EFFECTS OF CLASS-WIDE FUNCTION-RELATED INTERVENTION TEAMS (CW-FIT) ON OFF-TASK GROUP BEHAVIOR IN A PRESCHOOL CLASSROOM

by

LAUREN LAYMAN

(Under the Direction of KEVIN AYRES)

ABSTRACT

Disruptive behaviors such as elopement, calling-out, and aggression are often a major barrier to instruction in preschool classrooms. One widely used class-wide behavior management system built around an interdependent group contingency is Class-Wide Functionally-Related Intervention Teams (CW-FIT). To date, the first author has only been able to find one study on CW-FIT in a preschool setting, by Jolstead et al. (2017) which found a therapeutic change in off-task behavior and rates of teacher praise statements and reprimands. The current study used a withdrawal design to evaluate the effectiveness of CW-FIT in a preschool classroom in both large and small group settings on both off-task student behavior and teacher praise and reprimand behavior. Results suggested that the implementation of CW-FIT decreased off-task group behavior in both settings. Results for rates of teacher's praise and reprimand statements were, however, variable for the two settings. Limitations are also discussed.

INDEX WORDS: CW-FIT, off-task behavior, group contingency, interdependent group contingency, preschool, behavior game, problem behavior, challenging behavior, reprimands, praise, large group, small group

EFFECTS OF CLASS-WIDE FUNCTION-RELATED INTERVENTION TEAMS (CW-FIT) ON OFF-TASK GROUP BEHAVIOR IN A PRESCHOOL CLASSROOM

by

LAUREN LAYMAN

BS, The University of Georgia, 2018

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2018

EFFECTS OF CLASS-WIDE FUNCTION-RELATED INTERVENTION TEAMS (CW-FIT)

ON OFF-TASK GROUP BEHAVIOR IN A PRESCHOOL CLASSROOM

by

LAUREN LAYMAN

Major Professor:	Kevin Ayres
Committee:	Joel Ringdahl Scott Ardoin

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

TABLE OF CONTENTS

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

For the majority of students, preschool is their first introduction to an organized school environment where they must follow specific rules, perhaps except children who attended school-like daycare facilities. Thus, most teachers would probably agree that disruptions and problematic behavior are a big problem, especially in a preschool classroom. Children that have unresolved problem behaviors, such as aggression, disruption, elopement, etc., tend to have persistent behavior and academic difficulties (Wills et al., 2010). Typically, in a classroom when one or two students display problem behaviors, the teacher or a behavior analyst manage these behaviors by creating individual behavior intervention plans. However, multiple students in need of behavioral support make individual behavior management unfeasible. Therefore, teachers need clear and straightforward class-wide behavior management systems to help manage prevalent behavior problems. Many teachers have used group contingencies to increase class-wide on-task behavior in elementary classrooms, leading to decreased disruptions and more instructional time (Chow & Gilmour, 2015). One widely used class-wide behavior management system, which utilizes a token economy embedded in an interdependent group contingency, is Class-Wide Functionally-Related Intervention Teams (CW-FIT).

CW-FIT is a three-tiered system which has focused interventions at the class-wide and individual levels. Tier 1 utilizes a group contingency to focus on a whole class approach. The second tier adds self- and peer-management components for target students that do not respond to the tier 1 intervention alone. Students who fail to respond to both tier 1 and 2 interventions are

candidates for tier 3 which includes a functional behavior assessment to further understand the causes for each target student's disruptive behavior.

The current study focuses on the tier 1 portion discussed in Wills et al. (2010) on class-wide behavior management. The main components of the intervention include teaching communication/social skills (e.g., getting the teacher's attention), eliminating (both teacher and peer) attention for problem behavior, and reinforcement of appropriate behaviors through contingent rewards. To achieve these elements, teachers strategically split-up students into about four-six teams, ensuring that one group does not contain an unusually high proportion of students with behavioral problems. They then introduce the students to the three key communication/ social skills (getting the teacher's attention, ignoring the problem behavior of peers, and following directions the first time) through behavioral skills training (BST). Teachers also have the option of introducing additional skills depending on unique problem behaviors or other goals for the classroom. During the BST lesson, the teacher vocally instructs students on the skill, provides modeling of examples and non-examples, allows the students time to practice the skill, and gives feedback on their performance (Himle, Miltenberger, Gatheridge, & Flessner, 2004). Teachers may occasionally display posters of the skills including corresponding pictures in the room with four to five steps on how to implement each skill to remind students of the rules. For example, the steps to getting the teacher's attention appropriately include "step 1: look at the teacher, step 2: raise your hand, step 3: wait for the teacher to call on you, and step 4: ask your question or give your answer" (Wills, Kamps, Fleming, & Hansen, 2016, p. 64).

For each intervention session, the teacher begins with a review of the social skills and an announcement on the goal number of points for that session. They then continue with regular instruction while also providing praise and awards points after a timer for a predetermined

interval is heard. Teachers typically begin the intervals at two or three minutes and then progressively increases it to between eight and ten minutes. The teacher supplies these points to teams in which every member of that team was following the skills/rules outlined during the BST for the entire interval. They also display the points on a chart posted on the board or at the front of the room. At the end of the session, the teacher counts the number of points acquired by each team and distributes rewards (e.g., edibles or activity-based rewards like extra recess) to the teams that meet the predetermined goal. The teacher can gradually fade the intervention by using longer intervals or decreasing the number of times the game is played per day or week (Wills et al., 2010).

Researchers have shown CW-FIT to be effective in increasing on-task behaviors and praise for appropriate behaviors by teachers for kindergarten through seventh grade general education classes (Caldarella, Williams, Hansen, & Wills, 2015; Caldarella, Williams, Jolstead, & Wills, 2017; Conklin, Kamps, & Wills, 2017; Hansen, Caldarella, Williams, & Wills, 2017; Kamps et al., 2015; Kamps et al., 2011; Schmidt Naylor, Kamps, & Wills, 2018; Weeden, Wills, Kottwitz, & Kamps, 2016; Wills, Iwaszuk, Kamps, & Shumate, 2014; Wills et al., 2016; Wills et al., 2010; Wills, Kamps, Wehby, & Caldarella, 2016). Researchers have also evaluated CW-FIT for use in non-academic classes such as music classes (Caldarella et al., 2017), as well as for students with identified disabilities such as emotional and behavioral disorders (Kamps et al., 2011; Wills et al., 2016) and English Language Learners (Schmidt Naylor, Kamps, & Wills, 2018).

However, to date, researchers have only been able to locate one empirically evaluated study on CW-FIT in a preschool setting by Jolstead et al. (2017). In this study, the researchers used a multiple-baseline design with embedded withdrawal conditions across four separate

classrooms (of 13 or 14 children in each) in two different schools in Utah to evaluate the effectiveness of CW-FIT on increasing on-task behavior. The researchers identified the dependent variables in this study as group on-task behavior and rates of teachers' praises and reprimands. Jolstead et al. (2017) showed through their results that on-task behavior increased from the first baseline condition to the first intervention condition by 17.25% and from the second baseline to the second intervention conditions by 13.16%. These results also showed an overall, although variable, improvement in the ratio of teacher's praises to reprimands (Jolstead et al., 2017).

Teachers in the Jolstead et al. (2017) study implemented the intervention during circle time in which all children sat on the floor and participated in various activities (e.g., academic learning, dancing, singing, etc.) or during center-time in which children rotated between different centers focusing on letters, number, artwork, etc. The researchers did not separate or differentiate these settings when reporting results. However, large group (i.e. circle time) and small group (i.e. center time) settings similar to those used by Jolstead et al. (2017) would have most likely led to differentiated rates of reinforcement (especially between small groups that did or did not have a teacher attending to them), teacher-to-student ratios, and task difficulty. Another study by Wills et al. (2014), also looked at responses to CW-FIT intervention across three settings: spelling or writing, math, and science or social studies. However, they did not discuss the differences between the settings (e.g., independent versus teacher-led work, whole class versus small group activities, etc.).

The purpose of the current study was to further evaluate the effectiveness of CW-FIT in a preschool setting in both large group and small group settings. The first author ran two concurrent experiments to test the following research questions are: What effect will CW-FIT

have on group off-task behavior in a large group (i.e., whole class) setting? What effect will CW-FIT have on rates of teacher praise and reprimand statements in a large group setting? What effect will CW-FIT have on group off-task behavior in a small group setting? What effect will CW-FIT have on rates of teacher praise and reprimand statements in a small group setting?

CHAPTER 2

EXPERIMENT 1

Method

Participants. The present experiments both occurred in the same preschool classroom located within a Title I elementary school in Georgia. Preschool classrooms in the district are state-funded, and any child living within the district lines can attend as long as they are or will be four-years-old by September 1st of the current school year in which they enroll.

The teacher of this classroom had previously requested help with problem behavior and disruptions from the main researchers, two graduate-level behavior analyst interns. This teacher obtained her degree in early childhood education, preschool through fifth grade, and has been teaching for 15 years. The teacher spent five of these years teaching preschool; however, this was her first year at the present school. A paraprofessional also worked in the room during the entire day but had requested that her information and data not be reported. There were 21 total students in the classroom, with 11 (52.38%) boys and 10 (47.62%) girls. All students at the start of the study were four-years-old. Two students in the classroom were previously referred for language evaluations and have since begun one-on-one or small group instruction with a speech therapist or behavior analyst. As decided by the main researchers, one of these two students did not participate in the study since she struggled to follow one-step directions and therefore, did not possess the prerequisite skills to participate in the game. Her behavioral and academic needs were met individually outside of the planned intervention.

Setting and arrangements. In the first experiment, the researchers collected data during large group activities from 8:30-8:50 and 10:00-10:30. During all large group activities, students

sat in their assigned square on the carpet, and the teacher read a book to the students or lead the class in a class-wide activity (e.g., identifying letters, talking about things or people in their community). She encouraged student participation by asking them to give examples and answer questions (individually and chorally). A paraprofessional in the room usually participated in such activities by pulling misbehaving students from the carpet to sit by her in the back to keep their behavior better in check or issuing reprimands from the back of the room.

Before the study, the teacher reported that she occasionally used a classroom management system that utilized a chart at the front of the room which displayed envelopes labeled with each child's name that showed colored cards corresponding to the students' behavior. The teacher ranked the cards by color from best to worst in this order: blue, white, green, yellow, and red. Students started on the green card and could be moved up to a white card and then possibly to a blue card for being exceptionally well-behaved. The teacher awarded a sticker to students who ended the day with a blue or white card. Conversely, a yellow or red card indicated students that are disruptive or weren't following directions. Students with these colored cards were forced to sit out (two minutes for a yellow card or five minutes for a red card) of recess or center-time (free play time in the classroom at the end of the day) depending on the time of day. The teacher did not articulate clear rules for how students would proceed from one color to another. However, she did display written rules at the front of the classroom, even though the majority of the class could not read yet. Unfortunately, the teacher rarely reviewed these rules with the students before activities.

Materials. During the intervention, the teacher displayed posters listing the CW-FIT social skills (how to get the teacher's attention appropriately, how to ignore the inappropriate behavior of peers, and how to follow directions the first time) and steps for how to complete the

skills at the front of the room. The first author also helped to put colored tape (red, green, yellow, and purple) on the floor to visually mark where teams sat for large group activities. Corresponding colored cars were used on a number line to represent each team's progress during the game. The first author created this number line (see Appendix A) to resemble a road in which the cars could travel down to the "finish line." The first author decided to use this number line in place of a tally chart on the board in which the implementer tallies points for each team because many students in the classroom were not able to count or discriminate numbers. The teacher and data collectors also used three timers during the intervention: one to time the 10-minute sessions, another to take data on group-off task behavior using momentary time sampling (see dependent measures section), and the last to time intervals for points allocation.

Experimental design. The first experiment used a withdrawal design to evaluate for a functional relation between the implementation of CW-FIT and group off-task behavior in a large group setting and rates of teacher's praise and reprimands. The first author made decisions on when to advance to the following condition contingent on stable data of the primary dependent variable, group off-task behavior. For example, once there were four stable data points for both large-group and small-group settings during the first baseline condition, the first intervention condition began.

Baseline. The researchers collected baseline data for four sessions until data stabilized during the baseline condition. Baseline also lasted for a total of five sessions during the withdrawal condition. Classroom activities and management continued as described in the settings and participants section.

Teacher training. The main researchers trained the teacher in the implementation of Tier 1 CW-FIT after the baseline condition and before beginning the intervention. Training included

an informal explanation of the game and a permanent, written description of the reason for intervention, how to implement CW-FIT, and examples and non-examples of descriptive praise and reprimands. The researchers modeled implementation of the intervention for two weeks before the teacher took over implementation. The first author continued to provide coaching and feedback (in the form of answering questions and making suggestions) when needed. After the withdrawal of intervention, and before the return to the intervention condition, the teacher and first author went through a brief retraining and discussed areas for improvement. During the second intervention condition, the teacher resumed implementation of the intervention until the conclusion of the study.

Behavioral skills training (BST). The first author completed the behavioral skills training (BST) with the students in the classroom during a 15-minute lesson (Himle et al., 2004). She presented each skill separately following the BST protocol: verbal instruction of the behavior, modeling of the behavior, student practice of the behavior, and feedback from the first author (Himle et al., 2004). The skills taught include: how to appropriately get the teacher's attention, how to follow directions the first time, and how to the ignore the inappropriate behavior of peers (Wills et al., 2010). The teacher also displayed the CW-FIT posters with steps and pictures for how to complete the skill correctly at the front of the room during the intervention sessions.

The first author and teacher also led a shorter BST lesson (five minutes) that reviewed the skills/rules with the students in between the withdrawal and second intervention conditions. The teacher used brief mini-lessons (two to three minutes long) before each intervention session which included a quick review of the social skills poster, modeling of appropriate and inappropriate behaviors, student practice, and feedback from the teachers and researchers.

Intervention. The first author developed the intervention used in this experiment from the CW-FIT Tier 1 portion described in Wills et al. (2010) with a couple of changes, such as using the number line in place of a points chart, to make the intervention more accommodating for the younger students. The researchers presented the intervention to the students in the form of a game. The teacher implemented the intervention for three sessions in the first intervention condition and five sessions in the second intervention condition.

Teams. For large-group activities, the teacher assigned students to four groups of five students each based on their assigned seat on the carpet. The teacher designed the groups to have roughly the same number of problem students per group. The teacher also identified these problem students and formed groups before baseline. These groups remained the same throughout the entire study, and the teacher did not rearrange them for absent students to reduce confusion. This resulted in some groups having less than five students on some days.

Goals, points, and praise/reprimands. The teacher determined a goal number of points needed for each team to win and announced it to the class before the start of each intervention session. She started the initial goal at one point and had increased it to two points by the end of the study. The teacher placed a red sign that says "finish line" on the number line as a visual representation of the goal (see Appendix A). Each intervention session lasted between five and ten minutes. At one-minute intervals, the teacher awarded points to the groups that had followed the rules of the game during the entire interval. The first author had intended that the students receive frequent descriptive praise from their teacher and that attention to problem behaviors be reduced/ eliminated. This target, however, was not successful (see results on teacher praise and reprimands).

Rewards. At the end of the intervention session (or the end of the lesson if the lesson lasted longer than ten minutes), the teacher reviewed the final number line with the students. Teams that met the goal were once again praised for their good behavior and given a small reinforcer. Reinforcement items included small edibles (usually small candies like skittles or fruit snacks). The teacher also gave redirections and suggestions for the following game to the teams that did not meet the goal, and therefore, did not receive the reward.

Treatment fidelity. For each intervention session, the first author filled out a 14-point treatment fidelity checklist (see Appendix B). The teacher was able to implement CW-FIT Tier 1 with an average treatment fidelity of 74.70% (range: 71.43-83.33%). Although this fidelity is not ideal, this low percentage was consistently because teachers and paraprofessional were not able to maintain the goal number of praises (one or more per minute) and reprimands (one or less per five minutes) statements (LaBrot, Pasqua, Dufrene, Brewer, & Goff, 2016). Teacher's praise and reprimand statement fidelity were at 13% and 25%, respectively. Additionally, the paraprofessional's praise and reprimand statement fidelity were both at 0%. On the other hand, there were many steps that were carried out 100% fidelity including: teacher displaying materials (skills posters, number line, and cars) prominently in the room, splitting students up into three to five groups, reviewing the BST skills, using the timers appropriately for interval, recording points at the end of each interval, giving feedback when awarding points, reviewing the number line and number of points at the end of each trial, immediately giving out rewards, and recording data from the number line at the end of each trial. The teacher also discussed the goal number of points with the class with 88% fidelity.

Social validity. The main researchers sent the teacher an online questionnaire before baseline data collection. On the questionnaire, the teacher indicated the current level and system

of classroom management, the presence and severity of problem behaviors (e.g. aggression, disruption, out of seat/area, following directions, call-outs, etc.), the time during the day in which she needed the most help, and possible reinforcers that she was willing to try-out or use during the intervention conditions. The first author also administered a modified preference assessment to each student at the end of baseline to probe potential edible reinforcers. Preferred edibles were fruit snacks, smarties, and skittles.

Dependent measures and variables.

Group off-task behavior. The primary dependent variable used by Jolstead et al. (2017), group on-task behavior was similar to the primary dependent variable measured in the current study. Due to logistical reasons, the first author decided to measure group off-task behavior. The first author defined off-task behavior as nonengagement in the assigned classroom activity or any time when the student is focusing their attention anywhere other than the person/object the students were told to direct their attention towards including: watching another off-task peer, laying down on the floor, calling-out, talking with another peer, playing with or touching objects or other peers that are not part of the assigned activity, out of area/leaving the group, etc. Exceptions to the rule included if a student was fidgeting or playing with something (like their clothing), but their attention was still on (eyes pointed towards) the teacher or another specified object. For example, if a student is playing with his shoelaces, but still has his eyes on the teacher, he is on-task.

Momentary time sampling system by groups was used to measure group off-task behavior (Cooper, Heron, & Heward, 2007). Every 30 seconds the researcher(s) looked at each group separately and recorded if any student in that particular group was off-task. Therefore, if any student in the group was off-task, the entire group was marked as off-task. Only if all

students in a particular group were on-task at the moment of observation, would the data collector record the group as on-task. Groups used for data collection were identical to the teams used during the intervention and were determined pre-baseline.

Rates of teacher's praise and reprimands. The secondary dependent variables were rates of teacher praise and reprimanding statements. The primary researcher defined a praise statement as any verbal statement or gesture approving of appropriate behavior that is further than the acknowledgment of the behavior. On the other hand, the first author defined a reprimand as any verbal statement or gesture disapproving of inappropriate behavior. However, whole-group redirections in which the teacher did not identify individual students were not considered reprimands. For example, if the teacher said "remember class, if we want to earn points, we need to be following the directions" but did not single out any students, the data collectors did not count it as a reprimand. Redirections, unlike reprimands, call attention to the appropriate behavior and not to the inappropriate or problem behavior.

Data collectors originally recorded the frequency of each (teacher praise and reprimands) using a tally procedure per session, but they changed to a tally per interval after three sessions (see interobserver agreement section). The data collectors synchronized these 30-second intervals to the 30-second intervals used for the data collection of off-task behavior.

Interobserver agreement (IOA). Before baseline began, the first and second author (main researchers) were trained on and practiced distinguishing instances of off-task group behavior and teacher praise statements and reprimands. The majority of sessions (94.12%) were coded live by having two data collectors in the room. However, when two data collectors were not available to be in the classroom at the same time, video and audio recordings of the

classroom were taken in one session (5.88%). The second author coded these videos at a later time.

Off-task group behavior. When taking live IOA data, the primary data collector signaled which group to look at and when by holding up a number of fingers to indicate the group number. For example, when the primary data collector held up one finger to signal "group one," both data collectors directed their attention to group one and then recorded their observation. Then they both waited for the primary data collector to direct them to look at the next group. The data collectors used this procedure for every group at each 30-second interval. Data for off-task group behavior were collected simultaneously by two data collectors in 29.41% of all sessions (33.33% of baseline and 20.00% of intervention conditions). The first author calculated agreement percentages for off-task group behavior by using point-by-point agreement. This agreement averaged 75.69% overall (81.49% for baseline and 66.80% for intervention conditions) during the first experiment.

Rates of praise and reprimands. Data with IOA was collected in 29.41% of all sessions (33.33% of baseline and 20.00% of intervention conditions). Agreement data on teacher rates of praise and reprimand statements were originally calculated using a gross method (larger amount divided by the smaller amount) through the third session. However, after the fourth session calculating the data this way, the first author decided to use a mean-count per interval method (that combined data collection for both praise and reprimands) instead to ensure that both data collectors were recording the same instances of behavior. With this method, the IOA of each interval was calculated and then averaged together. This resulted in an overall average for IOA of 82.22% for praise and reprimands (82.11% during baseline and 82.38% during intervention).

Results

The first author used visual trend analysis, consistent with single-subject experimental design, to evaluate the results of both experiments by analyzing the level, trend, and variability of off-task group behavior and rates of teacher's praise and reprimand statements. Overall, results show that CW-FIT decreased group off-task behavior in the large group settings. Additionally, results were variable for rates of teacher praise and reprimands.

During the first experiment in the large group setting the researchers saw therapeutic changes to off-task group behavior (see Figure 1). Data was stable with a zero-celebration trend during baseline with an average of 78.12% off-task behavior (SD=7.47). During the first intervention condition, off-task group behavior was stable with a decreasing trend and averaged 45.43% (SD=6.06). During the withdrawal condition, levels of off-task group behavior returned to near baseline levels, averaging 75.66% (SD= 11.79) with an increasing trend. Behavior once again decreased to an average of 49.40% (SD= 3.99), where behavior was on a slightly decreasing trend.

Teacher behavior was highly variable during the first experiment (see Figure 2). During the baseline condition, the average rate of praise was 0.26 per minute (SD=0.21) and 0.71 per minute (SD= 0.46) for reprimands. During first intervention condition, the teacher was able to increase praise to an average of 0.64 per minute (SD= 0.25) and decrease reprimands to an average of 0.49 per minute (SD= 0.44). Levels of praise decreased to an average of 0.41 per minute (SD= 0.29), and levels of reprimands increased to an average of 0.73 per minute (SD=0.64) during the withdrawal condition. Praise decreased to 0.37 per minute (SD=0.11) and reprimands, however, decreased to 0.64 per minute (SD= 0.37) during the second intervention condition.

Figure 1

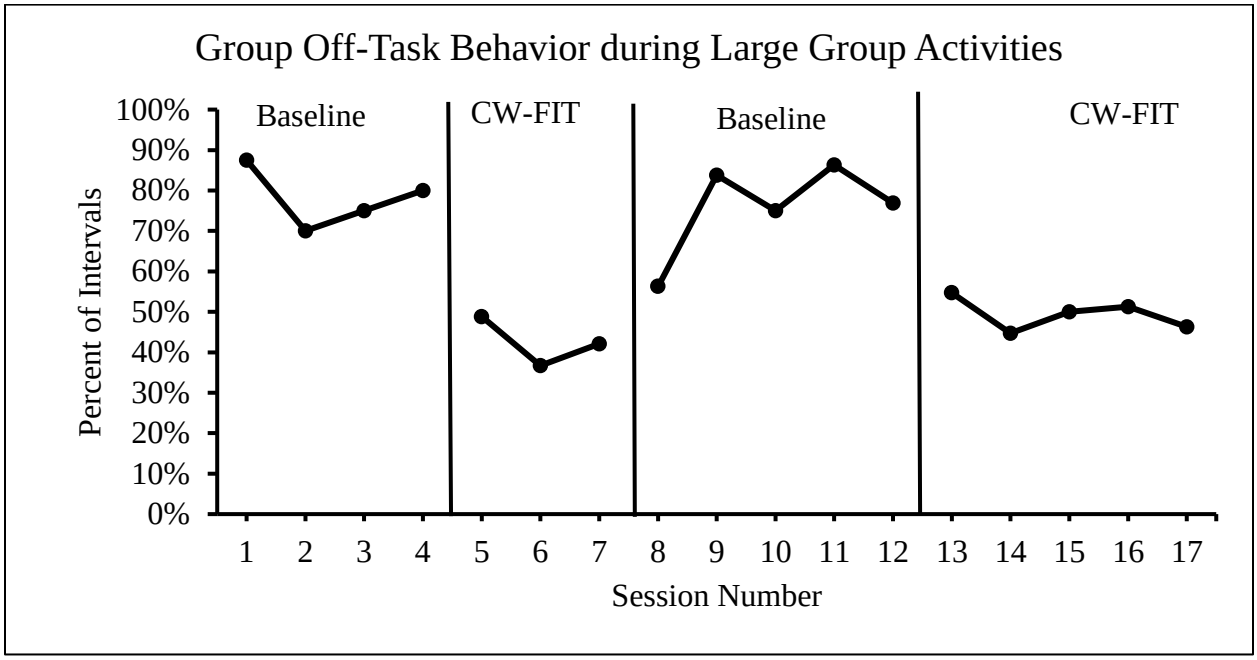
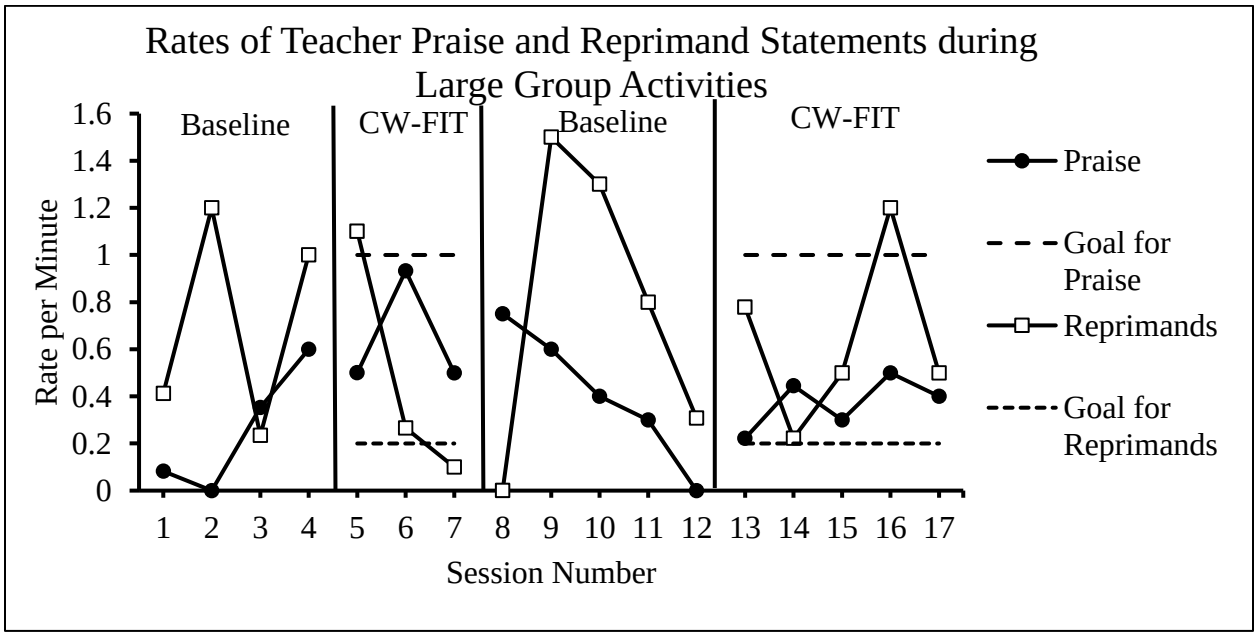


Figure 2



CHAPTER 3

EXPERIMENT 2

Method

For experiment 2, all procedures and protocols are identical to experiment 1, except where noted below.

Settings and arrangements. Small group activities took place between 8:50-9:10 and 10:30-10:45 in the morning during the second experiment. During these activities, the teacher divided students into four groups to complete different assignments in different areas of the room. The teacher led one group on the carpet, while the paraprofessional led another in the back of the room. The remaining two groups completed an independent activity at two separate tables. All groups were completing unconnected activities. The groups rotated between teacher-led and independent settings daily. The teacher used the same classroom management system during small group activities as she did during large group activities. The teacher was generally responsible for the behavior of the two independent groups as well as her group.

Materials. During small group activities, a colored, paper car that coordinated with the number line's colored cars was also placed on the table or floor in the middle of each small group to indicated the team's color for that activity.

Experimental design.

Baseline. Data collectors collected baseline data for four sessions during the baseline condition. The researchers also collected data during the withdrawal condition for a total of three sessions.

Intervention. The first intervention condition lasted two sessions (cut short by Thanksgiving Break) and the second intervention lasted for three sessions.

Teams. For small group activities, the teacher had already divided the classroom up into four groups of five students for academic instruction. These groups had been designed to even out the number of problem students (identified by their teacher) across the groups.

Treatment fidelity. The teacher was able to implement CW-FIT Tier 1 with an average treatment fidelity of 78.57% (range: 75.00-85.71%) during small group activities. As in experiment 1, the teacher and paraprofessional struggled to maintain the goal number of praises and reprimands. Fidelity for the teacher's praise and reprimands were 20.00% and 40.00%, respectively. The paraprofessional also had a 0% fidelity for both praise and reprimands. Additionally, the teacher only reviewed the BST skills and goal number of points with 80% fidelity. On the other hand, the teacher presented the CW-FIT materials (social skill posters, number lines, colored cars, etc.), split the class into groups, used the timer correctly, recorded points at each interval, provided feedback when awarding points, reviewed the points at the end of the session, provided rewards, and recorded the number of points for each team all with 100% fidelity.

Dependent measures and variables.

Interobserver agreement (IOA). Two data collectors in the room live-coded the majority of sessions (91.67%). The first author took video and audio recordings of the classroom during one session (8.33%).

Off-task group behavior. Data collectors simultaneously collected data for off-task group behavior in 41.67% of all sessions (43.86% of baseline and 40.00% of intervention conditions) in

the second experiment. This agreement averaged 77.08% overall (73.20% for baseline and 82.90% for intervention conditions).

Rates of praise and reprimands. Agreement percentages calculation were originally calculated using a gross method system until the fifth session when it was changed by the first author to a mean-count per interval method (see experiment one for explanation). Researchers collected IOA for this variable in 25.00% of all sessions (28.57% of baseline and 20.00% of intervention) and averaged 74.32% overall (65.49% for baseline and 92.00% for intervention conditions).

Results

During the second experiment in the small group setting, data on off-task group behavior looked similar to that in the first experiment (see Figure 3). The off-task group behavior during baseline was on a stable, accelerating trend, averaging 80.65% (SD= 4.67). This behavior averaged 46.90% (SD=4.38) with a decreasing trend during the first intervention condition. It once again returned towards baseline levels with a stable, zero-celeration trend during the withdrawal condition with an average of 72.17% (SD= 1.42). This behavior again decreased to an average of 46.27% (SD=10.23) with a zero-celerating trend during the second intervention condition.

Rates of teacher's praise and reprimand statements in the second experiment were less variable and showed a more clearly defined pattern as compared to these variables in the first experiment (see Figure 4). Praise and reprimand statements during baseline averaged 0.11 per minute (SD=0.08) and 1.47 per minute (SD=0.91), respectively. The data showed a therapeutic change for both of these variables during the first intervention condition where praise and reprimands averaged 1.24 per minute (SD=1.08) and 0.54 per minute (SD=0.09), respectively.

Praise again decreased, and reprimands increased during the withdrawal condition where rates of praise and reprimands were 0.40 per minute (SD= 0.14) and 1.15 per minute (SD=0.21), respectively. Mirroring the change from baseline to the first intervention condition, the average rate of praise was 1.00 per minute (SD=1.25) and the average rate of reprimands was 0.3 per minute (SD=0.17) in the second intervention condition.

Figure 3

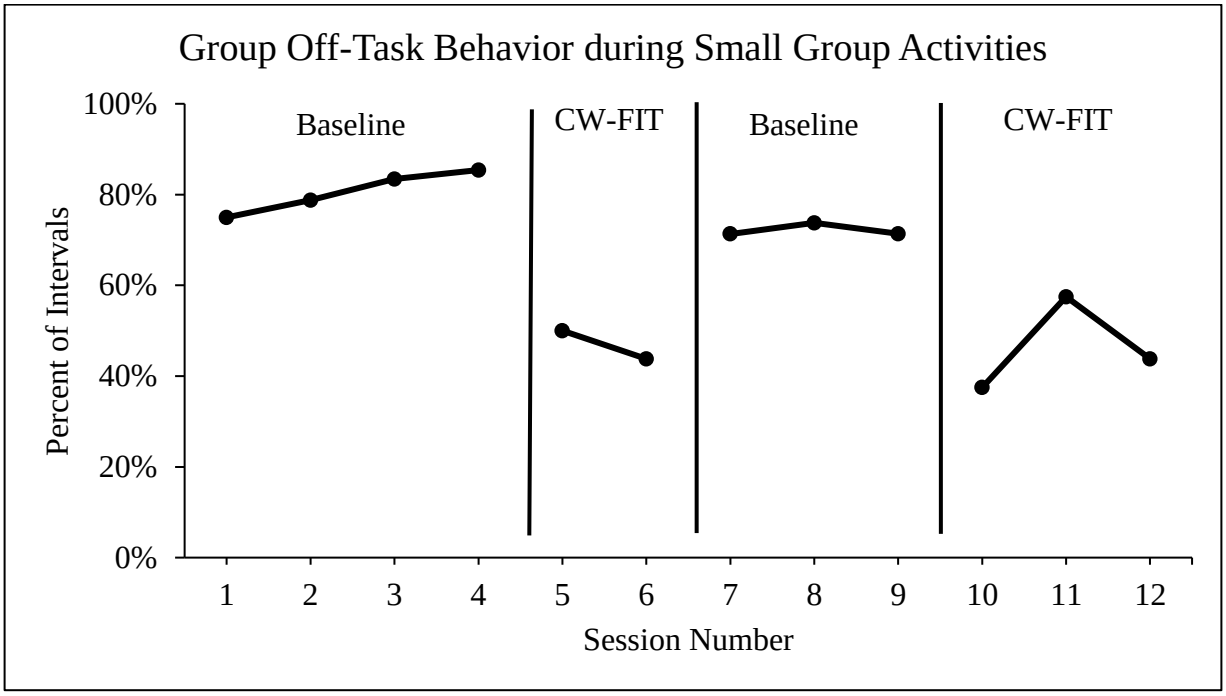
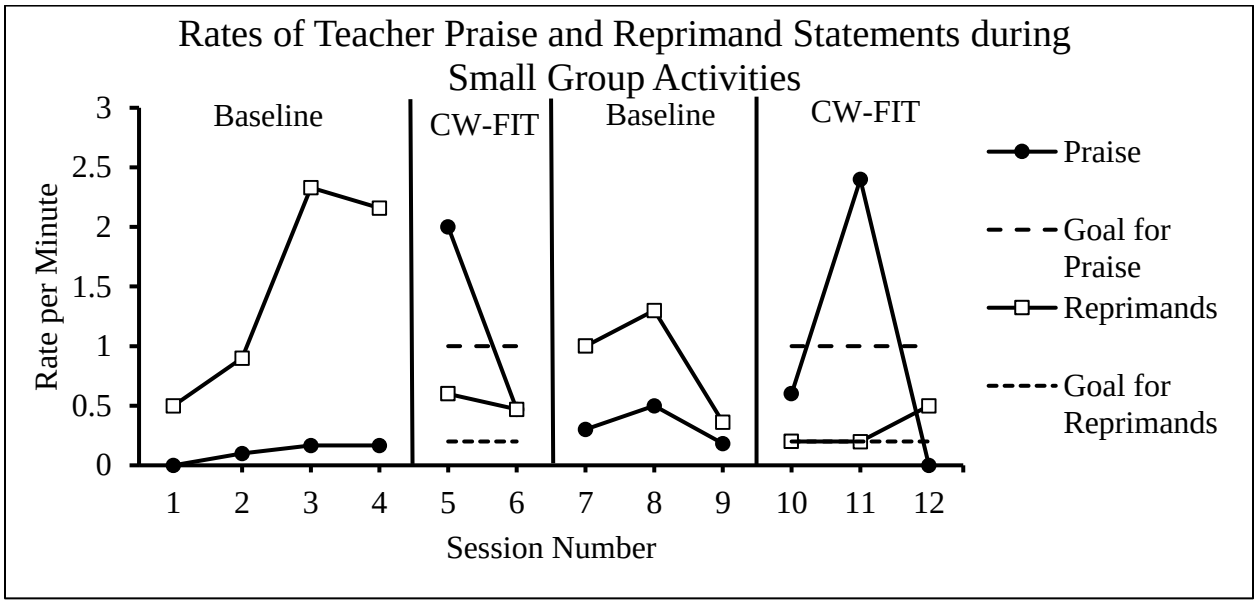


Figure 4



CHAPTER 4

GENERAL DISCUSSION

The purpose of this study was to evaluate the effectiveness of CW-FIT, an interdependent group contingency, in decreasing off-task group behavior and improving teacher behavior in both large group and small group settings. General finding from this study suggests that CW-FIT was effective in decreasing off-task group behavior, and therefore increasing on-task group behavior, during both large and small group settings. Effects of CW-FIT on rates of teacher praises and reprimands were, however, more variable and resulted in a clearer display of effects in the small group setting versus the large group setting. These findings are generally consistent with the findings of previous research on the effects of CW-FIT.

Similar patterns, which show an overall improvement on off-task group behavior regardless of settings, were seen during both large and small group settings (see Figure 5) across the two experiments. These findings are consistent with those found by previous research (e.g., Jolstead et al., 2017; Wills et al., 2010) which found increased on-task behavior correlated with the implementation of the CW-FIT intervention.

First, results from the first experiment showed that CW-FIT was effective in decreasing off-task group behavior during large group settings when the intervention was in place. This behavior was high during both baseline and withdrawal conditions and significantly decreased during both intervention conditions. Results from the second study, conducted in the small group setting, were also able to illustrate the effectiveness of CW-FIT in decreasing off-task group behavior. Off-task behavior was high during both the baseline and withdrawal conditions in this

experiment. This behavior decreased significantly between the baseline and first intervention conditions as well as between the withdrawal and second intervention conditions.

The first experiment was, however, unable to replicate the findings from previous research (e.g., Jolstead et al., 2017; Kamps et al., 2011) that showed that CW-FIT was effective in improving teacher behaviors. Rates of both praise and reprimand were also highly variable, especially for reprimands during the first experiment. Initially, praise was low in baseline and increased significantly during the first intervention condition. Rates of praise, however, did not return to baseline levels during the withdrawal condition and were nearly double the baseline rate. Additionally, rates of praise decreased from the withdrawal condition to the second intervention condition, showing a contra-therapeutic change. Rates of reprimands showed a more defined therapeutic change and were high during both baseline and withdrawal conditions. This rate also decreased during the first intervention condition. However, although reprimands did decrease from the withdrawal condition to the second intervention condition, this change was minimal. Data on reprimands was also very variable with two out of five data points being higher than the average for the withdrawal condition. Additionally, rates of praise never met the goal (1.00 praise statements per minute or higher) for individual sessions and rates of reprimands only met the goal (0.20 reprimand statements per minute or lower) once in the large group setting.

Although slightly variable, the second experiment was able to replicate previous findings (e.g., Jolstead et al., 2017; Kamps et al., 2011) that CW-FIT was effective in improving teacher rates of praise and reprimand statements, unlike the first experiment. Baseline data for praise was extremely low and increased significantly during the first intervention condition. Rates of praise decreased during the withdrawal condition (although not entirely equal to baseline levels) and then again increased during the second intervention condition. This therapeutic change from the

withdrawal condition to the second intervention condition was not as significant of an increase compared to the difference between the baseline and first intervention conditions for rates of praise. Rates of reprimands were also significantly high during the baseline condition and decreased during the first intervention condition. Similarly, reprimands were also high in the withdrawal condition and decreased during the second intervention condition. Also note that both averages (for praise) during the two intervention conditions met the goal for praise (1.00 per minute or higher), although individual sessions were more variable with only two of five sessions meeting this goal. Neither averages for the rates of reprimands met the goal of 0.20 per minute (or lower) during the intervention condition; however, two sessions in the second intervention condition did meet this goal.

Despite, positive results from the study there were, however, some limitations to its design and effectiveness. First, the first author was only able to include one classroom in the current study. Combined with the four classrooms used in the Jolstead et al. (2017) study, researchers have only empirically studied the CW-FIT intervention in a total of five preschool classrooms.

Second, the intervention may have been harder to implement because of the short one-minute intervals used for points allocation. The teacher indicated that these short intervals made it more difficult to teach whilst also being responsible for implementing the intervention. Additionally, the contingency to receive rewards was never able to progress past a requirement of two points. This means that groups of students only needed to be on-task for a minimum of two one-minute intervals out of a ten-minute lesson to receive a reward. This limitation may be due to the relatively fast duration of the entire study (two and a half months). Had there been more time to continue the study, it may have been possible to increase the intervals to a more

reasonable length and require students to be on-task more than a minimum of 20% of the lesson to receive a reward. Additionally, had the paraprofessional been more willing to take on more responsibility for the implementation, it may have relieved some of the stress on the teacher.

Third, as was discussed before, treatment fidelity was only moderately high at 74.70% in the first experiment and 78.57% in the second experiment. Again, this less than ideal level was mainly due to the teacher and paraprofessional's inability to maintain the goal level of praise and reprimand statements. Even though off-task behavior continued to improve with the intervention, the low fidelity that the teachers implemented the intervention with may have thwarted the intervention from reaching even higher levels of success.

Additionally, the criteria for meeting the praise and reprimand goals was somewhat arbitrary. The goal for the number of praises was one praise statement per minute. The first author had chosen this goal based on a paper by LaBrot et al. (2016) which had identified an expected rate of one praise statement per minute for teachers. In hindsight, the first author recognizes this goal may have been too challenging for the teacher in the current study to be able to achieve. Past researchers studying CW-FIT have identified increasing praise and decreasing reprimands to be integral parts of the implementation of CW-FIT and have included them as components of their treatment fidelity. However, these researchers have also failed to include in their reports the levels at which they considered praise and reprimanding to be adequate and in compliance with CW-FIT procedures (Calderella et al., 2015; Jolstead et al., 2017; Kamps et al., 2011; Weeden et al., 2016). In future research, the first author plans to find a more socially valid measure of whether teachers are providing sufficient praise to their students.

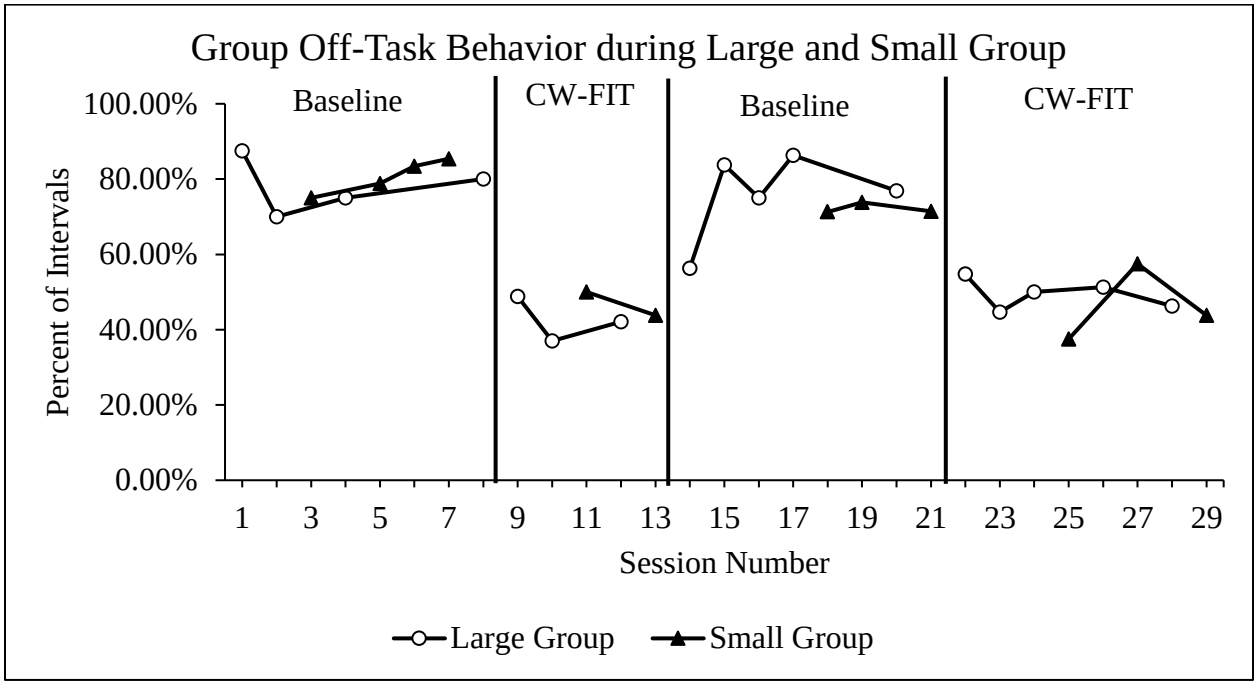
Last, another weakness from the design point of the study was that there were only two sessions in the first intervention condition in the second experiment. Unfortunately, this

oversight was due to the timeline of the study when it was interrupted by the Thanksgiving Break, which gave the students and teachers an entire week away from school. The break might have severely impacted the implementation of the study if the teacher had reinstated the intervention for only one day following a week-long break. Because of this reason, the first author decided to use the break as a natural removal of the intervention and indicator for when the withdrawal condition should begin. Nevertheless, having only two data points in this condition limit the power of the results obtained from the second experiment.

Despite the limitations discussed, overall, the implementation of CW-FIT was correlated with a decrease in off-task behavior in both large and small group settings. Trends and levels of off-task behavior were very similar for each condition across the two experiments and showed similar therapeutic changes between the baseline and intervention conditions consistent with previous research on the implementation of CW-FIT (e.g., Jolstead et al., 2017; Wills et al., 2010). CW-FIT implementation during the first experiment, however, did not show evidence to support previous findings (e.g., Jolstead et al., 2017; Kamps et al., 2011) that CW-FIT improved teacher's praise and reprimand statements. On the other hand, the first author was able to replicate this finding during the second experiment which found that CW-FIT increased rates of praise and decreased rates of reprimands.

Furthermore, CW-FIT was originally studied only in grades kindergarten through fifth (e.g., Wills et al., 2010; Kamps et al., 2011). Hence, the results of the current study and Jolstead et al. (2017) help to extend the capacity of CW-FIT as an intervention and provide evidence that CW-FIT is also effective in improving both student and teacher behavior in preschool classrooms.

Figure 5


Group Off-Task Behavior during Large and Small Group
Baseline
CW-FIT
Baseline
CW-FIT
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
Percent of Intervals
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
Session Number
Large Group
Small Group

References

Caldarella, P., Williams, L., Hansen, 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. *National Association for Music Education, 35,* 23-30. DOI: 10.1177/8755123315626229

Chow, J. C., & Gilmour, A. F. (2015). Designing and implementing group contingencies in the classroom. *Teaching Exceptional Children, 48,* 137-143. DOI: 10.1177/0040059915618197

Conklin, C. G, Kamps, D., & Wills, H. (2017). 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, New Jersey: Pearson Education, Inc.

Gast, D. L., & Spriggs, A. D. (2014). Visual analysis of graphic behavior. In D. L. Gast & J. R. Ledford (Eds.), *Single Case Research Methodology: Applications in Special Education and Behavior Sciences* (pp. 176-210). New York, NY: Routledge.

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

Himle, M. B., Miltenberger, R. G., Gatheridge, B. J., & Flessner, C. A. (2004). An evaluation of two procedures for training skills to prevent gun play in children. *Pediatrics, 113*, 70-77.

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

LaBrot, Z. C., Pasqua, J. L., Dufrene, B. A., Brewer, E. A., & Goff, B. (2016). In situ training for increasing head start after-care teachers' use of praise. *Journal of Behavioral Education, 25*, 32-48. DOI: 10.1007/s10864-015-9233-0

Schmidt Naylor, A., Kamps, D., & Wills, H. (2018). The effects of the CW-FIT group contingency on class-wide and individual behavior in an urban first grade classroom. *Education and Treatment of Children, 41*.

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. http://eds.b.ebscohost.com.proxy-remote.galib.uga.edu/eds/pdfviewer/pdfviewer?vid=1&sid=1e5fd1f4-813a-4ce2-aa7d-4ac9f784a1aa%40pdc-v-sessmgr01

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. http://eds.b.ebscohost.com.proxy-remote.galib.uga.edu/eds/pdfviewer/pdfviewer?vid=1&sid=8df95bad-467e-4be2-9d7e-1c59931847d2%40sessionmgr103

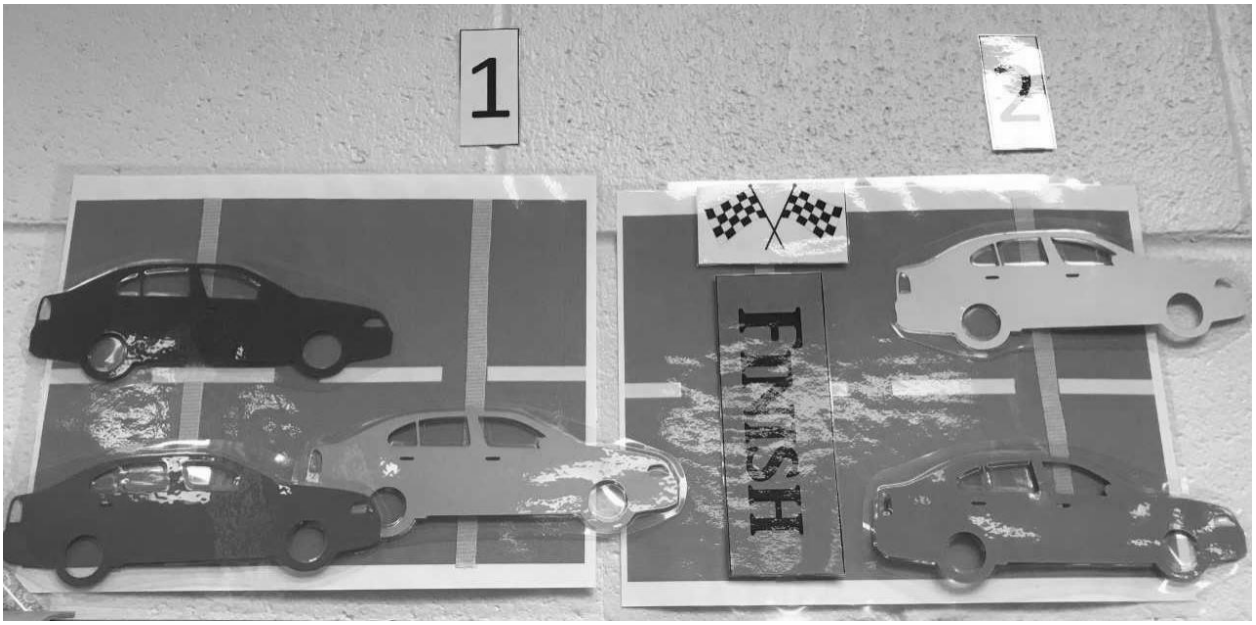
Wills, H. P., Kamps, D., Fleming, K., & Hansen, B. (2016). Student and teacher outcomes of the Class-Wide Function-Related Intervention Team efficacy trial. *Exceptional Children, 83,* 58-76. DOI: 10.1177/0014402916658658

Wills, H. P., Kamps, D., Hansen, B., Conklin, C., Bellinger, S., Neaderhiser, J., & Nsubuga, B, (2010). The Classwide Function-Based Intervention Team program. *Preventing School Failure, 54,* 164-171. DOI: 10.1080/10459880903496230

Wills, H. P., Kamps, D., Wehby, J. H., & Caldarella, P. (2016). *Improving elementary classroom and individual student outcomes with the CW-FIT program.* [PDF document]. Retrieved from: https://cwfit.ku.edu/sites/cwfit.ku.edu/files/docs/APBS%20CW-FIT%202016_Wills%20etal.pdf

Appendix A

Number Line



Appendix B

Procedural Fidelity

Before the game starts	
Skills posters, number line and cars are prominently displayed	Posters with each skill are posted in the room where students can see them. Number line is displayed so that the teams can see their progress.
Students split-up into 3-5 groups	Teams should be relatively even (same number of kids in each). Also try to split more disruptive students up between the groups.
BST of skills occurs at the beginning of the session	BST includes an explanation of the skill, model of the skill (adult or peer), whole class practice, and feedback. Altogether, reviewing the skills takes about 5 minutes.
Goal number of points was discussed with the students and indicated on the number line	Set an initial goal that they should be able to meet and slowly increase it. The point of the game is to give every group a chance to meet the goal and earn the reinforcer.
During Game:	
Timer used and reset for each appropriate interval	Have a timer to designate the end of the game, but have another that you set for 1-2 minutes (repeatedly through the game).
Points are recorded when the timer goes off (every 2 minutes)	Points are awarded at the end of the 1- or 2-minute timer. They should meet all the expectations throughout that time to earn a point.
Feedback is provided when points are awarded	Let teams who earned points know what they did well on and you can let teams who don't earn their points know which expectation they did

	not meet.
Teacher gave frequent descriptive praise	
Para-pro gave frequent descriptive praise	Provide praise even without the timer going off. Look for good things your students are doing throughout.
Teacher ignored problem behaviors (unless dangerous)	Ignoring inappropriate behavior refers to ignoring minor behaviors like call-outs (you can raise your hand quietly without looking at the student or call on a student with their hand raised and praise them), tapping on the desk, out of seat, and requests for clinic or bathroom (unless vomit is emanate). Major behaviors should be dealt with providing minimal attention (redirecting or blocking without talking to the student, providing only demands, or a quiet call for assistance).
Para-pro ignored problem behavior (unless dangerous)	
After Game:	
Immediately review points/number line at the end of each trial	At the end of the game or when all teams meet the goal, let the teams who met the goal know
Rewards are delivered immediately after the lesson	Immediately provide reinforcement to the teams that met the goal.
Take a picture of the number line or record group points in a spreadsheet	This is not vital to success, but it may help you trouble shoot if specific teams are consistently not meeting the goal.