrated an increase in the rate of praise in the generalization class period after the intervention was introduced in another class period. There was a gradual decline in praise as the intervention continued. Teachers 2 and 3, however did not demonstrate much of a change in praise. Reprimands were consistently absent or low for Teacher 1 and 3. Teacher 2, however had an increase in reprimands when her rate of praise decreased. Due to these inconsistent results as well as the insufficient number of baseline points for Teachers 1 and 2 as well as the gap in data points between the last baseline data point and the first intervention data point for Teacher 3, it is difficult to determine if any change in the rates or praise was due to the introduction of the intervention. Reprimand rates were low across phases and teachers; thus, it was not possible to determine if CW-FIT had an influence on reprimands because they were already occurring at a very low rate.

Generalization praise and reprimand maintenance. Class-wise praise and reprimand rates were maintained at 5 weeks and at 10 weeks. Importantly, Teacher 2 did not continue the intervention after data collection was completed, but the effects of the intervention appeared to remain in her classroom.

Research Question 3: CW-FIT and Other Positive Outcomes

The third research question addresses the impact of CW-FIT on other outcomes. Specially, does the implementation of CW-FIT produce other positive outcomes for students not investigated in previous research? It was hypothesized that there will: a) be a reduction in office referrals, b) be an increase in attendance, and c) be an improvement in classroom satisfaction.

Office discipline referrals. No office discipline referrals were reported across phases or across classrooms. Therefore, the impact of CW-FIT on office discipline referrals cannot be determined using the data from the current study.

Attendance. Student attendance was measured using the percentage of days that the students were absent before the intervention was introduced compared to the percentage of the days that the student was absent after the intervention was implemented. Results were inconsistent, and the percentage of days absent was actually higher during intervention compared to baseline for eight out of nine students. A test of proportions was conducted to determine if the percentage of absences was statistically significantly different between baseline and intervention phases for each student. An online calculator was used to conduct the calculations (www.socscistatistics.com/tests/ztest/default2.aspx). Due to the large number of comparisons, Bonferroni correction was used to control for Type I error (α = .05 / 9 comparisons = 0.006). None of the comparisons were statistically significant at the *p* < .006 level, indicating no difference in student attendance across phases. Therefore, based on the available data CW-FIT did not appear to have any impact on student attendance and the data were difficult to evaluate because some of the students began the intervention with zero absences.

Table 3.

		Baselir	ie		Intervention			
Participant	Absences	School Days	%	Absences	School Days	%	Ζ	р
T1 S1	4	28	14.29%	1	35	2.86%	1.67	0.95
T1 S2	2	28	7.14%	3	35	8.57%	-0.21	0.84
T1 S3	0	28	0.00%	3	35	8.57%	-1.59	0.11
T2 S1	2	32	6.25%	5	35	14.29%	-1.07	0.28
T2 S2	1	32	3.13%	2	35	5.71%	-0.51	0.61
T2 S3	0	32	0.00%	3	35	8.57%	-1.69	0.09
T3 S1	0	38	0.00%	2	36	5.56%	-1.47	0.14
T3 S2	1	38	2.63%	3	36	8.33%	-1.08	0.28
T3 S3	1	38	2.63%	1	36	2.78%	-0.04	0.97

Proportion of student absences before and after the intervention was implemented

Note. T = teacher; S = student; Students 1 and 2 are at-risk students and Student 3 is the comparison peer

Student satisfaction in the classroom. Another source of data was collected to

determine how engaged, satisfied, and happy the students were in the classroom. Students were

asked questions and their responses were monitored across phases.

Table 4.

Percentage of at-risk students that rated each item as True (3) or Very True (4)

Student Engagement Questions	Baseline	Time 1	Time 2	Time 3
In class, I work as hard as I can.	66.7%	83.3%	100.0%	100.0%
When I'm in class, I participate in class discussions.	100.0%	100.0%	100.0%	100.0%
When I'm in class, I listen very carefully.	83.3%	100.0%	100.0%	100.0%
When I'm in class, I feel good.	50.0%	50.0%	83.3%	100.0%
I enjoy learning new things in class.	83.3%	66.7%	66.7%	100.0%
When we work on something in class, I feel encouraged.	83.3%	83.3%	100.0%	100.0%
Class is fun for me.	50.0%	100.0%	100.0%	100.0%

Note. Time 1, 2, 3 = three different times during the intervention phase during which students were given the rating scale; Time 1 = 9/19/2019, Time 2 = 11/1/2019, Time 3 = 11/18/2019 (Teacher 1), 11/22/2019 (Teacher 2), 12/4/2019 (Teacher 3).

All six at-risk students were provided a questionnaire to assess their feelings about the

classroom. The questions were read aloud by the researcher. Table 4 demonstrates the percent of

at-risk students that responded to the questions with a rating of either a 3 (i.e., *True*) and 4 (i.e., *Very True*). The results show that all of the questions but one, "*I enjoy learning new things in class*", demonstrated progress across time. Specifically, on the third time point all of the at-risk students responded with a rating of a 3 or a 4 to every question, compared to the baseline time point during which only one question was answered with all students using those response options. Two of the items, "*When I am in class, I feel good*" and "*Class is fun for me*" showed the most improvement, from 50% during the baseline phase to 100% on time three of the administration of the survey. These results demonstrate evidence that all of the at-risk student's feelings about their classroom improved, as more positive responses were reported at the end of the intervention compared to the baseline phase. This suggests that the inclusion of CW-FIT in their classroom may have also improved their feelings about the class.

Peer comparison students from each classroom were also provided the questionnaire; however, all peer comparison students reported ratings of 3 or 4 across phases. This suggests that the peer comparisons were positively engaged in the classroom prior to the introduction of the intervention and continued to be positively engaged throughout the intervention. Additionally, by the third time point all at-risk students' ratings were similar to the peer-comparison's which indicates similar satisfaction and engagement to peer comparison students.

Social Validity

Peer comparisons. Peer comparison data was collected in each classroom, which is a direct measure of social validity that can be used to determine if the at-risk students' on-task behavior was a similar level to average classroom peers. Even though a criterion of 80% for on-task behavior was set, this provides a normative comparison of on-task behavior for the at-risk

students. Overall, on-task behavior for the peer comparisons was high during the baseline phase and did not change when the intervention was introduced across classrooms.

Intervention. During the intervention class (i.e., math) when CW-FIT was being implemented, in Teacher 2 and 3's classroom the on-task behavior for Student 2 often approached the peer's on-task behavior and at points reached the peer. Although the remaining students did not reach the peer comparison some neared the peer. Specifically, Student 1 from Teacher 2's classroom approached the peer comparison's data towards the end of the intervention phase. Student 1 from Teacher 3's classroom did not near the peer comparison's on-task behavior for most of the intervention data points, however his behavior approached that of the peer comparison two times during the intervention phase. Conversely, Student 1 from Teacher 1's classroom did not reach the peer comparison's on-task behavior.

Generalization. During the generalization class when CW-FIT was not implemented (i.e., reading), Student 2 from Teacher 1 and 2's classroom approached and reached the peer comparison's on-task behavior. The remaining students did not reach the peer's on-task behavior. However, Student 1 from Teacher 2's classroom neared the peer. Conversely, Student 1 from Teacher 1 and 3's classroom did not near the peer's on task behavior.

Summary. Overall, two students from the intervention class (i.e., math) and two students from the generalization class (i.e., reading) reached the peer comparison. However, one student from the intervention class and two students from the generalization class did consistently approach the peer comparison. Importantly, even though CW-FIT improved on-task behavior for all at-risk students to a minimum level (80%), these results indicate that the intervention produced improvements in some of the at-risk student's behavior that similar to students who are not considered at-risk.

Teacher questionnaire. Teacher social validity was evaluated using a modified version of the BIRP-15 (Elliot & Von Brock Treuting, 1991). Items on the scale were modified to reflect CW-FIT. Across all items on the BIRP-15, the average between all three teachers was a rating of 4.1 out of 5 (SD = 0.7). This indicates that, on average, the teachers chose *Slightly Agree* on the items. Teacher 1's mean was 4.1 (SD = 0.6), Teacher 2's mean was 3.7 (SD = 0.5), and Teacher 3's mean was 4.5 (SD = 0.8).

An examination of individual items provides a greater depth of understanding of their perception of the intervention. The following statements demonstrated the highest level of agreement (M = 4.7 across all items): "This was an acceptable intervention for increasing ontask behavior"; "I would suggest the use of this intervention to other teachers"; "I would be willing to use this intervention in the classroom setting"; "I liked the procedures used in this intervention"; "This intervention was a good way to handle off-task behaviors"; "Soon after using the intervention, I noticed a positive change in off-task behaviors". These results suggest that the teachers felt that CW-FIT produced a positive effect on off-task behaviors, was well liked, and the teachers are willing to continue to use it.

The following statement produced the lowest level of agreement (M = 2.7): "*The intervention produced enough improvement in the student's off-task behaviors so that the behavior no long is a problem in the classroom*". These ratings suggest that while teachers believe there was an effect on reducing off-task behavior, the intervention did not completely eliminate off-task behavior as a problem.

While examining the individual teachers, Teacher 3 had the lowest ratings on average. She rated a 2 (i.e., *Disagree*) on "*On-task behaviors will remain at an improved level even after the intervention is discontinued*" and a 3 (i.e., *Slightly Disagree*) on "*This intervention quickly* improved the child's behavior"; "The intervention provided a lasting improvement in student off task behavior"; "Using the intervention not only improved the child's behavior in the classroom, but also in other settings (e.g., other classrooms, home)"; "The intervention produced enough improvement in the student's off-task behaviors so that the behavior no long is a problem in the classroom". These results indicate that she felt that the intervention will not produce lasting effects, the intervention did not result in a fast decrease in on-task behaviors, and the intervention did not eliminate the problem of off-task behaviors.

In addition to the questions, a feedback section was included at the end of the survey and contained some of the following comments: "*Really enjoyed the intervention*!" "*This really helped all of the students in my classroom*", "*I am looking forward to doing this intervention during other parts of my day*!"

Student questions. In addition to the BIRP-15, all students in the three classrooms were asked questions regarding their opinion of CW-FIT. Qualitative data was obtained from their answers. Students in all three classrooms responded positively to the intervention and reported that they enjoyed it. Some shared that they enjoyed the rewards that they received at the end of the game and others reported that they liked getting points. When they were asked what they did not like about the intervention, students in Teacher 1's classroom reported that they wished they were able to play it more. Other students noted that sometimes it did not seem fair that they were working hard and their team did not get a point because of another student on their team. When asked if the game helped them focus better, students in the Kindergarten class (Teacher 2) and students in the second-grade class (Teacher 3) reported that it did not help them focus better, but it helped them focus better. When asked what they wished was different about the game many of

the students in all three classrooms agreed that they wished they could pick their own teams. Lastly, when they were asked if they would like to continue playing the game, students in all three classrooms reported that they would like to continue with the game.

Chapter V: Discussion

The purpose of the current study was to replicate findings from previous studies on the effectiveness of CW-FIT to increase on-task behaviors, increase teacher praise, and decrease teacher reprimands. A new addition to past research on CW-FIT that this study aimed to examine is if the aforementioned effects generalized to another part of the day when the intervention was not actively implemented. Additionally, other factors such as student attendance, enjoyment in the classroom, and office discipline referrals were collected to determine if the implementation of CW-FIT had an effect on those variables.

On-Task Behavior

All but one of the at-risk students showed substantial increases for on-task behavior across the three classrooms. As the intervention was introduced, their on-task behavior increased and either continued to increase or remained high as the intervention continued. All of the at-risk students' on-task behavior reached 80% or above over the duration of the intervention. Although students in Teacher 2's classroom did not demonstrate as much of a change in on-task behavior as the students the other classrooms, this was likely due to their on-task behavior already being high during baseline. On-task behavior for students improved only after the intervention was implemented in each classroom. These findings suggest there was a functional relation between the intervention and the increase in on-task behavior for five of the at-risk students whose on-task behavior was low (below 80%) during baseline. These effects were maintained five and ten weeks following the end of the intervention in all classrooms. Additionally, the effects were maintained in one classroom that did not continue the intervention (Teacher 2). The remainder of the teachers continue the intervention following the end of data collection.

Increase of on-task behaviors is consistent with previous CW-FIT studies where on-task behaviors increased as CW-FIT was introduced (Caldarella et al., 2015; Conklin et al., 2017; Kamps et al., 2011; Weeden et al., 2016; Wills et al., 2018; Wills et al., 2014). Other studies using CW-FIT also found similar results, but in those studies on-task behavior was measured across groups of student in the classroom rather than individual students (Caldarella et al., 2015; Conklin et al., 2017; Hansen et al., 2017; Hirsch et al., 2016; Wills et al., 2014). Overall, the current study replicated previous research suggesting CW-FIT has a positive influence on ontask behavior in the classroom.

Praise and Reprimands

Teacher praise. Overall, teacher praise increased when the intervention was introduced. These results are consistent with previous studies which found that the amount of praise is higher during the CW-FIT phase compared to when the intervention was not occurring (Caldarella et al., 2015; Conklin et al., 2017; Hansen et al., 2017; Kamps et al. 2011; Kamps et al. 2015; Weeden et al., 2016; Wills et al., 2014; Wills et al., 2016). However, as the intervention continued, Teacher 1 and 2's praise rate slowly decreased. Nevertheless, Teacher 2's praise rates continued to be above where they began during the baseline phase whereas Teacher 1's fell below baseline levels towards the end of the intervention. These effects were maintained over time. While past studies have not examined a trend in praise over time, but rather compared means between phases or treatment conditions, two studies found that when the intervention was re-introduced after a reversal phase the rate of praise was lower (Caldarella et al., 2017; Hansen et al., 2017). Conversely, findings from Kamps et al. (2011) and Jolstead et al. (2017) showed that teacher praise continued to increase when CW-FIT was re-introduced. It is possible that due to the student's positive response to the intervention, the teachers faded their praise as the students began to comply and demonstrate appropriate behaviors without additional praise from the teacher. This was also hypothesized by Caldarella et al. (2017) and may explain this trend in the current study. Additionally, teachers may have been using an intermittent reinforcement schedule of reinforcement for praise, during which they were not praising every single appropriate behavior. Previous literature has shown intermittent reinforcement to be a strong method that results in more consistent desired behavior in the future (Baer et al., 1987).

Teacher reprimands. All three teachers had very low rates of reprimands across phases, therefore it is difficult to determine if the introduction of CW-FIT impacted how often teachers used reprimands. The absence of reprimands was maintained over time. These findings are somewhat consistent with past research in which reprimands did not change significantly during the intervention compared to baseline (Caldarella et al., 2015; Conklin et al., 2017). Hansen et al. (2017) found that reprimands increased in some of the classes when CW-FIT was introduced or re-introduced following a reversal phase, but most studies have found reprimands decreased when the intervention was introduced (Caldarella et al., 2017; Kamps et al. 2015; Weeden et al., 2016; Wills et al., 2014). It is important to note that previous studies have not examined the trend of reprimands across phases, instead they examined them by comparing the average amount of reprimands across phases or conditions or comparing praise to reprimand ratios.

Generalization of CW-FIT Effects

On-task behaviors. It is difficult to determine if generalization of on-task behavior to other parts of the day occurred based on the results from the current study, primarily due to an insufficient number of baseline data points that were able to be collected due to personnel shortage and time constraints. Additionally, while there were three data points for Techer 3, due to parent teacher conferences, teacher and student absences, and days off from school for the

students, there was a large gap between the last baseline data point to the first intervention data point. The gap makes it difficult to determine if student on-task behavior remained at baseline levels when the intervention was introduced in Teacher 1 and Teacher 2's classroom, before the intervention was introduced in Teacher 3's classroom. While there was insufficient baseline data to determine if a change in behavior was due to the introduction of the intervention, this amount of baseline data is similar to a multiple probe design. In this type of design minimal data points are collected during the baseline phase and the observations in the classrooms are conducted at different times to determine if the observed behavior has continued to maintain on baseline levels. Therefore, while this was not the intention prior to beginning the study, this method may help in explaining the changes in the on-task behavior that were observed in some of the students and teachers (Horner & Baer, 1978).

Although there appears to be some behavioral change in some of the students in Teacher 1 and 2's classrooms, it is difficult to determine if those changes were due to the intervention because of very little baseline data for those two teachers. Additionally, the changes that occurred were inconsistent. Specifically, Student 1's on-task behavior in Teacher 1's classroom demonstrated a decline while the intervention was occurring. However, this change did not occur directly after the intervention was implemented. Conversely, Student 2 on-task behavior from Teacher 1's classroom demonstrated an increase while the intervention was occurring. Similar to Student 1, this change did not occur directly after the intervention was introduced.

Teacher 3 had a sufficient number of baseline data points, although there was a significant gap between the last baseline data point to the first intervention data point. This makes it difficult to determine if the student on-task behaviors remained at baseline levels when the intervention was implemented with Teachers 1 and 2. Nevertheless, the results showed that

one of the at-risk students demonstrated very little increase in on-task behavior and remained under 80% across both baseline and intervention phases. However, the other at-risk student demonstrated an increase in on-task behavior when the intervention was introduced which neared the peer comparisons. Five and ten weeks following the end of the intervention, results showed that four of the six at-risk students maintained their on-task behaviors similar to previous levels.

Praise and reprimand. Similar to the on-task generalization data, due to the lack of baseline data points it is difficult to determine if any changes in teacher behavior occurred due to the intervention. Nevertheless, Teacher 1 appeared to have an increase in praise over time. Conversely, Teachers 2 and 3 demonstrated very little change in their praise across phases. Furthermore, Teachers 1 and 3 had a minimal number of reprimands. However, Teacher 2 had an increase in reprimands as her praise rate decreased. Overall, it is difficult to determine if generalization of praise occurred as the results are inconsistent and an increase in praise was only observed in one teacher. Rates of praise and reprimand were maintained over time.

In previous studies the generalization of CW-FIT has not been examined. However, when CW-FIT was implemented throughout the day the effects of the intervention (i.e., increase ontask behaviors, increase in praise, and decrease in reprimands) were observed across all intervention times (Wills et al., 2014). Although this study attempted to fill this gap, unforeseen circumstances with data collection make it difficult to interpret the data that was collected and to form any strong conclusions about the generalization of CW-FIT.

Attendance/Student Satisfaction/Office Discipline Referrals

Attendance. Poor student attendance is related to poorer academic outcomes, behavior difficulties, and an increased risk of drop out (Appleton et al., 2008; Archambault et al., 2009). After examining attendance data across all students, no meaningful differences were found in

attendance across phases. After the intervention was introduced some of the students had no absences and some did have absences. However, it is difficult to determine the true reason for student absences. If a student misses a day of school due to illness, an intervention like CW-FIT will not be able to change their absences. Therefore, due to various factors that contribute to student attendance, it is difficult to determine if an intervention like CW-FIT would encourage students to attend school more because there are many reasons beyond simply student motivation that may influence whether they attend school.

Office discipline referrals. Similar to attendance, higher numbers of office discipline referrals negatively impact teacher-student relationships, increase time away from the classroom, and ultimately do not improve problem behaviors (Hawken, Sandra MacLeod & Rawlings, 2007). The rate of office discipline referrals was examined as part of the study, however there were no office discipline referrals that were reported for the targeted students as part of the study across phases. Thus, there was no way to determine whether there was a change in office discipline referral rate when CW-FIT was implemented because no students received any office discipline referrals.

Student classroom satisfaction. Students were provided a questionnaire, a modified version of the EvsD scale (Skinner et al., 2009), and asked about their feelings and effort in the classroom. The percent of at-risk students that responded with 3 (i.e., *True*) and 4 (i.e., *Very True*) increased across most questions during the intervention, including those asking if students were working as hard as they can, participating in class discussions, listening carefully, feeling good in the classroom, and feeling encouraged while working on something in the classroom increased. The only question that decreased from baseline to the intervention phase was students enjoying learning new things in the classroom. The intervention for all teachers was

implemented during math class. A new and more challenging curriculum, per teacher report, was introduced to the students this year, therefore this may explain why the percent of students that chose a rating of 3 and 4 remained low for this item across the intervention phase. Although the same size was small and these data should primarily be interpreted qualitatively, there may have been some influence of CW-FIT on how students felt about their classroom.

Overall, the at-risk students indicated they were feeling more positive in the classroom during the intervention phase. Although previous studies have not investigated how an intervention like CW-FIT may improve student's feelings about their classroom, past research has shown that strategies used in CW-FIT such as clear rules, positive reinforcement, and fair consequences can increase student motivation (Williams & Williams, 2011). Additionally, Johnson (2009) and Rhode et al. (1993) noted that positive student and teacher relationships are related to more positive student behaviors and a sense of belongingness. Therefore, it is possible that an intervention like CW-FIT, which includes all of the components mentioned above, leads to the students feeling more positively in the classroom. Peer comparisons were provided the questionnaire as well and their satisfaction was high throughout. Nevertheless, as all of the atrisk students were at 100% at the third time point, this may indicate that the at-risk students were just as satisfied and engaged in the classroom by the end of the intervention as the comparison peers.

Social Validity

A unique component of this study was measuring on-task behavior of peer comparison students, who could be used as normative comparisons and can be considered direct measures of social validity. This was done by determining if on-task behavior for at-risk students could reach levels similar to the peer comparison. The results showed that two students from the intervention
class when CW-FIT was implemented (i.e., math) and two students from the generalization class when CW-FIT was not implemented (i.e., reading) reached the peer comparison. However, only one student from the intervention class and two students from the generalization class showed on-task behaviors that neared the peer comparison's on-task behavior. These results suggest that on-task behaviors not only increased after the intervention was implemented, but the increases were to levels similar to that of a peer comparison who was not considered at-risk at the beginning of the study. Previous studies have not utilized a peer for comparison purposes when targeting individual students, thus this study provides additional evidence that on-task behavior not only increases, but for some students their on-task behavior approaches levels similar to average students in the class.

Another measure of social validity was the completion of the BIRP-15 by each teacher. Results showed teacher satisfaction scores overall was a mean of 4.1 (SD = 0.7) out of a possible score of 5, which indicates that the teachers, on average, responded "agree" to the statements. This indicates that, overall, they were satisfied with the intervention. Teacher 2 had the lowest ratings among all three teachers. This is possibly due to students in her classroom generally already demonstrating fairly high levels of on-task behavior during baseline, which did not allow for much change to be observed during intervention. Additionally, the teachers rated that the intervention improved student on-task behavior and they were willing to utilize the intervention. However, they felt that intervention that targets on-task behaviors. Although Teacher 2 discontinued the intervention following the end of the data collection period, she indicated that she enjoyed the intervention and it did appear to help some of her students. However, she chose to discontinue the intervention as she wanted to focus on the new math curriculum which was indicated was challenging to her and the students without any additional variables.

Limitations

While this study produced useful results, there are some limitations that must be considered when interpreting the results. First, it is possible that the students who were chosen as at-risk students were not correctly identified as those with the most significant difficulties in class. While the teachers used a validated tool to identify and rank students, the teachers may have ranked the most at-risk students incorrectly. Therefore, it may have been beneficial to include a secondary tool to determine which students to target. Also, as the intervention was introduced fairly close to the beginning of the year, it is possible that the students improved due to other factors such as learning classroom routines and becoming more comfortable with expectations. Furthermore, since the teachers were aware which students the primary investigator was targeting, it is possible that they paid closer attention to the student, provided them extra praise, which then caused them to improve rather than due to the intervention. Additionally, due to choosing the most severe students in the classroom, it is possible that the students showed improved due to regression to the mean, during which the student behavior may have moved closer to the average student regardless of intervention. However, given the design used in the current study, it can be observed that change in behavior only occurred in the classroom when the intervention was introduced. Therefore, it is unlikely that regression to the mean is the primary explanation for the changes in behavior.

A major limitation to the study is the lack of sufficient baseline generalization observations due to insufficient time for the researcher to conduct the observations. The lack of an appropriate number of observations did not allow the investigator to establish stability in the baseline data. Also, due to various factors such as absences, school events, and time off school there is a large gap between the last baseline data point to the first intervention data point for Teacher 3. The gap in data did not allow the investigator to determine if student on-task behavior remained at baseline prior to the introduction of the intervention.

The study included a small number of classrooms (n = 3) and an overall small number of students (n = 9). Although there were positive changes for five of the six at-risk students and all students demonstrated adequate levels of on-task behavior, a larger number of students would provide a more robust understanding of how CW-FIT improves on-task behavior. Another limitation that should be acknowledged is timed event recording was not used for praise/reprimand rate to allow for point by point agreement. This may have resulted in the two observers recording different statements as praise or reprimands but still showing adequate agreement.

Finally, the measure that was used to determine if the students were satisfied with their classrooms was a modified version of a validated measure. While the items used were from a validated measure, not all items were provided to the students and the items were modified for the current study. Moreover, the existing measure from which the questions were taken from, the EvsD (Skinner et al., 2009), was validated on students from third grade to tenth grade. Although one of the classrooms does fall within that age range, the remaining two do not. It is unknown whether this measure would generalize to younger students.

Recommendations for Future Research

There are several avenues for future research to extend the results of this study. First, future research on the CW-FIT intervention should include more baseline data points to examine generalization. Collecting data about generalization as a primary research question, where data

collection was focused on the generalization piece of the study, would be an important next step. There was some slight evidence that effects generalized, but the confidence in those findings is very small due to the previously mentioned limitations.

Additionally, it would also be important to track other behaviors that may change with CW-FIT such as remaining seated, raising hand, which are taught as part of the intervention. Conklin et al. (2017) tracked these behaviors during the intervention and noticed an increase in those behaviors. Therefore, it may be helpful to track these behaviors in addition to on-task behaviors to determine if generalization occurred with these other behaviors as well to other parts of the day during which the intervention is not implemented.

While the teachers did not use the materials such as posters, timer, or point sheet during the generalization time, two of the teachers were observed using the language that is on the poster to remind students of appropriate behaviors. Therefore, future research should measure teacher behaviors that are used as part of the CW-FIT intervention as well to see how their behavior generalizes to other periods of the day.

Furthermore, validating the modified EvsD rating scale that was used in the current study would be beneficial to further investigate student satisfaction in the classroom. As teachers were provided assistance in implementing the intervention with fidelity, future research examining the level of assistance teachers require may be beneficial in determining whether more assistance increases praise and on-task behaviors. Lastly, future studies should use more gender variety in the peer comparisons. Specifically, it would be beneficial to include two peer comparisons that represents each gender within one classroom as males tend to have more externalizing behaviors than females (Young et al., 2010).

The design that was used in the current study did not allow the researcher to determine if including or excluding various components of the intervention impacts outcomes. Therefore, future research should include a component analysis where the components of the intervention are introduced in phases. In other words, future researchers could implement the components of the intervention sequentially in a counterbalanced manner, such as beginning the intervention with teachers not providing rewards at the end of the session or not using praise or reprimands during the intervention session. Then those components could be introduced gradually to help determine which components of the intervention are necessary for improving on-task behavior.

Implications for Practice

CW-FIT is a well-established intervention that is easily implemented, requires only two hours of training, and few materials. It is well-accepted amongst students and teachers, both in the current study and in previous studies (Caldarella et al., 2015; Conklin et al., 2017; Hansen et al., 2017; Hirsch et al., 2016; Kamps et al., 2011; Weeden et al., 2016; Wills et al., 2018; Wills et al., 2014). The intervention improves not only student behavior but also teacher behavior. Students are more on-task during the intervention period and teachers praise more frequently. Additionally, as student teams are working towards the same goal, student relationships may improve as teamwork is reinforced. Students also feel better in the classroom while the intervention is being implemented. As teachers are providing frequent reinforcements and rewards, student-teacher relationships may improve as well as praise over reprimands is encouraged. Additionally, improved student attention to task has been shown to be an outcome of the intervention, which may improve student learning as they are more engaged in the lesson taught. Positive student behavior coupled with improved engagement to the material taught may improve student success and ultimately improve student outcomes in the future. Along with positive student outcomes, teachers may feel more successful and positive in their classroom which may result in a decrease in burnout and increase teacher job retention.

References

- Anderson, D. H., Munk, J. A. H., Young, K. R., Conley, L., & Caldarella, P. (2008). Teaching organizational skills to promote academic achievement in behaviorally challenged students. *Teaching Exceptional Children*, 40, 6–13. doi:10.1177/004005990804000401
- Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45, 369–386. doi:10.1002/pits.20303
- Archambault, I., Janosz, M., Fallu, J.S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence*, *32*, 651–670. doi:10.1016/j.adolescence.2008.06.007
- Baer, R. A., Blount, R. L., Detrich, R., & Stokes, T. F. (1987). Using intermittent reinforcement to program maintenance of verbal/nonverbal correspondence. *Journal of Applied Behavior Analysis*, 20, 179-184. doi:10.1901/jaba.1987.20-17
- Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research*, *19*, 3-30. doi:10.1177/0743558403258113
- 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, 119-124. doi:10.1901/jaba.1969.2-119
- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology*, 92, 568-582. doi:10.1037/0022-0663.92.3.568
- Beaman, R., Wheldall, K., & Kemp, C. (2007). Recent research on troublesome classroom behaviour: A review. *Australasian Journal of Special Education*, *31*, 45-60. doi:10.1017/S1030011200025586

- Berkeley, S., & Riccomini, P. J. (2011). Academic progress monitoring. In J. M. Kauffman & D.P. Hallahan (Eds.), *Handbook of special education* (pp. 334–347). New York, NY:Taylor & Francis.
- Brantley, D. C., & Webster, R. E. (1993). Use of an independent group contingency management system in a regular classroom setting. *Psychology in the Schools*, *30*, 60-66. doi: 10.1002/1520-6807(199301)30:1 3.0.CO;2-X
- Brown-Chidsey, R., & Steege, M. W. (2011). *Response to intervention: Principles and strategies* for effective practice. New York, NY: Guilford Press.
- Bru, E. (2006). Factors associated with disruptive behaviour in the classroom. *Scandinavian Journal of Educational Research*, *50*, 23-43. doi:10.1080/00313830500372000
- Bru, E. (2009). Academic outcomes in school classes with markedly disruptive pupils. Social Psychology of Education: An International Journal, 12, 461-479. doi:10.1007/s11218-009-9095-1
- Bryan, J., Day-Vines, N. L., Griffin, D., & Moore-Thomas, C. (2012). The disproportionality dilemma: Patterns of teacher referrals to school counselors for disruptive behavior. *Journal of Counseling & Development*, 90, 177–190. doi:10.1111/j.1556-6676.2012.00023.x
- Caldarella, P., Williams, L., Hansen, B. D., & Wills, H. (2015). Managing student behavior with Class-Wide Function-Related Intervention Teams: An observational study in early elementary classrooms. *Early Childhood Education Journal*, 43, 357–365. doi:10.1007/s10643-014-0664-3
- Caldarella, P., Williams, L., Jolstead, K. A., & Wills, H. P. (2017). Managing student behavior in an elementary school music classroom: A study of Class-Wide Function-Related

Intervention Teams. *Update: Applications of Research in Music Education*, *35*, 23–30. doi:10.1177/8755123315626229

- Conklin, C. G. (2010). The effects of Class-Wide Function-Related Intervention Teams (CW-FIT) on students' prosocial classroom behaviors (Unpublished doctoral dissertation). University of Kansas, Lawrence, KS.
- Conklin, C. G., Kamps, D., & Wills, H. (2017). The effects of Class-Wide Function-Related Intervention Teams (CW-FIT) on students' prosocial classroom behaviors. *Journal of Behavioral Education*, 26, 75–100. doi:10.1007/s10864-016-9252-5
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Merrill.
- Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology*, *36*, 59-82. doi:10.1016/S0022-4405(97)00050-2
- Cramer, M., Hirano, S. H., Tentori, M., Yeganyan, M. T., & Hayes, G. R. (2011). Classroombased assistive technology: Collective use of interactive visual schedules by students with autism. *CHI* 11,1-10.
- Curtis, M. J., Grier, J. E., & Hunley, S. A. (2004). The changing face of school psychology: Trends in data and projections for the future. *School Psychology Review*, *33*(1), 49-66. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquestcom.www2.lib.ku.edu/docview/620361655?accountid=14556
- De Pry, R. L., & Sugai, G. (2002). The effect of active supervision and pre-correction on minor behavioral incidents in a sixth grade general education classroom. *Journal of Behavioral Education*, 11, 255-267. doi:10.1023/A:1021162906622

Devadoss, S., & Foltz, J. (1996). Evaluation of factors influencing student class attendance and performance. American Journal of Agricultural Economics, 78, 499–507. doi:10.2307/1243268

Domitrovich, C. E., Bradshaw, C. P., Greenberg, M. T., Embry, D., Poduska, J. M., & Ialongo, N. S. (2010). Integrated models of school-based prevention: Logic and theory. *Psychology in the Schools*, 47(1), 71-88. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquest-com.www2.lib.ku.edu/docview/622098999?accountid=14556

- Doyle, W., & Carter, K. (1987). Choosing the means of instruction. In V. Koehler (Ed.), *Educators' handbook: A research perspective* (pp.188-206). New York: Longman
- Dunlap, G., Strain, P. S., Fox, L., Carta, J. J., Conroy, M., Smith, B. J., ... Sowell, C. (2006).
 Prevention and intervention with young children's challenging behavior: Perspectives regarding current knowledge. *Behavioral Disorders*, *32*, 29–45.
 doi:10.1177/019874290603200103
- Eagle, J. W., Dowd-Eagle, S. E., Snyder, A., & Holtzman, E. G. (2015). Implementing a multitiered system of support (MTSS): Collaboration between school psychologists and administrators to promote systems-level change. *Journal of Educational and Psychological Consultation*, 25, 160–177. doi:10.1080/10474412.2014.929960
- Elliott, S. N., & Von Brock Trueting, M. (1991). The Behavior Intervention Rating Scale:
 Development and validation of a pretreatment acceptability and effectiveness
 measure. *Journal of School Psychology*, 29, 43-51. doi:10.1016/0022-4405(91)90014-I

- Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review*, *5*, 273-297. doi: 10.1023/A:1020977107086
- Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review*, 11, 75-113. doi: 10.1007/s10567-008-0036-x
- Ferguson, E., & Houghton, S. (1992). The effects of contingent teacher praise, as specified by canter's assertive discipline programme, on children's on-task behaviour. *Educational Studies*, 18, 83-93. doi:10.1080/0305569920180108
- Fraser, B. J., & Fisher, D. L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal*, 19, 498-518. doi:10.2307/1162539
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109. doi:10.3102/0034654307400105
- Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, 47, 434–465. doi:10.3102/0002831209350494

Gresham, F., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment:
Principles, procedures, and future directions. *School Psychology Review*, *30*(2), 156-172.
Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquest-com.www2.lib.ku.edu/docview/619587146?accountid=14556

- Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA:
 Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools*, 43, 753–770. doi:10.1002/pits.20186
- Hawken, L. S., Sandra MacLeod, K., & Rawlings, L. (2007). Effects of the behavior education program (BEP) on office discipline referrals of elementary school students. *Journal of Positive Behavior Interventions*, 9, 94-101. doi:10.1177/10983007070090020601
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72, 625–638. doi:10.1111/1467-8624.00301
- Hansen, B. D., Caldarella, P., Williams, L., & Wills, H. P. (2017). Managing student behavior in dual immersion classrooms: A study of Class-Wide Function-Related Intervention
 Teams. *Behavior Modification*, 41, 626–646. doi:10.1177/0145445517698418
- 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. doi:10.1177/019874290503000204
- Hawken, L. S., & Horner, R. H. (2003). Evaluation of a targeted intervention within a schoolwide system of behavior support. *Journal of Behavioral Education*, *12*, 225-240. doi:10.1023/A:1025512411930
- Hawken, L. S., MacLeod, K. S., & Rawlings, L. (2007). Effects of the behavior education program on office discipline referrals of elementary school students. *Journal of Positive Behavior Interventions*, 9, 94-101. doi:10.1177/10983007070090020601

- 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.
- Hintze, J. M., & Christ, T. J. (2004). An examination of variability as a function of passage variance in CBM progress monitoring. *School Psychology Review*, *33*(2), 204-217.
 Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquest-com.www2.lib.ku.edu/docview/620462577?accountid=14556
- Hintze, J. M., Christ, T. J., & Methe, S. A. (2006). Curriculum-based assessment. Psychology in the Schools, 43, 45–56. doi:10.1002/pits.20128
- Hirsch, S. E., Healy, S., Judge, J. P., & Lloyd, J. W. (2016). Effects of an interdependent group contingency on engagement in physical education. *Journal of Applied Behavior Analysis*, 49, 975–979. doi:10.1002/jaba.328
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, 11, 189–196. Doi: 10.1901/jaba.1978.11-189
- Irvin, L. K., Tobin, T. J., Sprague, J. R., Sugai, G., & Vincent, C. G. (2004). Validity of office discipline referral measures as indices of school-wide behavioral status and effects of school-wide behavioral interventions. *Journal of Positive Behavior Interventions*, 6, 131-147. doi:10.1177/10983007040060030201
- Jamieson, D. W., & Thomas, K. W. (1974). Power and conflict in the student-teacher relationship. *Journal of Applied Behavioral Science*, 10, 321-336. doi: 10.1177/002188637401000304

- Jenkins, J. R., Zumeta, R., & Dupree, O. (2005). Measuring gains in reading ability with passage reading fluency. *Learning Disabilities Research and Practice*, 20, 245–253. doi:10.1111/j.1540-5826.2005.00140.x
- Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (Eds.). (2007). Handbook of response to intervention: The science and practice of assessment and intervention. New York, NY: Springer.
- Jolstead, K. A., Caldarella, P., Hansen, B., Korth, B. B., Williams, L., & Kamps, D. (2017). Implementing positive behavior support in preschools: An exploratory study of CW-FIT tier 1. *Journal of Positive Behavior Interventions*, 19, 48–60. doi:10.1177/1098300716653226
- Kamps, D., Wills, H., Dawson-Bannister, H., Heitzman-Powell, L., Kottwitz, E., Hansen, B., & Fleming, K. (2015). Class-Wide Function-Related Intervention Teams "CW-FIT" efficacy trial outcomes. *Journal of Positive Behavior Interventions*, *17*, 134–145. doi:10.1177/1098300714565244
- Kamps, D., Wills, H. P., Heitzman-Powell, L., Laylin, J., Szoke, C., Petrillo, T., & Culey, A.
 (2011). Class-Wide Function-Related Intervention Teams: Effects of group contingency programs in urban classrooms. *Journal of Positive Behavior Interventions*, *13*, 154–167. doi:10.1177/1098300711398935
- Kauffman, J. (2001). *Characteristics of emotional and behavioural disorders of children and youth* (7th ed.). Upper Saddle River, NJ: Merrill.
- Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers.*British Journal of Educational Psychology*, 77, 229-243. doi:10.1348/000709905X90344

Lane, T. W., & Murakami, J. (1987). School programs for delinquency prevention and intervention. In E. K. Morris & C. J. Braukmann (Eds.), *Behavioral approaches to crime and delinquency: A handbook of application, research, and concepts* (pp. 305-330).
Boston, MA: Springer.

Langdon, C. A. (1996). The third Phi Delta Kappa poll of teachers' attitudes toward the public schools. *Phi Delta Kappan*, 78(3), 244. Retrieved from http://search.ebscohost.com.www2.lib.ku.edu/login.aspx?direct=true&db=aph&AN=961 1087988&site=ehost-live

- Lannie, A. L., & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom:
 Effects of the good behavior game on student and teacher behavior. *Education and Treatment of Children, 30*, 85-98. doi:10.1353/etc.2007.0002
- Lembke, E. S., & Stichter, J. P. (2006). Utilizing a system of screening and progress monitoring within a three-tiered model of instruction: Implications for students with emotional and behavioral disorders. *Beyond Behavior*, *15*, 3-9.
- Lewis, R. (2001). Classroom discipline and student responsibility: The students' view. *Teaching* and Teacher Education, 17, 307-319. doi:10.1016/S0742-051X(00)00059-7
- Litoe, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. Journal of Applied Behavior Analysis, 8, 341–347. doi:10.1901/jaba.1975.8-341
- 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. doi:10.1177/0145445514554393
- Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students

with challenging behavior. *Journal of School Psychology*, *50*, 625-654. doi: 10.1016/j.jsp.2012.06.001

- Martinez, A., McMahon, S. D., & Treger, S. (2016). Individual and school-level predictors of student office disciplinary referrals. *Journal of Emotional and Behavioral Disorders*, 24, 30–41. doi:10.1177/1063426615588289
- McComas, J. J., Goddard, C., & Hoch, H. (2002). The effects of preferred activities during academic work breaks on task engagement and negatively reinforced destructive behavior. *Education and Treatment of Children*, 25(1), 103. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquestcom.www2.lib.ku.edu/docview/619740820?accountid=14556
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3), 276-282.
- McIntosh, K., Bohanon, H., & Goodman, S. (2010). Toward true integration of response to intervention systems in academic and behavior support: Part three: Tier 3 support. *Communique*, *39*(4).
- McIntosh, K., & Goodman, S. (2016). Integrated multi-tiered systems of support: *Blending RTI and PBIS*. New York, NY: Guilford Publications.
- Merrell, K. W., & Gueldner, B. A. (2010). *Social and emotional learning in the classroom: Promoting mental health and academic success*. New York, NY: Guilford Press.
- National Center for Education Statistics. (2003). *Status and trends in the education of racial and ethnic minorities*. Washington, DC: Author.

- Nelson, J. R., & Roberts, M. L. (2000). Ongoing reciprocal teacher–student interactions involving disruptive behaviors in general education classrooms. *Journal of Emotional* and Behavioral Disorders, 8, 27-37. doi:10.1177/106342660000800104
- Nichols, J. D. (2003). Prediction indicators for students failing the state of indiana high school graduation exam. *Preventing School Failure: Alternative Education for Children and Youth*, 47, 112–120. doi:10.1080/10
- Olive, M. L., & Smith, B. W. (2005). Effect size calculations and single subject designs. *Educational Psychology*, 25, 313-324. doi:10.1080/0144341042000301238 459880309604439
- O'Shaughnessy, T. E., Lane, K. L., Gresham, F. M., & Beebe-Frankenberger, M. E. (2003).
 Children placed at risk for learning and behavioral difficulties: Implementing a schoolwide system of early identification and intervention. *Remedial and Special Education*, 24, 27–35. doi:10.1177/074193250302400103
- Perry, D. F., Dallas Allen, M., Brennan, E. M., & Bradley, J. R. (2010). The evidence base for mental health consultation in early childhood settings: A research synthesis addressing children's behavioral outcomes. *Early Education and Development*, 21, 795-824. doi: 10.1080/10409280903475444
- Peterson, S. M. P., Caniglia, C., Royster, A. J., Macfarlane, E., Plowman, K., Baird, S. J., & Wu, N. (2005). Blending functional communication training and choice making to improve task engagement and decrease problem behaviour. *Educational Psychology*, 25, 257-274. doi: 10.1080/0144341042000301193
- Public Agenda. (2004). *Teaching interrupted: Do discipline policies in today's public schools foster the common good?* Retrieved from https://www.publicagenda.org

- Rausch, M. K., & Skiba, R. (2004). Disproportionality in school discipline among minority students in Indiana: Description and analysis (Children Left Behind Policy Briefs, Supplementary Analysis 2-A). Bloomington, IN: Center for Evaluation and Education Policy (ERIC Document Reproduction Service No. ED488897)
- Rhode, G., Jensen, W. R., & Reavis, H. K. (1993). *The tough kid book: Practical classroom management strategies*. Longmont, CO: Sopris West
- Romeo, F. F. (1998). The negative effects of using a group contingency system of classroom management. *Journal of Instructional Psychology*, *25*(2), 130-133. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquest-com.www2.lib.ku.edu/docview/619326888?accountid=14556
- Rose, T. L. (1988). Current disciplinary practices with handicapped students: Suspensions and expulsions. *Exceptional Children*, 55(3), 230-239. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquestcom.www2.lib.ku.edu/docview/617554569?accountid=14556
- Scates, V. E. (2005). Managing student behavior is necessary for academic success. In S. W.
 Watson (Ed.), *Culminating experience empirical and theoretical research projects*.
 Chattanooga: University of Tennessee.
- Scott, T. M., Park, K. L., Swain-Bradway, J., & Landers, E. (2007). Positive behavior support in the classroom: Facilitating behaviorally inclusive learning environments. *International Journal of Behavioral Consultation and Therapy*, *3*, 223-235. doi:10.1037/h0100800
- Sheldon, S. B. (2007). Improving student attendance with school, family, and community partnerships. *The Journal of Educational Research*, 100, 267–275. doi:10.3200/JOER.100.5.267-275

- Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 85, 257-268. doi:10.1093/ptj/85.3.257
- Skalski, A. K., Minke, K., Rossen, E., Cowan, K. C., Kelly, J., Armistead, R., & Smith, A. (2015). NASP. Practice Model Implementation Guide.
- Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline:
 Sources of racial and gender disproportionality in school punishment. *The Urban Review*, *34*, 317-342. doi:10.1023/A:1021320817372
- Skinner, C. H., Cashwell, C. S., & Dunn, M. S. (1996). Independent and interdependent group contingencies. *Special Services in the Schools*, 12, 61–78. doi:10.1300/J008v12n01_04
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic?. *Journal of Educational Psychology*, *100*, 765. doi: 10.1037/a0012840
- Skinner, E., Kindermann, T., & Furrer, C. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and 94 emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, 69, 493-525. doi: 10.1177/0013164408323233
- Spence, S. H., Donovan, C., & Brechman-Toussaint, M. (1999). Social skills, social outcomes, and cognitive features of childhood social phobia. *Journal of Abnormal Psychology*, 108, 211-221. doi:10.1037/0021-843X.108.2.211
- Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, 41, 225-234. doi:10.1037/0012-1649.41.1.225

- Sprick, R. S., Borgmeier, C., & Nolet, V. (2002). Prevention and management of behavior problems in secondary schools. In M. R. Shinn, H. M. Walker & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches; interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 373-401) National Association of School Psychologists, Washington, DC.
- 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. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquest-com.www2.lib.ku.edu/docview/619216701?accountid=14556
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*, 349–367. doi:10.1901/jaba.1977.10-349
- Sugai, G., & Homer, R. R. (2002). A promising approach for expanding and sustaining schoolwide positive behavior support. *School Psychology Review*, 35(2), 245–259. Retrieved from http://www2.lib.ku.edu/login?url=https://search-proquestcom.www2.lib.ku.edu/docview/621620444?accountid=14556
- Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., . . . Ruef, M. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions*, 2, 131-143. doi: 10.1177/109830070000200302
- Sugai, G., Simonsen, B., Freeman, J., & La Salle, T. (2016). Capacity development and multitiered systems of support: Guiding principles. *Australasian Journal of Special Education*, 40, 80–98. doi:10.1017/jse.2016.11

- Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behaviorspecific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders*, 8, 2-8, 26. doi:10.1177/106342660000800101
- Taylor-Greene, S., Brown, D., Nelson, L., Longton, J., Gassman, T., Cohen, J., . . . Hall, S.
 (1997). School-wide behavioral support: Starting the year off right. *Journal of Behavioral Education*, 7, 99-112. doi:10.1023/A:1022849722465
- Theodore, L. A., Bray, M. A., Kehle, T. J., & Dioguardi, R. J. (2004). Contemporary review of group-oriented contingencies for disruptive behavior. *Journal of Applied School Psychology*, 20, 79–101. doi:10.1300/J370v20n01_06
- Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The Good Behavior Game: 1969-2002. *Behavior Modification*, *30*, 225–253. doi:10.1177/0145445503261165
- Van Houten, R., & Nau, P. A. (1980). A comparison of the effects of fixed and variable ratio schedules of reinforcement on the behavior of deaf children. *Journal of Applied Behavior Analysis*, 13, 13–21. doi:10.1901/jaba.1980.13-13
- Vitaro, F., Brendgen, M., Larose, S., & Trembaly, R. E. (2005). Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of Educational Psychology*, 97, 617–629. doi:10.1037/0022-0663.97.4.617
- Vitaro, F., Larocque, D., Janosz, M., & Tremblay, R. E. (2001). Negative social experiences and dropping out of school. *Educational Psychology*, 21, 401-415. doi: 10.1080/01443410120090795
- Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). Antisocial behavior in school: Evidencebased practices (2nd ed.). Belmont, CA: Wadsworth/Thompson Learning.

- Walker, H. M., & Severson, H. H. (1990). Systematic Screening for Behavior Disorders (SSBD). Longmont, CO: Sopris West.
- Wallace, J. M., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991-2005. *The Negro Educational Review*, 59(1–2), 47–62. Retrieved from https://www-ncbi-nlm-nihgov.www2.lib.ku.edu/pmc/articles/PMC2678799/
- Weber, J. (2011). *Individual and family stress and crises*. Thousand Oaks, CA: Sage Publications.
- Weeden, M., Wills, H. P., Kottwitz, E., & Kamps, D. (2016). The effects of a class-wide behavior intervention for students with emotional and behavioral disorders. *Behavioral Disorders*, 42, 285–293. doi:10.17988/BD-14-12.1
- Wehby, J. H., & Kern, L. (2014). Intensive behavior intervention: What is it, what is its evidence base, and why do we need to implement now?. *Teaching Exceptional Children*, 46, 38-44. doi:10.1177/0040059914523956
- Wehby, J. H., Symons, F. J., & Shores, R. E. (1995). A descriptive analysis of aggressive behavior in classrooms for children with emotional and behavioral disorders. *Behavioral Disorders*, 20, 87-105. doi: 10.1177/019874299502000207
- Whitcomb, S. A., & Merrell, K. W. (2013). *Behavioral, social, and emotional assessment of children and adolescents* (4th ed.). New York, NY: Routledge/Taylor & Francis Group.
- Williams, K.C., & Williams, C.C. (2011). Five key ingredients for improving student motivation. *Research in Higher Education Journal*, 12(2), 1–23.

- Wills, H. P., Iwaszuk, W. M., Kamps, D., & Shumate, E. (2014). CW-FIT: Group contingency effects across the day. *Education and Treatment of Children*, 37, 191-210. doi: 10.1353/etc.2014.0016
- Wills, H. P., Kamps, D., Hansen, B., Conklin, C., Bellinger, S., Neaderhiser, J., & Nsubuga, B. (2010). The Classwide Function-Based Intervention Team Program. *Preventing School Failure: Alternative Education for Children and Youth*, 54, 164–171. doi:10.1080/10459880903496230
- Wills, H., Kamps, D., Fleming, K., & Hansen, B. (2016). Student and teacher outcomes of the class-wide function-related intervention team efficacy trial. *Exceptional Children*, 83, 58-76. doi:10.1177/0014402916658658
- Wills, H., Wehby, J., Caldarella, P., Kamps, D., & Romine, R. S. (2018). Classroom management that works: A replication trial of the CW-FIT program. *Exceptional Children*, 84, 437–456. doi:10.1177/0014402918771321
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior. *American Journal of Preventive Medicine*, 33, S130–S143. doi:10.1016/j.amepre.2007.04.011
- Winn, J. B. (2006). The Effects of Known Versus Unknown Criteria on Journal Writing Performance of Elementary Students Using an Independent Group Contingency (Unpublished doctoral dissertation). Retrieved from University of Tennessee, Knoxville, Tennessee
- Young, E. L., Sabbah, H. Y., Young, B. J., Reiser, M. L., & Richardson, M. J. (2010). Gender differences and similarities in a screening process for emotional and behavioral risks in

secondary schools. Journal of Emotional and Behavioral Disorders, 18, 225-235.

doi:10.1177/1063426609338858

Z Score Calculator for 2 Population Proportions. (n.d.). Retrieved January 10, 2020, from https://www.socscistatistics.com/tests/ztest/default2.aspx.

Appendix A Approval Letter from the University of Kansas Institutional Board



Date: May 31, 2019

TO: Lina Kitson, (lgoldenberg@ku.edu)

FROM: Alyssa Haase, IRB Administrator (785-864-7385, irb@ku.edu)

RE: Approval of Initial Study

The IRB reviewed the submission referenced below on 5/31/2019. Approval expires on 5/30/2022.

IRB Action: APPROVED		Effective date: 5/31/2	019	Expiration Date : 5/30/2022
STUDY DETAILS				
Investigator:	Lina Kitson			
IRB ID:	STUDY00144006			
Title of Study:	Class-wide Function-Related Intervention Teams			
Funding ID:	None			
REVIEW INFORMATION				
Review Type:	Initial Study			
Review Date:	5/31/2019			
Documents Reviewed:	Assent Script, CW-FIT Assent Script, CW-FIT Parent Consent Form, CW-FIT			
	Teacher Consent Form, • CW-FIT Teacher Email, • Lina Kitson CW-FIT Protocol, •			
	Rollin	g Ridge Principal Letter		
Expedited Category(ies):	 (7)(b) Social science methods 			
	• (5) I	Data, documents, records, or s	pecimens	
	• (7)(a) Behavioral research		
Special Determinations:	Chil	dren		
Additional Information:				

KEY PROCEDURES AND GUIDELINES. Consult our website for additional information.

- Approved Consent Form: You must use the final, watermarked version of the consent form, available under the "Documents" tab, "Final" column, in eCompliance. Participants must be given a copy of the form.
- Continuing Review and Study Closure: Submit a <u>Continuing Review</u> request and required attachments at least 4 weeks in advance of the expiration date. If Continuing Review is not approved before 5/30/2022, the study approval will expire on that date and all human subjects research activities must stop. Please close your study to IRB oversight once your study meets the first 4 milestones, as outlined in the <u>Closing a Study guidance</u>.
- Modifications: Prior to making any significant changes to the project, a <u>Modification</u> request must be submitted and approved.
- Add Study Team Member: <u>Complete a study team modification</u> if you need to add investigators not
 named in original application. Note that new investigators must take <u>the online tutorial</u> prior to being
 approved to work on the project.
- 5. Data Security: University data security and handling requirements apply to your project.
- 6. Submit a Report of New Information (RNI): If a subject is injured in the course of the research procedure or there is a breach of participant information, an RNI must be submitted immediately. Potential non-compliance may also be reported through the RNI process.
- Consent Records: When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity.

Human Research Protection Program Youngberg Hall | 2385 Irving Hill Rd | Lawrence, KS 66045 | (785) 864-7429 | research.ku.edwhrpp



Study Records must be kept a minimum of three years after the completion of the research. Funding
agencies may have retention requirements that exceed three years.

Appendix B Modified Version of the EvsD Scale

Please pick the number/face that best describes your feelings about each statement.



- 3. When I'm in class, I listen very carefully.
 - 1 2 3 4



Emotional Engagement

4. When I'm in class, I feel good.



5. I enjoy learning new things in class.



6. When we work on something in class, I feel encouraged.



7. Class is fun for me.



Appendix C CW-FIT Procedural Fidelity Checklist

CW-FT Procedural Fidelity Checklist

Modified from CW-FIT Teacher Implementation Manual

Date:	Teacher:	Time:
Jate:	reacher:	nme:

Subject:

Class Activity: Whole Group \Box Small Group \Box Independent \Box One-on-One \Box Transition \Box

CW-FIT Procedures	Rating
1. Getting attention skill displayed on poster.	Yes/No
2. Following directions skill displayed on poster.	Yes/No
3. Ignoring behavior skill displayed on poster.	Yes/No
4. Team point chart displayed.	Yes/No
5. Daily point goal posted.	Yes/No
6. Pre-corrects on skill "getting attention"	Yes/No
7. Pre-corrects on skill "following directions"	Yes/No
8. Pre-corrects on skill "Ignoring behavior"	Yes/No
9. Timer used during the entire session.	Yes/No
10. Timer is set at appropriate intervals.	Yes/No
11. Points awarded to teams for use of skills.	Yes/No
12. Praise/points to reprimand ratio approximately 4:1.	Yes/No
13. Teacher provides specific praise for use of skills.	Yes/No
14. Points were tallied at the end of the CW-FIT session.	Yes/No
15. Winners were announced.	Yes/No
16. Rewards were delivered to winning teams.	Yes/No

Appendix D Consumer Satisfaction Survey (Social Validity) Teacher Name:

Consumer Satisfaction Survey

Please circle the number which best describes your agreement or disagreement with each statement. Please answer each question.

1=Strongly Disagree	4=Slightly Agree
2=Disagree	5= Agree
3=Slightly Disagree	

1. This was an acceptable intervention for increasing on task behavior.

1	2	3	4	5

2. Most teachers would find this intervention appropriate for increasing on task behavior.

1 2 3 4 5

3. This intervention proved effective in increasing on task behavior.

1 2 3 4 5

4. I would suggest to use of this intervention to other teachers.

1 2 3 4 5

5. The problem of off task behaviors is severe enough to warrant use of this intervention.

1 2 3 4 5
	1	2	3	4	5	
7. I would be willing to use this intervention in the classroom setting.						
	1	2	3	4	5	
8. This intervention did not result in negative side effects for the students.						
	1	2	3	4	5	
9. This intervention was appropriate for a variety of children.						
	1	2	3	4	5	
10. This intervention is consistent with those I have used in classroom settings.						
	1	2	3	4	5	
11. The intervention was a fair way to handle off-task behaviors.						
	1	2	3	4	5	
12. This intervention was reasonable for decreasing off-task behaviors.						
	1	2	3	4	5	

13. I liked the procedures used in this intervention.

1	2	3	4	5	
14. This intervention	was a good wa	ay to handle off-ta	isk behaviors.		
1	2	3	4	5	
15. Overall, this intervention was beneficial for students.					
1	2	3	4	5	
16. This intervention quickly improved the child's behavior.					
1	2	3	4	5	
17. The intervention provided a lasting improvement in student off task behavior.					
1	2	3	4	5	
18. The intervention deviate from other c	improved the lassmates' beh	student's behavio avior.	r to the point t	hat it would not not	iceably
1	2	3	4	5	
19. Soon after using the intervention, I noticed a positive change in off-task behaviors.					
1	2	3	4	5	

20. On-task behaviors will remain at an improved level even after the intervention is discontinued

	1	2	3	4	5	
21. Usi setting	ng the intervo s (e.g., other	ention not only i classrooms, hon	mproved the chine).	ild's behavior in	the classroom, b	out also in other
	1	2	3	4	5	
22. Wł interve	nen comparing ention, the off	g off-task studer f-task students a	nts with well beh nd peer's behav	naved peers befo ior was more ali	ore and after use ike after using th	of the e intervention.
	1	2	3	4	5	
23. The behavi	e interventior or no long is a	n produced enou a problem in the	gh improvemen classroom.	t in the student'	's off-task behavi	iors so that the
	1	2	3	4	5	
24. Otł	ner behaviors	related to off-ta	ask behaviors als	o improved by t	he intervention.	

1 2	3	4	5
-----	---	---	---

Appendix E Student Consumer Satisfaction Questions (Social Validity)

Student Consumer Satisfaction Questions

- 1. Did you like playing the game in your classroom?
- 2. What did you like about it?
- 3. What did you not like about it?
- 4. Did you feel like the game helped you stay focused better?
- 5. What do you wish was different about the game?
- 6. Do you want to continue playing the game?