Improving Student Behavior in Middle School Art Classrooms: Initial Investigation of CW-FIT Tier 1

Kimberly D. Monson, EdS[1], Paul Caldarella, PhD[1] iD**,
Darlene H. Anderson, PhD[1], and Howard P. Wills, PhD[2]**

Abstract
Classroom management is commonly challenging in middle schools. Class-wide function-related intervention teams (CW-FIT) is a multitiered intervention designed to decrease problem behaviors at the classroom level. It is comprises evidence-based practices such as teaching classroom expectations, increasing teacher praise, and using positive reinforcement in an interdependent group contingency. CW-FIT has shown promise in a variety of school settings, but it has not been tested in middle school art classrooms. This initial investigation examined the effects of CW-FIT using a single-subject ABAB design in two middle school art classrooms. Results indicated that class on-task behavior increased by more than 25% and teacher praise-to-reprimand ratios more than doubled during CW-FIT implementation compared with baseline levels. Results also indicated that on-task behavior for students identified as at risk for behavioral disorders improved by more than 18% during the intervention. Teachers and students found the intervention to be socially valid. Resulting implications were promising for using CW-FIT in other middle school art classrooms.

The transition from elementary to middle school includes multiple challenges that may cause students to lose motivation, disengage, and behave inappropriately (Symonds & Hargreaves, 2016). Transitioning students have to cope with starting a new school as they are contending with physiological, social, and emotional challenges, which can be difficult (Susman & Rogel, 2004; Young, Caldarella, Richardson, & Young, 2012). Unfortunately, relationships between middle school students and their teachers are often strained (Symonds & Hargreaves, 2016). Factors such as the ever-increasing class size in middle schools and the higher demands and expectations on students can cause students to feel an uncertainty of fit between themselves and their environment (Beaty-O'Ferrall, Green, & Hanna, 2010; Eccles, 2004). Many students become increasingly apathetic toward school and act out, thus leading to poor academic outcomes and classroom misbehavior (Young et al., 2012).

Challenges unique to middle school can lead to student disengagement and disruptive behavior. Most commonly, as Harrison, Vannest, Davis, and Reynolds (2012) discovered, middle school students are punished or sent to the principal's office for being disrespectful, defiant, and disruptive. As noted by Gottfredson, Gottfredson, and Hybl (1993) over 25 years ago, student apathy, loss of motivation, and

inappropriate behavior not only stunt student growth academically, socially, and emotionally, but can also be detrimental to the class as a whole. Despite the need to improve behavior management practices, many middle school teachers find it difficult to manage their classrooms with 65% reporting they receive little or no training to address behavioral issues (Reinke, Stormont, Herman, Puri, & Goel, 2011). In fact, student discipline is a significant reason why many teachers choose to leave the profession (Harrell, Leavell, van Tassel, & McKee, 2004).

Middle School Art Classrooms

Middle school art classes, although uniquely structured and different from core classes, can in many ways exacerbate students' problem behaviors common to middle schools

[1]Brigham Young University, Provo, UT, USA
[2]The University of Kansas, USA

Corresponding Author:
Paul Caldarella, Center for the Improvement of Teacher Education and Schooling, Brigham Young University, 149 MCKB, Provo, UT 84602, USA.
Email: paul_caldarella@byu.edu

Action Editor: Wendy Oakes

(Jason & Kuchay, 2001; Susi, 1995). Much of the literature surrounding behavior management in art classes was written 20 to 30 years ago and has not been updated; however, these findings still appear relevant today. Such studies suggest that art classrooms tend to be less restrictive, as they offer students a space to be creative, which can negatively affect student behavior (Larochelle, 1999). Students often view the art class as free (nonacademic) time, and they tend to ignore school rules and engage in off-task behavior (Larochelle, 1999). As with middle schools in general, art classes are becoming larger, leading many teachers to feel overwhelmed as they attempt to meet student needs (Kuster, Bain, Newton, & Milbrandt, 2010). An increase in students and a less structured environment can lead to higher instances of students behaving inappropriately.

Despite the difficulties art teachers face, they are expected to be both masters in their subject and masters in classroom management. According to Saunders (1989), secondary art teachers are expected to have competency in the art world as well as strong pedagogical skills. They must function as both teacher and manager over all aspects of the class and must be able to anticipate problem behavior before it happens. Such skills are essential, but unfortunately, many art teachers enter the field feeling less than prepared to manage a classroom effectively.

Kowalchuk (1999) found that difficulty with classroom management was one of the most commonly stated teacher problems, and many art teachers felt their previous education had not prepared them to handle such issues. In this study, only 6% of art teachers were successful in managing their classes, whereas 29% struggled to maintain control of the classroom. Other art teachers did not know how to determine whether a situation was going to escalate or whether they would be better off ignoring minor problem behaviors. Art teachers with high school and middle school-aged students struggled with their students' lack of motivation (Kuster et al., 2010). Overall art teachers tended to feel like they did not have the proper training to deal with problem behaviors and that lack of classroom management training should be addressed.

Ineffective Punishment Tactics

Many schools rely on punishment tactics such as suspension and expulsion to control student behavior, although such policies have been shown to be ineffective (Sugai & Horner, 2002). Zero-tolerance policies set to keep students safe by enforcing strict rules have led to a dramatic increase in students being expelled and suspended with patterns that disproportionality affect certain populations (e.g., students of color, students with disabilities; Skiba, 2014). The intent is that harsh punishments will curb future disruption by sending a message that specific behaviors will not be tolerated. However, excluding students from school is ineffective at

improving the school environment and may have lasting negative effects on academic and social growth (Skiba, 2014). The development of positive behavior interventions and supports (PBIS) was in part a reaction by parent and disability advocates against such zero-tolerance policies. As an alternative to these zero-tolerance policies, universal PBIS programs, using evidence-based practices, have had measurable positive outcomes on the behavior problems students and educators are facing (Simonsen, Sugai, & Negron, 2008). Class-wide function-related intervention teams (CW-FIT) is an example of one such program.

CW-FIT

CW-FIT is a multi-tier interdependent group contingency aligned with PBIS (Wills et al., 2010). This proactive approach is an attempt to mitigate negative student behavior without coercive consequences. According to the three-tiered continuum of support, 80% to 90% of students can be positively affected through Tier 1 interventions implemented with fidelity (Sugai & Horner, 2002; Young et al., 2012). Feasible to implement at the Tier 1 classroom level, CW-FIT has been found to increase teaching time by reducing the number of disruptions and problem behaviors that negatively affect classroom learning (Wills et al., 2010). Like other PBIS interventions, CW-FIT utilizes evidence-based practices such as teaching classroom expectations and positive behaviors followed by reinforcement of those skills and behaviors through praise and rewards (Sugai & Horner, 2002; Wills et al., 2010). Group contingencies and self-management strategies help to increase student engagement, prosocial skills, and on-task behavior (Kamps, Conklin, & Wills, 2015).

CW-FIT has two tiers with varying levels of intervention to accommodate and assist students who may not be responding adequately to a primary level of treatment (Wills et al., 2010). Tier 1 of CW-FIT focuses on teaching classroom expectations while extinguishing negative behaviors by reinforcing positive behaviors with praise and rewards implemented with a group contingency. Tier 2 of CW-FIT is structured to assist students who did not respond to Tier 1 interventions and continue with ongoing problematic behaviors. Students who are identified for Tier 2 through behavioral screening assessments are then taught self-management skills.

CW-FIT has shown positive results in many elementary school classrooms (see, for example, Wills, Kamps, et al., 2018; Wills, Wehby, Caldarella, Kamps, & Swinburne-Romine, 2018) but has not been widely tested in middle schools. Out of 15 published articles studying the effects of CW-FIT, 14 were based on elementary schools or early education settings, and only one was implemented in a middle school (Conklin, Kamps, & Wills, 2016). All of the studies found an increase in overall student on-task behavior and

teacher praise with CW-FIT. Nelson and colleagues (2018) sought to address the needs of art teachers who felt unprepared to manage student behavior, testing CW-FIT in one fifth-grade and two third-grade art classes. On-task behavior and teacher praise significantly improved during CW-FIT, and the intervention was found by the participating teacher and students to be socially valid.

Study Purpose

Due to the unique challenges in middle schools and art classrooms, as well as the promising results found in classrooms with similar needs, this initial investigation explored the effects of CW-FIT in two middle school art classes. The current study sought to expand on the results gathered by Nelson and colleagues (2018) and address some of their limitations. Nelson and colleagues tested CW-FIT only in an elementary school and considered only overall class-wide data. It is important that the field increases research examining the efficacy of first implementing Tier 1 before implementing Tier 2, particularly for students identified as at-risk.

The current study implemented CW-FIT Tier 1 in two middle school art classes, both to test the intervention with the entire class and to support students at risk for behavior disorders within those classes. This study was undertaken to answer the following questions: (a) Can middle school art teachers implement CW-FIT with fidelity? (b) Will CW-FIT increase overall student on-task behavior in middle school art classes? (c) Will CW-FIT specifically increase on-task behavior for students at risk for behavioral disorders? (d) Will implementing CW-FIT increase teachers' praise-to-reprimand ratio (PRR)? (e) Do middle school art teachers and students consider CW-FIT to be socially valid?

Method

Settings and Participants

This study was conducted in two seventh-grade art classes in a Title 1 middle school in a Mountain West state in the United States; at this school, 51% of students received free or reduced-price lunch. Two art teachers participated in this study. These classes were chosen to participate due to inappropriate behaviors and lack of motivation from the students. The teacher in Classroom 1 was a 29-year-old female with a bachelor's degree and 6 years of teaching experience. A 44-year-old male with a master's degree and 4 years of teaching experience taught Classroom 2. Both classes were taught during the 2017–2018 school year. Classroom 1 consisted of 30 students (average age = 12.30 years), and Classroom 2 included 26 students (average age = 12.23), for a total of 56. Of the total, 27 were male and 29 were female, with 35 Caucasian and 21 Hispanic students.

Both art classes were taught Monday through Friday for 45 min, except for "late start Mondays" and assembly days, when the class was shortened to 35 min. Both classes focused primarily on visual arts; during the study, they were working on ceramic and painting projects. Classes varied slightly in structure depending on the teachers' goals for the day. For example, some class periods consisted almost entirely of a lecture, whereas other class periods were focused entirely on starting or completing art projects.

Target students. Participating teachers were asked to nominate students at risk for behavioral disorders based on internalized and externalized behaviors using Stage 1 of the Systematic Screening for Behavior Disorders (Walker, Severson, & Feil, 2014). After students were nominated, teachers were asked to rank their nominations according to the degree of each student's exhibited behaviors. Parents of these students were sent consent forms to sign. Following the consent process, student behavior was assessed using the *School Social Behavior Scales*, Second Edition (SSBS-2; Merrell, 2002). The SSBS-2 assesses students based on different subscales including Self-Management, Social Competence, Relations with Peers, and Academic and Antisocial Behavior. Selection criteria required that students be ranked as at risk in at least one of the SSBS-2 scales. After this assessment, researchers observed the nominated students during baseline, with selection criteria showing an average on-task score of 70% or lower. None of the nominated students was receiving special education or resource accommodations at the time of the study. Initially, four students were nominated across the two classrooms, but two of them were not included as target students because of not being ranked as at risk on any subscales of the SSBS-2 or having an average on-task percentage higher than 70% during baseline. The remaining two students, one in Classroom 1 and one in Classroom 2, met selection criteria. Both were in the seventh grade: one was a 12-year-old Caucasian male and the other a 14-year-old Hispanic male.

Procedures

Consent. Both teachers signed consent forms to participate. Modified consent forms were mailed to all the students' parents in Classrooms 1 and 2, as approved by the participating university and school district institutional review boards. Consent forms were also mailed to the parents of the target students, who signed and returned them to the researchers. Target students also provided signed assent.

Baseline. Implementing CW-FIT into the classrooms involved five data points collected in Classroom 1 and six collected in Classroom 2. Data represented group and target student on-task behavior, as well as teacher praise and reprimands. During baseline, teachers conducted their class as

they normally would. Expectations in Classroom 1 were (a) use materials properly, (b) be positive and polite, (c) keep hands to yourself, and (d) use time wisely. Students had assigned seating and were allowed to leave their seats only to get supplies, wash paint brushes, or use the lighted tracing tables at the back of the class. The teacher in Classroom 1 responded to disruptions by pulling students aside to talk with them privately to correct behavior. She then called parents for concerns about behavior and if needed sent students to the school office. As a reward, if students were meeting her expectations, she would allow them to sit wherever they chose. In Classroom 2, the teacher's expectations were to (a) produce quality work, (b) be responsible, and (c) take care of the materials. He had a positive and negative consequence system that he would go over with students at the beginning of the semester and refer to afterward as needed. Students were allowed to sit by and talk with friends, as well as listen to music, as positive consequences for meeting class expectations. Students received negative consequences when they were disruptive or disrespectful to the teacher or classmates. These consequences included being sent to their assigned seat and not being allowed to talk to friends; after a student had received three warnings, his parents were contacted. During baseline, no changes were made to the routines, instruction, or classroom management style of either teacher.

Training. After baseline, in October, teachers were trained in a 45-min session on CW-FIT procedures. All components of CW-FIT were reviewed: (a) teaching classroom expectations, (b) providing behavior-specific praise, (c) awarding points to teams in response to the timer and providing group rewards, and (d) understanding the rationale behind each aspect. Training occurred with explanations of the major intervention components with detailed examples. The teachers were able to practice components and ask questions to increase their understanding. The training was led by researchers using a PowerPoint presentation and videos showing CW-FIT in a classroom. After the first training session, researchers provided the teachers with feedback on CW-FIT as implementation began. Teachers were given 2 days to become familiar with CW-FIT, and researchers remained available to consult with them as needed. During the initial 2 days of implementation, researchers monitored teachers to make sure they were implementing CW-FIT with fidelity and consulted with teachers regarding rewards that would work best in their classrooms. Fidelity was monitored using a fidelity checklist that observers completed concurrently with on-task data observations. Throughout the study, researchers provided brief 1 to 2 min consultations and corrections on aspects of CW-FIT, such as reminding teachers to do pre-corrections before class, giving behavior-specific praise, placing expectation posters where the whole class could see them, and using the timer at appropriate intervals.

Intervention. The independent variable, Tier 1 of CW-FIT, was then implemented in the classrooms: teaching classroom expectations, using an interdependent group contingency, giving praise, and awarding points redeemable for a reward.

Classroom expectations. The teachers chose which classroom expectations their class needed to learn or strengthen. The teacher of Classroom 1 chose *listen* and *work*. For Classroom 2, the teacher focused on *showing respect* and *staying on task*. Using a script to guide them, both teachers taught one expectation each of the first 2 days of intervention. The script required that students participate in the lessons. After the teacher had described the target expectation, the students were organized into groups to discuss which skills would be necessary to meet that expectation. After 2 min, the students were asked to share what they had discussed. For example, for the expectation *listen*, the students responded that a skill to meet the expectation could be keeping "eyes [on] the teacher, book, or screen." Students were then asked to discuss in their groups why they believed the expectation was important. After 2 min, students gave their rationale for the expectation, such as listening helps students "know what to do." Posters of each expectation taught were created and displayed in a prominent place in the classroom easily visible to all students. Both teachers gave their classes 2- to 3-min pre-corrects as a reminder of the expectations before each subsequent class.

Teams. Both teachers organized students into six teams comprising approximately three to six students per team. Teams were organized by tables at which students were sitting. Target students were not placed on their own team. Depending on the art project, students could leave their desks to obtain supplies or use lighted tables for tracing, but they remained on the same team. Most students remained in their assigned seats during art instruction.

Timer. Both teachers set their timers at 5-min intervals in line with the 35- to 45-min class timeslot and personal preference. The timers had both a beeping and a vibrating function, but teachers felt the students would display desired behavior only when they heard the audible beep to get points. For this reason, they switched the function to vibrate, as has sometimes been done in past studies of CW-FIT (see, for example, Nelson et al., 2018). Teachers awarded points after every 5 min when they felt the timer vibrates.

Points, praise, and goals. A point goal was set by the teachers at the beginning of each class, and teams reaching the point goal were given a reward. Daily point goals were based on 75% to 85% of the number of times the timer would vibrate during the class period (e.g., for nine timer signals, a point goal of 6 or 7). The teachers used a chart

to track each team's points, placing it in a prominent area where students could see it. Teachers were trained to scan every group at each timer vibration; each group with all members on task received a point. If at least one person in a group was off task, no point was given. Teachers could also award bonus points to groups or individuals who were meeting expectations between timer intervals to help those who were behind to catch up.

Along with awarding points to teams meeting expectations, both teachers were trained to provide behavior-specific praise. At the timer vibrations, teachers would praise groups and individuals while awarding points: for example, "Great job Team 1 for keeping your eyes on me while I'm teaching; you earned a point!" If a team did not earn a point, the teacher was to remind the students of the expectations and encourage them to work for the next point or try to earn bonus points: for example, "Team 5, you were not working, so I cannot give you a point. Get to work so you will be able to earn the next one."

Reward. At the end of the class period, the teachers tallied up the points for each team and provided rewards to teams reaching the point goal. Rewards were chosen based on student input and teacher practicality. During the first 2 days of CW-FIT, both teachers asked their students to suggest rewards that would be desirable incentives to help them stay on-task. Classroom 1 asked to receive tangible rewards: for example, lip balm, flexible pencils, mini flashlights, erasers, and mini pencil sharpeners. Classroom 2, meeting during the final period of the school day, decided that leaving class 2 to 3 min early should be their reward. Teams that did not meet the point goal were not given a reward. Target students were not given extra individualized rewards, but were rewarded along with their teams.

Dependent Variables and Measures

Researchers chose dependent variables and measures based on previous CW-FIT studies so results could be compared (Conklin et al., 2016; Nelson et al., 2018; Wills et al., 2010). Group on-task behavior data were collected first, followed by the collection of individual target student on-task behavior data.

Group on-task behavior. Researchers used a momentary time sampling method to record group on-task behavior in each classroom. Trained undergraduate and graduate researchers observed each class using 30 s intervals during 20-min observation sessions. Using a digital stopwatch, observers would visually scan each team every 30 s to look for on- and off-task behaviors. On-task behaviors included students working on art projects, watching the teacher demonstrate how to use different supplies, and listening to the lecture. Students were off task when they were disengaging, getting

supplies without permission, or talking while the teacher was talking. As CW-FIT uses an interdependent group contingency, a team was on task only if every team member was on task. Using paper and pencil, observers would mark groups as on task with a plus sign and as off task with a minus sign. Once the 20-min observation was completed, observers calculated the overall class on-task percentage.

Target student on-task behavior. Target student data were collected similar to overall class data, with a few differences. After all the teams were scanned, each target student was individually observed for on- and off-task behavior. Immediately following the group on-task observation, researchers recorded individual on-task percentages for both target students.

Teacher praise and reprimand. Observers recorded the frequency of teachers providing praise and reprimands during the 20-min observation sessions. Tally marks were recorded using paper and pencil each time the teachers praised or reprimanded groups or individuals. Praise was defined as a verbal statement of approval: "Everyone is waiting patiently for permission to get supplies, nice job!" Reprimands were defined as verbal statements in which the teacher told individuals or groups to stop an undesired behavior: "Team 5, you need to stop goofing off and get back to work."

Treatment fidelity. A nine-item checklist was completed at the end of each 20-min observation to determine whether the teachers were implementing CW-FIT with fidelity. The checklist enumerated CW-FIT procedures—such as "Classroom expectations clearly posted" and "Points awarded to teams for the use of skills"—and a separate column indicated whether these items were observed.

Social validity. At the end of the study, a researcher asked the teachers to complete a seven-item questionnaire assessing the social validity of CW-FIT. On five questions the teachers ranked their thoughts and experiences with CW-FIT using a 4-point Likert-type scale, ranging from 1 = *very true* to 4 = *not true*. These questions dealt with the practicality of CW-FIT and the amount of support the teachers received during the study. Two open-ended questions were included regarding what the teachers would recommend modifying about CW-FIT, and what might have helped them implement the intervention more easily. A questionnaire was also completed by students in both classrooms, consisting of four questions regarding what the students liked and disliked about CW-FIT and why; the target students received the same questionnaire as the others.

Interobserver Agreement

Interobserver agreement data were collected for 38.09% of the observations in Classroom 1 and 25.00% of the

observations in Classroom 2. Before beginning the study, researchers trained observers (three graduates and one undergraduate) to record class members' on-task, target students on-task, teacher praise, and teacher implementation fidelity. Before assisting in the study, observers memorized definitions and practiced gathering data by watching a video of CW-FIT occurring in a classroom. Accuracy was tested by comparing the observers' data with a master code file; 90.00% reliability was found across three sessions. Observers then achieved 90.00% reliability compared with the research coordinator by observing a nonstudy classroom across three sessions. Interobserver agreement for observational data averaged 96.63% for class on-task behavior, 98.59% for target student on-task behavior, 96.83% for teacher praise statements, and 92.56% for teacher reprimand statements.

Design and Analysis

A single case reversal design (ABAB; Cooper, Heron, & Heward, 2006) was used in this study. Observers in Classroom 1 collected five data points for baseline, two for training, five for the first intervention and reversal, and six for the final intervention. All phase changes were based on stability of group on-task data, because this was the primary dependent variable of interest. Observers in Classroom 2 collected six data points for baseline, two for training, and five for the first intervention, reversal, and final intervention. Changes in level, trend, and variability of group on-task, target student on-task, and teacher praise and reprimand were analyzed visually. Percentages of overall on-task data were calculated by totaling each instance of on-task behavior and then dividing that sum with the total possible number of intervals. To determine whether teachers were implementing CW-FIT with fidelity over the course of intervention, researchers averaged the percentages from teacher fidelity forms gathered after each observation. Praise to reprimand ratios were calculated by first calculating the average number of praise statements and the average number of reprimands across for each observation session. Ratios were then created comparing the average praise statements divided by the average number of reprimands across study phases. Researchers also conducted *Tau-U* analyses (Parker, Vannest, Davis, & Sauber, 2010) to compare the averages between phases using an online calculator (Vannest, Parker, Gonen, & Adiguzel, 2016). Baseline data were compared with intervention data and corrected when the trend of baseline data was increasing to the point it could have confounded the researcher's interpretation of the effectiveness of the intervention (Bruni et al., 2017). Researchers did not correct for baseline instances of clear differences between baseline and intervention data points.

Results

Treatment Fidelity

The teacher in Classroom 1 implemented CW-FIT Tier 1 with an average fidelity of 97.02% (*SD* = 3.34) during intervention phases. Similarly, during intervention the teacher in Classroom 2 implemented CW-FIT Tier 1 with an average of 95.89% (*SD* = 3.67). During baseline and reversal, the fidelity averages were 1.14% (*SD* = 0.57) in Classroom 1 and 5.55% (*SD* = 7.85) in Classroom 2.
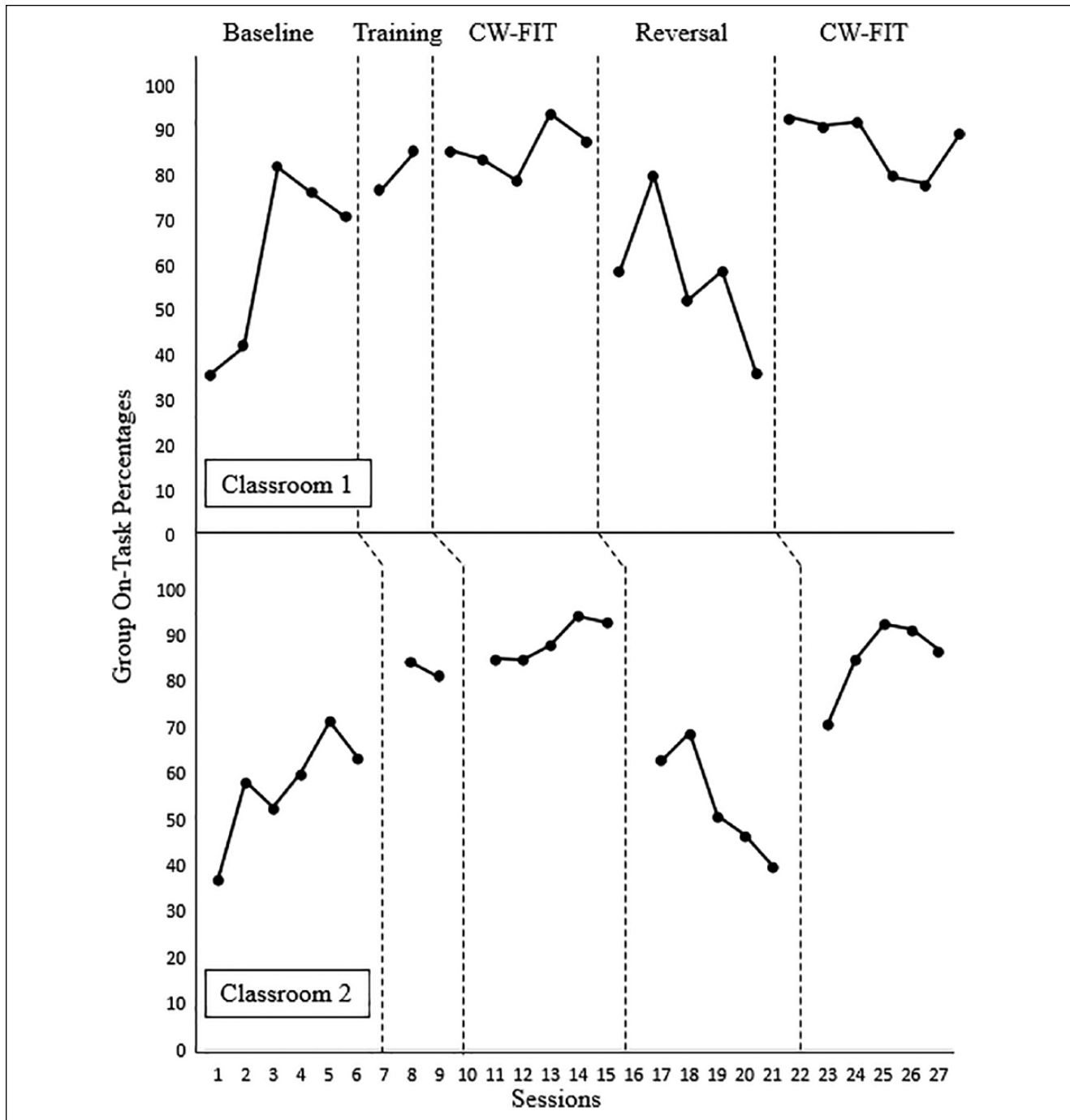
Group On-Task Behavior

During baseline, Classroom 1 had a group on-task average of 61.35% (*SD* = 20.94), with an upward trend and high variability, as shown in Figure 1. During training, the level of on-task behavior immediately increased to 86.07% (*SD* = 5.35), with one overlapping data point. During CW-FIT, the average was 88.94% (*SD* = 4.41), with a moderate upward trend and high variability, and one overlapping data point. During reversal, the level decreased to 57.25% (*SD* = 15.70), with a downward trend and moderate variability. During the final intervention, the average level of on-task behavior increased to 87.34% (*SD* = 6.48), with a downward trend and moderate variability and with two overlapping data points. Significant differences were found between baseline and intervention for this classroom (*Tau-U* = 0.90; *p* < .01).

In Classroom 2, baseline group on-task percentages averaged at 57.06% (*SD* = 11.66), with an upward trend and low variability. Similar to Classroom 1, on-task behavior increased immediately during training, averaging 80.95% (*SD* = 6.15). During CW-FIT, the average level increased slightly to 82.50% (*SD* = 2.12), with a high upward trend and moderate variability. There were no overlapping data points between baseline and intervention. During reversal, the average decreased to 53.81% (*SD* = 11.80), with a downward trend and moderate variability. During the final intervention phase, the average level of on-task behavior increased to 85.16% (*SD* = 8.66), with a slight upward trend and low variability. As with Classroom 1, significant differences were found between baseline and intervention (*Tau-U* = 0.98; *p* < .01).

Target Student On-Task Behavior

Data from each target student are discussed individually below (see Figure 2). During baseline, Target Student 1 had an on-task average of 70.77% (*SD* = 19.29), with an upward trend and low variability. During training, his average level of on-task behavior increased to 90.00% (*SD* = 0) with a stable trend. When CW-FIT was introduced, he decreased to an average of 73.50% (*SD* = 15.37), with an upward
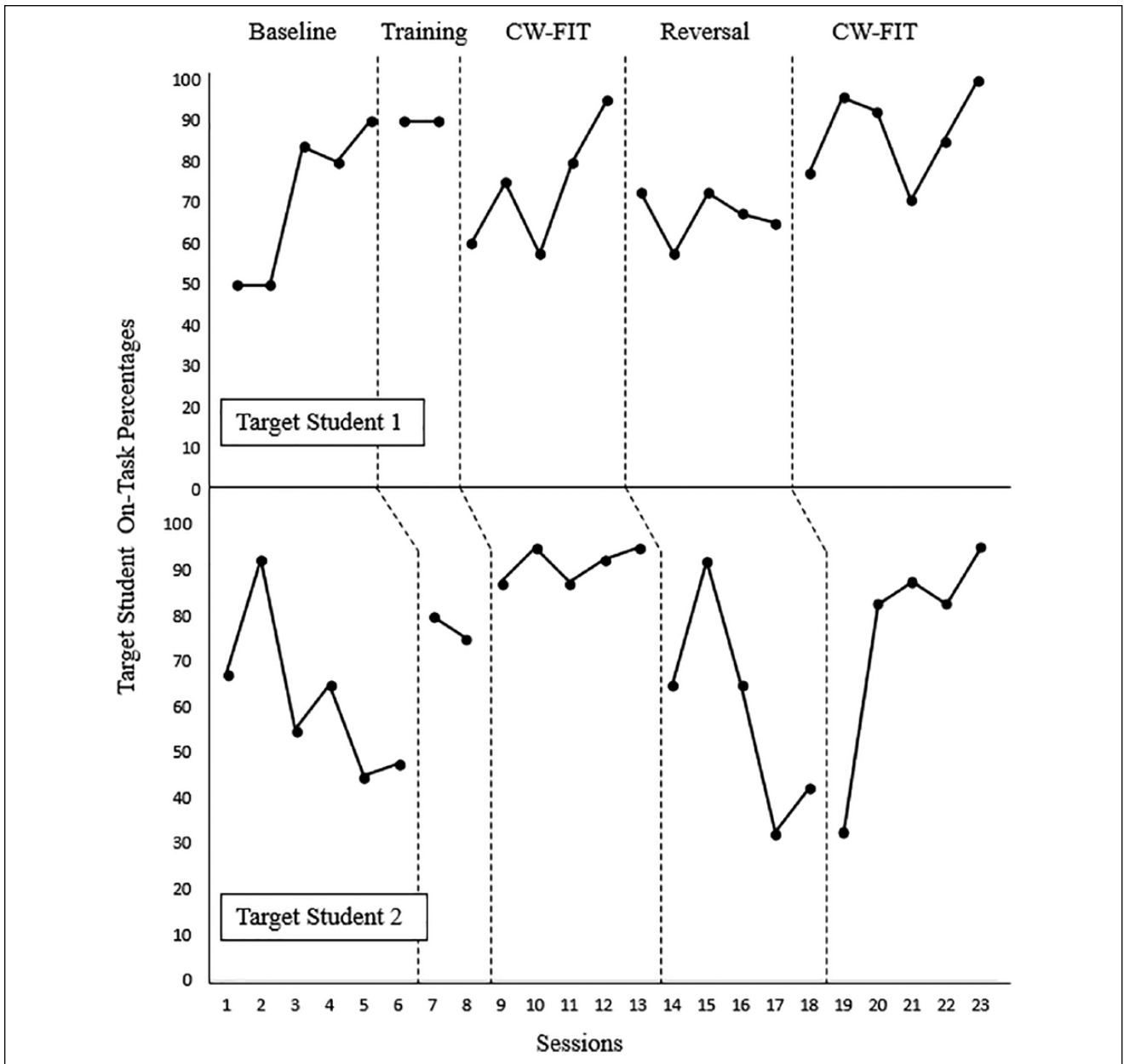
Figure 1. Average group on-task behavior percentages across phases.
Note. CW-FIT = class-wide function-related intervention teams.

trend and low variability. During reversal, his average decreased further to 67.00% (*SD* = 6.22), with a stable trend and low variability. When CW-FIT was reintroduced, the average level of on-task behavior for Target Student 1 increased to 86.88% (*SD* = 11.34), with a slight upward

trend and moderate variability. No significant changes were found between baseline and intervention for Target Student 1 (*Tau-U* = −0.08; *p* = .75).

Target Student 2 had an average on-task behavior of 62.08% (*SD* = 17.42) during baseline, with a downward

Figure 2. Average target student on-task behavior percentages across phases.
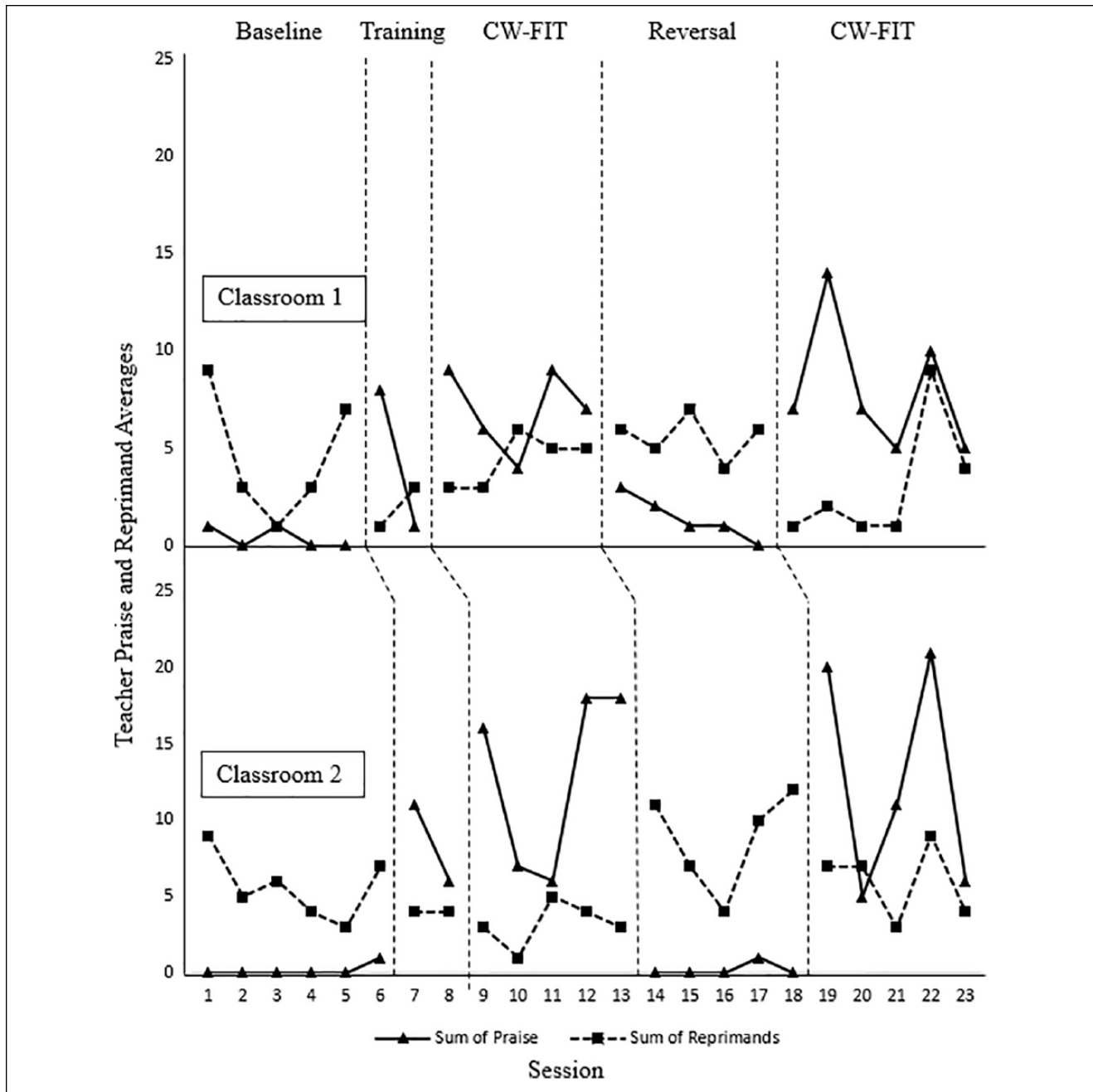*Note.* CW-FIT = class-wide function-related intervention teams.

trend and moderate variability. During training, he immediately increased to an average on-task behavior level of 77.50% (*SD* = 3.53). When CW-FIT was introduced, his average continued increasing to reach 91.50 (*SD* = 3.79), with an upward trend and moderate variability. During reversal, his average decreased to 59.40% (*SD* = 23.10), with a downward trend and high variability; however, when CW-FIT was reintroduced, the average level of on-task behavior increased to 76.00% (*SD* = 24.85), with an upward trend and moderate variability. A significant difference was found between overall baseline and intervention

for Target Student 2 when correcting for baseline (*Tau-U* = 0.80; *p* < .01). Other significant changes were found for Target Student 2 from baseline to intervention (*Tau-U* = 0.83; *p* = .02) and from intervention to reversal (*Tau-U* = −0.84; *p* = .03).

Teacher Praise and Reprimand

During baseline, the teacher in Classroom 1 praised the students an average of 0.04 times (*SD* = 0.55) and her rate of reprimands was 4.6 times (*SD* = 3.29) per 20-min

Figure 3. Average number of teacher praise versus reprimand in classrooms across phases.
Note. CW-FIT = class-wide function-related intervention teams.

observation session, as shown in Figure 3. The baseline PRR was 0.01:1, with a slight downward trend with low variability for praise, whereas reprimands showed a slight downward trend with high variability. During training, the praise rate in Classroom 1 increased to 4.50 (*SD* = 4.95) and the reprimand rate decreased to 2.00 (*SD* = 1.41) for a PRR of 2.25:1. Training resulted in an immediate increase in praise and an immediate decrease in reprimands. When

CW-FIT was introduced, the praise rate in Classroom 1 increased to 7.00 (*SD* = 2.12) with a reprimand rate of 4.40 (*SD* = 1.34), resulting in a PRR of 1.59:1. Praise showed a stable trend with moderate variability, whereas reprimand had a slight upward trend with low variability. During reversal, the praise rate in Classroom 1 fell to 1.40 (*SD* = 1.14), with a reprimand rate of 5.60 (*SD* = 1.14) and a PRR of 0.25:1. Praise during reversal had a downward trend with

low variability and reprimand had a slight downward trend with high variability. When CW-FIT was reintroduced, Classroom 1 had a praise rate of 8.00 (*SD* = 3.46) and a reprimand rate of 3.00 (*SD* = 3.16) for a PRR of 2.67:1. This showed praise with a downward trend and moderate variability, whereas reprimands showed an upward trend with moderate variability. No significant decreases were found in reprimands between baseline and intervention for Classroom 1 (*Tau-U* = −0.37; *p* = .15).

In Classroom 2, at baseline the teacher praised the students an average of 0.17 times (*SD* = 0.41) and reprimanded them an average of 5.67 times (*SD* = 2.16). The baseline PRR was 0.03:1, with a stable trend and very low variability for praise, and a slight downward trend with low variability for reprimand. During training, the praise rate in Classroom 2 increased to 8.50 (*SD* = 3.53) with a reprimand rate of 4.00 (*SD* = 0) for a PRR of 2.13:1. Training resulted in an immediate increase in praise and a slight decrease in reprimands. When CW-FIT was introduced, the praise rate in Classroom 2 increased to 13.00 (*SD* = 6) with a reprimand rate of 3.20 (*SD* = 1.48), resulting in a PRR of 4.06:1. During this phase, praise showed a moderate upward trend with moderate variability, whereas reprimands showed a slight upward trend with low variability. During reversal in Classroom 2, the praise rate was 0.20 (*SD* = 0.45) with a reprimand rate of 8.80 (*SD* = 3.27) for a ratio of .02:1. Praise during reversal showed a stable trend with very low variability, and reprimands had a slight upward trend with moderate variability. When CW-FIT was reintroduced, Classroom 2 had a praise rate of 12.60 (*SD* = 7.57) and a reprimand rate of 6.00 (*SD* = 2.45) for a PRR of 2.1:1. Praise during this phase showed a downward trend and moderate variability, and reprimand showed a stable trend and low variability. There were no significant decreases in reprimands between phases for Classroom 2 (*Tau-U* = −0.48; *p* = .06).

Social Validity

Teacher. The teacher in Classroom 1 answered *mostly true* for two questions: "The CW-FIT program was easy to learn and implement in my classroom" and "The training I received was adequate." She responded *very true* for two questions: "The Procedural Fidelity sheet was an effective teaching tool" and "In class support and feedback provided by the researcher was helpful." She answered *somewhat true* for "I will continue to use CW-FIT in my classroom." Expressing what was most helpful in learning how to implement CW-FIT, she responded that having a copy of the fidelity checklist was helpful, but that initial training packets could be condensed, and the first training only clarified a few points on how to implement the program. When asked how she would modify CW-FIT for the future, she responded that the timer could always be on a vibrate

function, students could have individual plans and goals rather than be in a group, and the prizes could be weekly instead of at the end of each day.

The teacher in Classroom 2 answered *very true* for three items: "The CW-FIT program was easy to learn and implement in my classroom," "The Procedural Fidelity sheet was an effective teaching tool," and "In class support and feedback provided by the researcher was helpful." He answered *mostly true* for "The training I received was adequate" and "I will continue to use CW-FIT in my classroom." He considered the daily feedback to be helpful in learning how to implement CW-FIT. Asked what he would modify about CW-FIT in the future, he responded that he was unsure and needed more time to process this question. Approximately 1 week after completing the survey, he was asked the same question by a researcher and answered that he did not have any changes to suggest at that time.

Students. A total of 48 students across the two classrooms (85.71% of all participating students) were also surveyed. Of these respondents, 95.83% responded that they enjoyed CW-FIT, whereas 4.17% indicated that they did not enjoy the intervention. Like the teachers, students were given open-ended questions as to what they liked and did not like about CW-FIT. The most common answers expressing what students liked about the intervention were "getting prizes/rewards" (*n* = 32), "it helps people work harder/stay on-task" (*n* = 5), and "it is a simple way to quiet our class down" (*n* = 3). Asked which aspects they did not like about CW-FIT, 21 (43.75%) students stated there were no aspects of CW-FIT that they disliked. However, others commented they disliked having other group members off task, so it was hard to earn points (*n* = 6) and being distracted by the timer every 5 min (*n* = 4); others merely stated that the intervention was "unfair" (*n* = 4). Some students also said they disliked losing points (*n* = 4), which is not an aspect of CW-FIT; one teacher had started taking away points as a negative reinforcer, without the knowledge of researchers; she was corrected when the researchers realized she was doing this. When asked whether CW-FIT would help other classes, 89.58% of students said *yes*, 8.33% said *no*, and 2.08% said *sometimes*. Common responses to the question of why CW-FIT would be helpful were "the prizes are motivating" (*n* = 26) and it helps keep students quiet and on-task (*n* = 5). Only four students stated that it would not be helpful in other classes, explaining that some students may not find the prizes very motivating or some students only pretend to be on task when the timer goes off.

Discussion

The purpose of this study was to determine the effectiveness of CW-FIT Tier 1 (Wills et al., 2010), a classroom management intervention based on PBIS principles, when

implemented in two middle school art classrooms. Classroom management is a frequent problem in middle schools, and art teachers feel underprepared to successfully manage classroom behaviors, especially when entering the field (Harrison et al., 2012; Kuster et al., 2010). Finding an effective classroom management tool can help teachers have more instruction time (Carter & Pool, 2012). CW-FIT has been shown to increase teaching time and improve class-wide behavior in middle school art classrooms, as demonstrated by the present study.

Results of the study indicated that middle school art teachers were able to implement the intervention with levels of fidelity consistent with fidelity found in other CW-FIT studies (Caldarella, Williams, Hansen, & Wills, 2015; Nelson et al., 2018; Wills, Iwaszuk, Kamps, & Shumate, 2014). The fidelity for most items ranged from 96.67% to 100%. Teacher praise increased in both classrooms, though changes in teacher reprimands were not as consistent. Overall, significant increases in teacher praise were found between all baseline and intervention phases in both classrooms. The teacher from Classroom 1 had a PRR of 0.14:1 during baseline and a ratio of 2.08:1 during intervention. Similarly, the Classroom 2 teacher had a PRR of 0.03:1 during baseline and a ratio of 2.58:1 during intervention. The results were consistent with CW-FIT elementary school studies in which PRRs increased significantly (Kamps, Wills, et al., 2015; Nelson et al., 2018; Wills et al., 2014). However, there were no significant differences in reprimands for either classroom, possibly due to the difficulties of managing middle school classrooms (Beaty-O'Ferrall et al., 2010; Eccles, 2004; Reinke et al., 2011) as reflected in such teachers often being more reactive to student misbehavior than are elementary school teachers (Reddy, Rhodes, & Mulhall, 2003).

Percentages of group on-task behavior improved with implementation of CW-FIT Tier 1, particularly for Classroom 2 in which there were two replications of effects and no overlapping data points between baseline and intervention. The average group on-task percentage during baseline in Classroom 1 was 59.30%, which increased to 84.79% during intervention—a 25.50% improvement for this classroom. Similarly, in Classroom 2, the average percentage of group on-task behavior during baseline was 55.44%, which increased to 85.53% during intervention, a 30.09% average improvement. These increases are consistent with previous CW-FIT elementary school studies (Kamps, Wills, et al., 2015; Kamps et al., 2011) and somewhat higher than previous results in elementary art classrooms (Nelson et al., 2018).

Although there were definite increases in on-task behavior for both target students, there were no statistically significant increases found for Target Student 1. Overall, the combined target student on-task percentage increased by 17.75%, which was lower than previous CW-FIT studies with target elementary students (Caldarella et al., 2015;

Conklin et al., 2016; Wills et al., 2014). The on-task percentage for Target Student 1 increased from an average of 68.89% during baseline to 83.46%, leading to a 14.57% increase, which was not statistically significant and was lower than changes in individual target students found in previous CW-FIT studies (Wills et al., 2014). This was likely due to higher on-task percentages during baseline and reversal for this student. Target Student 2 did have a significant increase in on-task behavior and went from 60.74% during baseline to 81.67% during intervention, leading to a 20.93% increase, which was statistically significant and similar to previous CW-FIT elementary school studies (Conklin et al., 2016; Wills et al., 2010).

Finally, participating teachers and students found CW-FIT Tier 1 to be helpful and socially valid. This is also consistent with previous studies in elementary school (Caldarella et al., 2015; Nelson et al., 2018; Wills et al., 2010). Both teachers stated that the intervention was relatively easy to implement in their classrooms and noticed a positive change in their students. Likewise, both teachers found the fidelity checklist to be an effective training tool. One teacher indicated that she did not like the beeping timer and felt that the initial training material could have been clearer with its expectations of where and how to implement the intervention. The vast majority of students indicated that they liked CW-FIT and thought that it would be beneficial in other classrooms. Most students enjoyed earning prizes and getting to leave class early, as well as benefiting from improved class behavior.

Limitations and Areas for Future Research

Although the results of this initial investigation were positive, there were some limitations. First, the study was conducted in only two art classrooms. To improve generalizability to other middle school art classes, the intervention should be replicated in more classrooms. Also, students were either Caucasian or Hispanic; the intervention should be studied in more diverse settings. Second, only two target students participated in the study. Significant changes were found for only one of them, which leads to concern about how target students were selected and whether modifications could be made to selection criteria. Replications should also include more target students to effectively examine the intervention's effect on student behavior.

Ascending on-task behavior baselines in both classes limit the confidence in a functional relationship between CW-FIT and on-task behavior. More baseline data could have been collected to determine whether on-task behavior more fully leveled out before starting the intervention. However, the last baseline data point was down before researchers began the first intervention. Future studies would benefit from implementing the intervention in classrooms with more stable baselines.

CW-FIT is a multitiered intervention, yet this study implemented only Tier 1. Some previous studies of CW-FIT have implemented both Tier 1 and Tier 2 (Caldarella et al., 2015; Kamps, Conklin, et al., 2015; Wills et al., 2010). Both teachers in the current study felt positive about their students' progress and did not feel like Tier 2 was necessary. We suggest that both Tiers of CW-FIT be investigated in middle school art classrooms according to student needs.

Finally, the teacher in Classroom 1 spontaneously began taking points away from teams as punishment for off-task behavior. In CW-FIT, each group maintains the points they have earned, even if they do not earn a point for the next time signal (Wills et al., 2010). Points were taken away approximately three times. The first time this happened, a member of the research team corrected the teacher after class. During the next data point, the teacher took points away again and was again corrected. She stopped taking points away, but four students noted on their social validity surveys that they did not like having points taken away during CW-FIT. It is recommended that future studies add a section on the fidelity checklist or training materials that specify and give rationale for the instruction that points should not be taken away.

Conclusion

This initial investigation has shown that CW-FIT Tier 1 can be effective as a classroom management tool for middle school art classroom teachers. Overall findings suggest that CW-FIT Tier 1 was effective at improving behaviors at class-wide and individual target student levels. This is significant, given the need for more research examining the efficacy of first implementing Tier 1 interventions before implementing Tier 2 for students identified as at-risk. As noted earlier, research into classroom management in middle school art classes is limited. Although further replications of this study need to be conducted to verify the results, CW-FIT Tier 1 shows promise for application in middle school art classrooms. Results from this investigation indicated that middle school art teachers can implement the intervention with fidelity, which leads to increased teacher praise and the improvement of group and target student on-task behavior. Teachers and students found the intervention easy to implement, enjoyed it, and considered it socially valid. These results suggest that CW-FIT Tier 1 could be a useful tool for classroom management in other middle school art classrooms.

Declaration of Conflicting Interests

Funding

ORCID iD

Paul Caldarella (iD) https://orcid.org/0000-0002-0883-8890

References

Beaty-O'Ferrall, M. E., Green, A., & Hanna, F. (2010). Classroom management strategies for difficult students: Promoting change through relationships. *Middle School Journal*, *41*, 4–11. doi:10.1080/00940771.2010.11461726

Bruni, T. P., Drevon, D., Hixson, M., Wyse, R., Corcoran, S., & Fursa, S. (2017). The effect of functional behavior assessment on school-based interventions: A meta-analysis of single-case research. *Psychology in the Schools*, *54*, 351–369. doi:10.1002/pits.22007

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

Carter, D. R., & Pool, J. L. (2012). Appropriate social behavior: Teaching expectations to young children. *Early Childhood Education Journal*, *40*, 315–321. doi:10.1007/s10643-012-0516-y

Conklin, C. G., Kamps, D., & Wills, H. (2016). 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. (2006). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Eccles, J. S. (2004). Schools, academic motivation, and stage-environment fit. In R. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 125–153). New York, NY: Wiley.

Gottfredson, D. C., Gottfredson, G. D., & Hybl, L. G. (1993). Managing adolescent behavior: A multiyear, multischool study. *American Educational Research Journal*, *30*, 179–215. doi:10.3102/00028312030001179

Harrell, P., Leavell, A., van Tassel, F., & McKee, K. (2004). No teacher left behind: Results of a five-year study of teacher attrition. *Action in Teacher Education*, *26*, 47–59.

Harrison, J. R., Vannest, K., Davis, J., & Reynolds, C. (2012). Common problem behaviors of children and adolescents in general education classrooms in the United States. *Journal of Emotional and Behavioral Disorders*, *20*, 55–64. doi:10.1177/1063426611421157

Jason, L. A., & Kuchay, D. A. (2001). Ecological influences on school children's classroom behavior. *Education*, *105*, 411–413.

Kamps, D., Conklin, C., & Wills, H. (2015). Use of self-management with the CW-FIT group contingency program.

Education and Treatment of Children, *38*, 1–32. doi:10.1353/etc.2015.0003

Kamps, D., Wills, H. P., 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

Kowalchuk, E. A. (1999). Perceptions of practice: What art student teachers say they learn and need to know. *Studies in Art Education*, *41*, 71–90. doi:10.2307/1320251

Kuster, D. A., Bain, C., Newton, C., & Milbrandt, M. K. (2010). Novice art teachers: Navigating through the first year. *Visual Arts Research*, *36*, 44–54. doi:10.1353/var.2010.0001

Larochelle, L. (1999). Classroom management in the art room. *Arts & Activities*, *126*, 62–63.

Merrell, K. W. (2002). *School social behavior scales* (2nd ed.). Eugene, OR: Assessment-Intervention Resources.

Nelson, M. A., Caldarella, P., Hansen, B. D., Graham, M. A., Williams, L., & Wills, H. P. (2018). Improving student behavior in art classrooms: An exploratory study of CW-FIT tier 1. *Journal of Positive Behavior Interventions*, *20*, 227–238. doi:1098300718762744

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2010). Combining non-overlap and trend for single case research: Tau-U. *Behavior Therapy*, *42*, 284–299. doi:10.1016/j.beth.2010.08.006

Reddy, R., Rhodes, J. E., & Mulhall, P. (2003). The influence of teacher support on student adjustment in the middle school years: A latent growth curve study. *Development and Psychopathology*, *15*, 119–138. doi:10.1017/s0954579403000075

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, *26*, 1–13. doi:10.1037/a0022714

Saunders, R. J. (1989). How to select an effective art teacher. *NASSP Bulletin*, *73*, 54–60.

Simonsen, B., Sugai, G., & Negron, M. (2008). Schoolwide positive behavior supports: Primary systems and practices. *Teaching Exceptional Children*, *40*, 32–40. doi:10.1177/004005990804000604

Skiba, R. J. (2014). The failure of zero tolerance. *Reclaiming Children and Youth*, *22*, 27–33.

Sugai, G., & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy*, *24*, 23–50. doi:10.1300/J019v24n01_03

Susi, F. D. (1995). *Teacher Resource Series. Student behavior in art classrooms: The dynamics of discipline*. Reston, VA: National Art Education Association.

Susman, E. J., & Rogel, A. (2004). Puberty and psychological development. In R. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 15–44). New York, NY: Wiley.

Symonds, J., & Hargreaves, L. (2016). Emotional and motivational engagement at school transition: A qualitative stage-environment fit study. *The Journal of Early Adolescence*, *36*, 54–85. doi:10.1177/0272431614556348

Vannest, K. J., Parker, R. I., Gonen, O., & Adiguzel, T. (2016). Single Case Research: Web based calculators for SCR analysis (Version 2.0) [Web-based application]. College Station: Texas A&M University. Available from http://www.single-caseresearch.org

Walker, H. M., Severson, H. H., & Feil, E. G. (2014). *Systematic screening for behavior disorders* (2nd ed.). Eugene, OR: Pacific Northwest.

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., Caldarella, P., Wehby, H. P., & Swinburne-Romine, R. S. (2018). Class-wide function-related intervention teams "CW-FIT": Outcomes from a multi-site randomized replication trial. *The Elementary Education Journal*, *119*, 29–51. doi:10.1086/698818

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. P., Wehby, J., Caldarella, P., Kamps, D., & Swinburne-Romine, R. (2018). Classroom management that works: A replication trial of the CW-FIT program. *Exceptional Children*, *84*, 437–456. doi:10.1177/0014402918771321

Young, E. J., Caldarella, P., Richardson, M. J., & Young, K. R. (2012). *Positive behavior support in secondary schools: A practical guide*. New York, NY: The Guilford Press.