

Managing Student Behavior in Middle School Using
Class-wide Function-Related Intervention Teams

Paul Caldarella

Brigham Young University

Howard P. Wills

University of Kansas

Darlene H. Anderson, Leslie Williams

Brigham Young University

Author Notes

Address all correspondence to Paul Caldarella, PhD, Center for the Improvement of Teacher Education and Schooling, Brigham Young University, 149 MCKB, Provo, UT 84602.

Telephone: 801-422-5081. Fax: 801-422-0199. Email: paul_caldarella@byu.edu

Acknowledgement

The research reported in this article was supported in part by a grant from the Institute of Education Sciences and the U.S. Department of Education (R324A160279) awarded to the University of Kansas in cooperation with Brigham Young University. The opinions presented in this article are those of the authors and no endorsement by the agency is intended or implied.

The citation for this paper in its' final published version is as follows:

Caldarella, P., Wills, H. P., Anderson, D. H., & Williams, L. (2019). Managing student behavior in the middle grades using Class-wide Function-Related Intervention Teams. *Research in Middle Level Education*, 42(7), 1-15. doi:10.1080/19404476.2019.1654799

Abstract

During the transition from elementary to middle school, adolescents may become increasingly vulnerable for emotional and behavioral problems. Class-Wide Function-Related Intervention Teams for middle school (CW-FIT MS), the independent variable examined in this study, was developed to improve teaching and learning by fostering student engagement. The study's purpose was to examine the feasibility and effects of CW-FIT MS Tier 1 implementation across nine middle school classes using a single-subject ABAB withdrawal design. Participants were 234 students in grades 6-8, including 23 identified as at risk for emotional behavioral disorders. Results of CW-FIT MS Tier 1 implementation showed significant improvement in on-task behavior for groups as well as individual target students, increased teacher praise-to-reprimand ratios, and decreased target student disruptions. Outcomes of social validity surveys were consistent with earlier studies. Study findings extend previous results by demonstrating positive intervention effects in novel settings with a greater number of participants. Study limitations and areas for future research are discussed.

Keywords: middle school, teacher praise, student engagement, group contingency

Managing Student Behavior in Middle School Using
Class-Wide Function-Related Intervention Teams

Student problem behavior continues to be a major concern for teachers (Bushaw & Lopez, 2010; Emmer & Stough, 2001; Harrison, Vannest, Davis, & Reynolds, 2012). The middle school years (Grades 6-8) mark the beginning of a particularly vulnerable time for developing problem behavior (Bernstein, 2002). Evidence-based practices and interventions are available to address inappropriate classroom behavior and improve academic engagement (Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008); however, teachers report inadequate training as a significant barrier to implementing research-validated curricula (Reinke, Stormont, Herman, Puri, & Goel, 2011). To bridge the gap between research and practice, experts recommend increasing the training, coaching, and support offered to teachers of students manifesting behavioral difficulties (Reinke et al., 2011).

The transition from elementary to middle or junior high school is significant and often stressful for students (Chung, Elias, & Schneider, 1998), involving difficult changes in social and academic contexts (Chung et al., 1998; McIntosh, Flannery, Sugai, Braun, & Cochrane, 2008) that can increase anxiety, distractibility, hyperactivity, and difficulty in following directions and completing assignments (Harrison et al., 2012). Disengagement (Sinclair, Christenson, Evelo, & Hurley, 1998) and accompanying declines in academic achievement are reported as additional concerns (Chung et al., 1998; Young, Caldarella, Richardson, & Young, 2012).

Managing student behavior tends to be more complex for teachers in middle school than for teachers in earlier grades (Chung et al., 1998). For example, disruptive behavior poses a significant problem to middle school teachers and students, placing such students at increased risk for reactive discipline practices such as reprimands and suspensions (Närhi, Kiiski, Peitso, &

Savolainen, 2015). Such reactive practices are potentially harmful to teaching and learning; in particular, harsh reprimands prove less effective than positive behavioral classroom management strategies (Reinke, Herman, & Stormont, 2013).

Fortunately, evidence-based approaches implemented with fidelity can improve teaching and maximize instruction time as well as foster student engagement and learning (Emmer & Stough, 2001; Muscott, Mann, & LeBrun, 2008). For example, Sinclair and colleagues (1998) identified the following teacher actions to improve school-related outcomes: (a) create supportive and caring learning environments, (b) provide opportunities for students to succeed, and (c) help learners understand connections between school achievement and future life events. McIntosh et al. (2008) further identified school engagement as a critical variable in secondary school transitions leading to positive long-term outcomes. As findings indicated that students with behavior issues were likely to experience problems with academics as well, these researchers recommended identifying risk factors (e.g., low grades and misbehavior prior to Grade 10) and providing students with academic and behavioral support (McIntosh et al., 2008). Re-examining the academics-behavior relationship in terms of the current emphasis on prevention in school-wide systems of support, Algozzine, Wang, and Violette (2011) advised practitioners to “[Teach] behavior as relentlessly as . . . reading or other academic content” (p. 13) to strengthen student performance.

Classroom management strategies designed to increase students' on-task behavior and decrease disruptive behavior rely largely on five teacher behaviors: (a) providing sufficient class structure, (b) establishing clear expectations, (c) interacting actively with students to increase engagement, (d) using multiple strategies to respond to appropriate behavior, and (e) using a continuum of strategies to deal with inappropriate behavior (Simonsen et al., 2008). Suggested

strategies also include group contingencies, differential reinforcement, functional assessment, and appropriate behavior cues (Stage & Quiroz, 1997).

Behavior specific praise also improves student engagement (Marchant & Anderson, 2012; Reinke et al., 2013). Of increasing interest for improving student behavior is teachers' praise-to-reprimand ratios (PRR), a comparison of the number of praise statements delivered to the number of reprimands (Caldarella, Larsen, Williams, Wills, & Wehby, in press). In an empirical investigation of the critical indicators of effective teaching (Espin & Yell, 1994), the teachers rated as most competent elicited high rates of student academic responding and delivered frequent positive responses (e.g., praise); they also gave infrequent negative responses (e.g., reprimands) in the presence of low rates of student off-task behavior. It is thus important that teachers receive training and support in the delivery of contingent, behavior specific praise and decreased reprimands as a means to strengthen desired student behavior (Simonsen et al., 2008).

Class-Wide Function-Related Intervention Teams

Class-Wide Function-Related Intervention Teams (CW-FIT), which incorporates the above evidence-based strategies, has been empirically evaluated in elementary school classrooms with positive results (Wills, Kamps, Fleming, & Hansen, 2016). As adapted for middle school, CW-FIT MS incorporates a teaching component with emphasis on helping students learn necessary behavioral skills. Intervention features include (a) directly teaching classroom expectations and prosocial skills, (b) using an interdependent group contingency with differential reinforcement of desired behavior, and (c) minimizing teacher attention for inappropriate behavior by planned ignoring. These specific components, considered part of a Tier 1 intervention or primary prevention effort, are implemented with the entire class (Wills et al.,

2016).

The CW-FIT MS interdependent group contingency involves (a) assigning students to teams based on seating or instructional arrangements, (b) using a class reward menu to support differential reinforcement of appropriate classroom behavior, and (c) providing students with behavior-specific praise to recognize and reward desired behavior. Teachers use a timer set at 5-minute intervals to prompt feedback and delivery of points. A typical intervention session proceeds as follows: (a) the teacher pre-corrects or prompts skills, (b) a daily point goal is set, (c) regular classroom instruction begins, (d) the teacher manages timer intervals, (e) feedback/points are delivered contingently to groups every 5 minutes for engaging in appropriate classroom behavior, (f) points are tallied at end of class period, and (g) teams that meet the daily point goal receive a group reward.

Initially teachers implementing the program receive training and in-class coaching, as well as ongoing support as needed. CW-FIT MS includes individualized Tier 2 interventions (e.g., self-management) for students who do not respond adequately to Tier 1 alone. The present study examined intervention effects achieved with the student participants at the Tier 1 or primary intervention level (Wills et al., 2016; Wills et al., 2010).

Previous Research

Previous studies of CW-FIT in elementary schools have consistently shown improved prosocial behavior and reduced disruptive behavior during instruction, as well as increased rates of teacher praise and decreased reprimands across age groups and subject areas (Caldarella, Williams, Hansen, & Wills, 2015; Caldarella, Williams, Jolstead, & Wills, 2017; Kamps et al., 2015; Kamps et al., 2011; Weeden, Wills, Kottwitz, & Kamps, 2016; Wills, Iwaszuk, Kamps, & Shumate, 2014; Wills et al., 2010). The intervention has also been found to yield high levels of

treatment integrity and to be perceived by teachers and students as being socially valid (Wills et al., 2016).

As part of the process of developing CW-FIT MS, a series of qualitative focus groups were conducted to invite middle school teachers, administrators, and students to provide feedback on CW-FIT MS components (see Wills, Caldarella, Mason, Lappin, & Anderson, in press). While the middle school and elementary versions were quite similar, several revisions were made.

1. Only two classroom expectations were taught instead of three, since teachers reported limited class time, with one of the lessons focused on showing respect in the classroom, since this concern was frequently identified during focus groups.
2. Classroom expectation lessons included more active student participation, in response to the need for greater middle school student involvement and buy-in.
3. Longer timer intervals were used (average 3-minute elementary school intervals increased to 5 minutes in middle school).
4. Teacher training was shortened and external coaches provided targeted fidelity feedback at the end of class periods, instead of in class demonstrations of intervention components (as had been done in elementary school classrooms).

This is the third study focused on the impact of CW-FIT in middle schools. Results of two prior studies using single subject designs appeared promising. Orr, Caldarella, Hansen, and Wills (2019) used an ABAC design to examine intervention effects on student on-task behavior and teacher praise rates in a self-contained special education classroom for students with severe disabilities and typically developing peer tutors. Results of this study suggested that implementation of CW-FIT was associated with improvements in student on-task behavior and

teacher praise rates, especially when peer tutors were included in the intervention and when the timer was silent. Both teachers and students reported the intervention to be socially valid.

Wills et al. (in press) examined the effects of CW-FIT MS Tier 1 in three middle school classrooms using a single-subject ABAB withdrawal design. Results showed a noticeable improvement in class-wide on-task behavior from a baseline average of 56% to an implementation average of 78%. Teacher praise rates also improved significantly, from a baseline mean of 1.0 to an intervention mean of 5.0 per 10-minute observation. Findings also indicated that six students identified as at risk for emotional and behavioral disorders (EBD) demonstrated improvements in their on-task behavior.

Study Purpose

The promising results achieved by Orr et al. (2019) and Wills et al. (in press) affecting the unique behavioral challenges occurring in middle schools affirmed the importance of evaluating the feasibility of implementation of intervention components in novel environments with a larger number of teachers and students, including students with or at risk for EBD. Prior to beginning the study, the researchers hypothesized that increased rates of student on-task behavior and teacher praise would be observed during CW-FIT MS Tier 1 implementation. They also theorized that teachers and students would approve of the intervention's goals, procedures, and outcomes when asked to complete social validity ratings. The following research questions were addressed: (a) Can middle school teachers implement CW-FIT MS with fidelity? (b) Will CW-FIT MS increase overall student on-task behavior in middle school classes? (c) Will CW-FIT MS increase at-risk students' on-task behavior and decrease their disruptive behavior? (d) Will implementing CW-FIT MS increase teachers' praise-to-reprimand ratios? (e) Do middle school teachers and students consider CW-FIT MS to be socially valid?

Method

Settings and Participants

This study was conducted with 234 students enrolled in nine middle school classes in Grades 6-8 across five Title I schools. The participating schools represented different geographic and socioeconomic areas (Table 1). Classrooms 1-5 were located in an urban Midwestern U.S. city, while Classrooms 6-9 were in two interurban communities in the Western U.S.

Teachers. Seven female and two male teachers participated, ranging in age from 24 to 44 years ($M = 32.4$; $SD = 6.8$). Six of the teachers had bachelor's degrees, two had master's degrees, and one had an educational specialist degree. Teaching experience ranged from 1 to 12 years ($M = 5.9$; $SD = 3.6$), five having taught in different schools and four in just one school. For each teacher all data collection took place in the class period she or he identified as having the most challenging student behavior. Subject areas taught in four of the nine classes were social studies, language arts, or a combination of social studies and language arts; three of the classes were math and two were science.

At-risk students. Each teacher nominated and ranked two to four students in the target class using Stage 1 of the *Systematic Screening for Behavior Disorders* (SSBD; Walker & Severson, 1992), which identifies students as at risk for externalizing (off-task, disruptive) or internalizing (anxious, withdrawn) problem behavior. The SSBD is a standardized, norm-based screening tool with a multiple gating system on which the teacher (a) evaluates all students in the class on internalizing and externalizing dimensions, (b) nominates and ranks the most severe cases, and (c) completes ratings of maladaptive and adaptive behavior. Students whose scores exceed normative criteria are then directly observed to confirm the presence of problem behaviors. The present study used Stage 1 of the SSBD, which consists of evaluating and ranking

nominated students. Each student for whom individual data were reported was nominated as at risk for EBD. Direct observations confirmed that at-risk students displayed levels of on-task behavior below 70% and/or engaged in at least three disruptions per 10-minute interval, similar to criteria used in other studies (Kamps et al., 2011; Caldarella et al., in press).

With the described screening procedure, 23 students (17 male and 6 female) were identified across the nine classes as at risk for EBD: three sixth graders, nine seventh graders and 11 eighth graders. Demographic information for these target students is displayed in Table 2; detailed behavioral information is shown in Tables 3 and 4.

Procedures

Consent. Teachers signed forms consenting to participate. Modified consent forms were mailed to the parents of all of the students in participating classes, as approved by the affiliated universities, and an individualized consent form was delivered or mailed to the target students and their parents, who signed and returned them to the researchers. The baseline phase began when consent forms were returned for the target students and selected classrooms.

Baseline. Five behavioral observation data points were collected in each class for the group on-task measure during baseline. Due to absences some of the target students were observed fewer than five times, but all were observed on at least three occasions to help researchers examine trends in the students' behavior. Baseline consisted of one or two 20-minute observations per day. The instructional activity determined the number of data collection sessions.

Training. Teacher training began after five completed baseline classroom observations. Teachers received instruction in the protocol in one approximately 60-minute session or in two approximately 30-minute sessions. Training consisted of a PowerPoint slide presentation with

embedded video clips of CW-FIT MS implementation in other middle school classrooms, with particular emphasis on behavior specific praise. Teachers also received a copy of the procedural fidelity checklist. The researchers offered coaching and modeling during the first two or three “start-up” lessons (classroom expectations taught in 10-minute lessons). Behavioral observation data were not recorded during the training phase. The intervention phase began when the teacher implemented procedures with at least 85% fidelity. Teachers continued to receive researcher feedback and support throughout the intervention phase.

Intervention. The independent variable, CW-FIT MS Tier 1, was implemented in the classes. Program components included teaching classroom expectations, using an interdependent group contingency, giving praise, and awarding points redeemable for a reward, as described below.

Classroom expectations. The primary Tier 1 component involved teaching students functional replacement behaviors for problem behaviors they had been using to obtain attention (adult or peer), escape work, or gain access to desired activities and materials. Students were taught specifically defined skills generated by the classroom teacher with student input. Behaviors selected by study classes included how to show respect, as well as one of the following (a) how to follow directions, (b) how to listen to the teacher, (c) how to stay on task, and (d) how to collaborate with peers. To teach the 10-minute expectation lessons, teachers used a guide: a partially scripted lesson plan comprised of a rationale, discussion, student practice, and teacher feedback. The expectations were posted where all students could see them.

Group contingency. In the next step of CW-FIT MS implementation, the teacher organized the class into teams of approximately three to five students and explained that a timer would be used and points would be earned by individual teams for using the posted skills. The

teacher displayed the chart on which points were recorded and explained that daily points would be earned towards a point goal earning a group reward. Teacher and students generated a list of possible rewards: for example, school supplies, snacks, a positive ticket for use at the school store, 1–2 minutes of free time, opportunity to play a game, a free assignment, or privilege of leaving class 1–2 minutes early.

After the expected skills had been introduced and the point system explained, teachers started each instructional period by reviewing the expectations (providing *precorrects*). They also completed the point chart, reminded students of the goal and reward, and scanned groups when a timer went off to determine whether student groups had earned a point. The intervals between opportunities for teams to earn points was 5 minutes. The teachers were also trained to use behavior specific praise in response to on-task behaviors. If a group did not earn points, specific feedback was provided by the teacher reminding students of the classroom behavior skills they had been taught.

Dependent Variables and Measures

Researchers chose dependent variables and measures based on previous CW-FIT studies so results could be compared (Caldarella et al., 2017; Wills et al., 2016; Wills et al., in press). Data were collected for (a) on-task behavior of all students in the class, (b) on-task behavior of the students identified as at risk for EBD, (c) disruptive behavior of the at-risk students, and (d) teacher praise and reprimand rates.

On-task behavior. The on-task behavior of all students in the class, the primary dependent variable, was evaluated using a paper and pencil momentary time sampling procedure. Data were recorded every 30 seconds during each 20-minute observation for each student team. A group would receive a point at the sound of the timer if every group member was on task:

following teacher instruction and complying with classroom expectations. Specific on-task behavior included looking at the teacher or materials, cooperating with peers during group work, reading, and writing as directed by the teacher. Every 30 seconds, observers quickly scanned the groups and recorded + if all students in the group were on task and – if any student in the group was off task.

Additionally, researchers documented the on-task behavior of each target student in the same manner as the groups of students. At the end of each 30-second interval the target students' on-task behavior was separately recorded after the group data were recorded.

Disruptive behavior. Observers also recorded the number of at-risk students' disruptions: any verbal or motor behavior that interfered with the target student's participation or other students' learning. Observers noted whether target students had been talking to a peer, calling out, making inappropriate physical contact or gestures, or engaging in other distracting activities not related to teacher instruction. A disruption was recorded at the end of each 30-second interval as a discrete event.

Praise and reprimands. During the 20-minute observation period, observers concurrently recorded the frequency of teacher praise and reprimands delivered in response to student behavior. Praise was recorded when the teacher made a verbal statement indicating approval of student behavior beyond acknowledgment of adequate performance or correct response. Teacher reprimand was defined as a verbal comment intended to stop the student(s) from misbehaving. A reprimand could reference a future negative consequence and might include negative comments and redirection.

Intervention fidelity. Fidelity of the CW-FIT MS intervention was collected on 100% of the sessions. A nine-item fidelity form was completed by observers at the end of each

implementation period. Indicators such as "timer used and set at appropriate intervals," "points awarded to teams for use of skills," and "points tallied and reward delivered" were either recorded as NP (not present) or P (present) on a 3-point Likert scale indicating implementation quality.

Social validity. Following completion of the study the teachers responded to a social validity survey including five questions on a 4-point Likert scale and two open-ended questions. The first open-ended question asked teachers what they had found most helpful in learning how to implement the CW-FIT MS program. The second open-ended question asked them to offer suggestions to modify the program for future use. A student survey included five questions: two with a *yes* or *no* response and three that were open ended. Students were asked to indicate in the open-ended questions what they liked most about the CW-FIT MS program, what (if anything) they did not like about it, and whether the program could help students get more work done in their classes.

Interobserver agreement

Before the study began, all data collectors practiced recording with the study observation techniques in middle school classes until reaching the criterion of 85% reliability across three consecutive sessions. Interobserver agreement (IOA) was collected on 28% of all observations during baseline, intervention, and withdrawal conditions. A second individual (e.g., graduate research assistant) collected the IOA data. Across all conditions, IOA was 92.67% for group on-task behavior, 95.22% for target student on-task behavior, 94.78% for target student disruptions, 85.78% for teacher praise, and 85.22% for teacher reprimands. Intervention fidelity IOA averaged 95.67% (range 91%–100%).

Design and Analysis

A single-subject ABAB withdrawal design was used to evaluate the effects of the intervention across study phases (Kazdin, 2010): including baseline, CW-FIT MS, withdrawal, and return to CW-FIT MS in the final phase. Phase change decisions were based on visual analysis of data trend and variability in the primary dependent variable—class-wide on-task behavior. The researchers' objective was to obtain five data points per condition for all participants. Changes in level, trend, and variability of group on-task behavior were analyzed visually. Changes in target student on-task and disruptive behavior, as well as teacher praise and reprimand, were compared using Cohen's *d* effect sizes to examine standardized mean differences across baseline and CW-FIT MS phases. To determine if teachers were implementing CW-FIT MS with fidelity over the course of intervention, researchers averaged the percentages from teacher fidelity forms collected following each observation. Researchers summarized the results of the teacher and student social validity questionnaires using descriptive statistics and qualitative coding of participants' open-ended responses.

Results

Treatment Fidelity

Teachers implemented CW-FIT MS with high fidelity averaging 91% (range 72%–100%) during intervention phases, compared with an average of 2.76% (range 0%–22%) during baseline phases. Suggestions were provided verbally to teachers in the intervention classes when fidelity scores were less than 80%, offering support on specific items on the fidelity checklist with low scores (e.g., increase specific praise, reduce reprimands, and increase timer frequency to prompt attention to appropriate behavior). Low fidelity occurred for 19.50% of sessions during intervention, and two teachers received feedback for more than two sessions of low fidelity.

Group On-Task Behavior

Figure 1 shows group on-task behavior across the nine participating classes. Baseline group on-task behavior averaged 40.30% (range = 30.50%–49.60%; $SD = 5.66$). During the first CW-FIT MS intervention phase, average group on-task behavior increased to 83.39% (range = 74.53%–95.51%; $SD = 8.07$). When CW-FIT MS was removed, percentages of group on-task behavior returned to near baseline levels, averaging 49.99% (range 36.20%–64.10%; $SD = 9.77$). When CW-FIT MS was reintroduced into classrooms, the average group on-task behavior increased to 84.37% (range 69.30%–96.33%; $SD = 8.53$).

Target Student On-Task Behavior

Table 3 shows on-task behavior across the 23 target students. Baseline target student on-task behavior averaged 42.62% (range = 17.00%–67.80%; $SD = 6.45$). During the first intervention phase, average on-task behavior increased to 84.88% (range = 62.00%–100.00%; $SD = 5.37$). Average target student on-task behavior decreased to 57.72% when the intervention was removed (range = 24.50%–85.36%; $SD = 6.24$). During the second CW-FIT MS implementation phase, average target student on-task behavior increased to 87.74% (range = 73.70%–99.00%; $SD = 3.73$).

Target Student Disruptive Behavior

Table 4 shows disruptive behavior across the 23 target students. Baseline target student disruptive behavior averaged 10.18 (range = 1.83–21.60; $SD = 1.93$). During the first intervention phase, average target student disruptive behavior decreased to 2.54 (range = 0.00–9.80; $SD = 1.80$). When the intervention was removed, average target student disruptive behavior increased to 6.47 (range = 0.00–15.80; $SD = 1.33$). During the second CW-FIT MS implementation phase, average target student disruptive behavior decreased to 2.28 (range =

0.00–6.80; $SD = 1.38$).

Teacher Praise to Reprimand Ratio

Table 5 shows teacher PRR across the nine participating classes. Baseline PRR averaged 0.36:1 (range = 0.00:1–1.60:1). During the first CW-FIT MS intervention, average PRR increased to 10.56:1 (range = 0.28:1–39.00:1). During the withdrawal phase, average PRR was close to baseline ($M = 0.73:1$, range = 0.12:1–2.00:1). When CW-FIT MS was reintroduced, average PRR was 6.82:1 (range = 0.29:1–27.20:1).

Social Validity

Teachers. The nine participating teachers expressed overall satisfaction with the CW-FIT MS intervention, rating the statements concerning acceptability and effectiveness of the intervention as *very true*, *mostly true*, or *somewhat true*. No items were marked *not true*. All teachers reported receiving adequate training on the CW-FIT MS program; two-thirds responded *very true* that the program was easy to learn and implement, the remainder reported *mostly true*. Teachers responded *mostly true* and *very true* regarding the effectiveness of the procedural fidelity feedback form as a teaching tool. All but one of the teachers indicated the in-class support and researcher feedback were helpful (the other teacher indicating *somewhat true*). On the item regarding whether they planned to use CW-FIT MS again in the future, two-thirds of the teachers responded *very true* or *mostly true*, and one-third responded *somewhat true*.

Teachers reported collaboration with researchers or other teachers who had implemented CW-FIT MS, and watching videos of intervention implementation, as being most helpful. Other helpful features of the training included learning about fidelity and observing the effects of praise on behavior. Suggestions for improvement included being more flexible in using the intervention based on the educational activity. One teacher suggested using CW-FIT MS only during

independent learning, while another wanted to use it for the entire 90 minutes of class. A third teacher suggested adjusting the length of the timer interval to better align with the class activity. Other teachers mentioned giving more genuine praise at various times throughout the lesson, not limited to the timer beeps; some suggested not using a timer. Another teacher requested a list of non-monetary rewards for middle school classes. One teacher suggested setting her class goal lower, allowing students more opportunities to earn points so she would feel more comfortable withholding points when the class did not earn them.

Students. A total of 221 students (94.44%) across the nine participating classes completed social validity surveys; 87.33% indicated they enjoyed CW-FIT MS. On the open-ended questions as to what they liked and did not like about CW-FIT MS, the most common *likes* were prizes/rewards ($n = 162$), increased productivity/learning (e.g., increased focus, better behavior, etc.; $n = 51$), and a more positive environment (e.g., more respect, better relationships, calmer classroom; $n = 14$). Concerning aspects of CW-FIT MS they did not like, 120 students (54.30%) stated they did not dislike any aspects. Others commented they disliked not earning points/not winning ($n = 22$), being distracted by the timer intervals ($n = 14$), having classmates not cooperate or having a “bad” team ($n = 8$), disagreeing with how the points were allocated ($n = 10$), and having observers present ($n = 19$). Others had more general comments: “boring,” “annoying,” “distracting,” “too difficult,” “too easy,” “doesn’t help,” “everything”; $n = 26$).

When asked if CW-FIT MS could help students get more work done in their classes, 89.59% of students responded *yes*. Common responses to the question of why CW-FIT MS would be helpful included rewards ($n = 96$), increased productivity ($n = 63$), increased motivation and fun in the classroom ($n = 53$), more positive environment ($n = 14$), increased accountability and life skills learned ($n = 21$), and increased learning ($n = 10$). Only 19 students

stated that CW-FIT MS would not be helpful in other classes, explaining that it could be distracting, students might pretend to work or care more about the points than working, it could create too much pressure, students might forget or not listen, and students should just be on task anyway.

Discussion

The purpose of the present study was to investigate the feasibility and effects of CW-FIT MS Tier 1 implementation in five middle schools. Results are discussed in terms of the five research questions. First, concerning implementation fidelity, data collected across the nine classes during all intervention periods resulted in high fidelity, in contrast to lack of fidelity during baseline conditions. These findings are consistent with previous research documenting high levels of fidelity during CW-FIT MS implementation at the middle school level (Orr et al., 2019; Wills et al., in press) and in elementary school classes (Kamps et al., 2015; Wills et al., 2016).

Second, results suggested that on-task behavior improved significantly during CW-FIT MS implementation compared to baseline conditions. The data displayed in Figure 1 show a marked and immediate increase in group on-task behavior for all nine classes at the beginning of the first intervention phase. An immediate decrease in level was subsequently noted in each class when CW-FIT MS was removed. A positive change in level and trend of group on-task behavior was also observed across all classes at the beginning of the second CW-FIT MS phase. In all classes, the data suggested a functional relationship between the intervention and improved on-task behavior. These findings are consistent with the results of earlier studies conducted in elementary (Caldarella et al., 2015; Kamps et al., 2011) and middle school (Orr et al., 2019; Wills et al., in press) classes.

Third, results indicated marked improvement in the on-task behavior of target students (Table 3), with a corresponding reduction in their disruptions (Table 4). Similar improvement in on-task behavior and increased engagement for students identified as at risk for EBD following CW-FIT implementation has also been documented in previous research with younger students (Weeden et al., 2016; Wills et al., 2016).

Fourth, findings revealed desired changes in teachers' praise-to-reprimand ratios (Table 5). The data disclosed a significant increase in the delivery of praise during both CW-FIT MS phases compared to baseline and reversal phases. All nine teachers also demonstrated decreases in reprimands during CW-FIT MS implementation. The current findings corroborate results achieved in prior studies (e.g., Orr et al., 2019; Wills et al., in press).

Finally, concerning social validity, results of the satisfaction survey indicated that overall the teachers approved of the intervention and found it easy to implement. Suggestions for improvement included being more flexible with the intervention, including the timer, depending on the class activity. Students also responded positively, with results indicating that most enjoyed the intervention, especially the rewards, and believed that CW-FIT MS improved behavior and productivity in the classroom. Students' dislikes included not earning points and being distracted by the observers or timer. Social validity findings aligned favorably with similar results attained in elementary school and middle school special education classrooms (Caldarella et al., 2015; Orr et al., 2019).

Limitations and Areas for Future Research

Although study results substantiated previous CW-FIT MS research in a number of salient ways, several limitations should be noted. For example, the sample of participating students and teachers involved only nine classes in five different schools across two primary

research sites. The research sample must therefore be considered localized rather than nationally representative, limiting inferences that can be made concerning external validity and generalizability of the findings.

Limitations involving the classroom observations weakened the internal validity of the study. Logistical constraints prevented observers from being blind to condition, increasing the possibility of observer bias. Additionally the limited number of teachers prevented random assignment of classes to experimental or control conditions. Due to absence or suspension, behavior data (on-task and disruptive) were deficient for Target Student 11 during the withdrawal and final intervention phases and for Target Student 12 during the withdrawal phase (see Tables 3 and 4).

Participant characteristics might similarly be considered limiting. Although target students were identified as being at risk for EBD using Stage 1 of the SSBD, the majority of student participants did not have an EBD classification with an accompanying individualized education program or 504 Plan. However, earlier studies conducted with elementary and middle school students who had been identified with cognitive or behavioral disability, and were receiving special education services, demonstrated improved behavioral performance and increased teacher praise subsequent to CW-FIT implementation (Orr et al., 2019; Weeden et al., 2016), which gives credence to assumptions made despite this weakness.

Research is needed to address these limitations and to extend current findings to additional student populations, age groups, settings, and behaviors. Identifying a way to accurately assess the impact of CW-FIT MS on students' academic performance is also considered a research priority (Wills et al., in press). In addition, future examinations could benefit from randomly assigning middle school classes to treatment or control conditions using

larger sample sizes.

To more thoroughly examine the beneficial effects of CW-FIT MS, researchers would do well to recruit larger numbers of students with disabilities, including those with EBD (Weeden et al., 2016). Though current findings suggest that CW-FIT MS Tier 1 may be sufficient for the majority of students, further research is needed to investigate a variety of Tier 2 interventions that might be successfully implemented with non-responders in general education classes and in more restrictive settings such as special education classes (Wills et al., 2016; Wills et al., in press).

Conclusion

Current findings are consistent with results of earlier studies documenting the positive effects of CW-FIT on teacher praise and reprimands, as well as classroom behavior and student engagement across a range of age groups and ability levels, including students at risk for EBD (Caldarella et al., 2015; Kamps et al., 2015; Orr et al., 2019; Weeden et al., 2016; Wills et al., in press). The present study also extends the results of recent research (Orr et al., 2019; Wills et al., in press) by demonstrating positive intervention effects in novel middle school environments during a feasibility trial with CW-FIT MS implementation.

References

- Algozzine, B., Wang, C., & Violette, A. S. (2011). Reexamining the relationship between academic achievement and social behavior. *Journal of Positive Behavior Interventions* 13(1), 3–16. doi:10.1177/1098300709359084
- Bernstein, E. (2002). *Middle school and the age of adjustment: A guide for parents*. Westport, CT: Bergin & Garvey.
- Bushaw, W. J., & Lopez, S. J. (2010). A time for change: The 42nd annual Phi Delta Kappa/Gallup Poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 92(1), 8-26.
- Caldarella, P., Larsen, R. A. A., Williams, L., Wills, H. P., & Wehby, J. H. (in press). Teacher praise-to-reprimand ratios: Behavioral response of students at risk for EBD compared with typically developing peers. *Education and Treatment of Children*.
- Caldarella, P., Williams, L., Hansen, B. D., & Wills, H. P. (2015). Managing student behavior with Class-Wide Function-Related Intervention Teams: An observational study in early elementary classrooms. *Early Childhood Education Journal*, 43, 357–365. doi:10.1007/s10643-014-0664-3
- Caldarella, P., Williams, L., Jolstead, K. A., & Wills, H. P. (2017). Managing student behavior in an elementary school music classroom: A study of class-wide function-related intervention teams. *Update: Applications of Research in Music Education*, 35(3), 23–30. doi:10.1177/8755123315626229
- Chung, H., Elias, M., & Schneider, K. (1998). Patterns of individual adjustment changes during middle school transition. *Journal of School Psychology*, 36(1), 83–101. doi:10.1016/S0022-4405(97)00051-4

- Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology: With implications for teacher education. *Educational Psychologist, 36*(2), 103–112. doi:10.1207/S15326985EP3602_5
- Espin, C., & Yell, M. (1994). Critical indicators of effective teaching for preservice teachers: Relationships between teaching behaviors and ratings of effectiveness. *Teacher Education and Special Education, 17*, 154–169. doi:10.1177/088840649401700303
- 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*(1), 55–64. doi:10.1177/1063426611421157
- 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., Heitzman-Powell, L., Laylin, J., Szoke, C., Hobohm, 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:10.1177/5109830071139893
- Kazdin, A. E. (2010). *Single case research designs: Methods for clinical and applied settings*. New York, NY: Oxford University Press.
- Marchant, M., & Anderson, D. H. (2012). Improving social and academic outcomes for all learners through the use of teacher praise. *Beyond Behavior, 21*(3), 22–28.
- McIntosh, K., Flannery, K. B., Sugai, G., Braun, D. H., & Cochrane, K. L. (2008). Relationships between academics and problem behavior in the transition from middle school to high

- school. *Journal of Positive Behavior Interventions*, *10*, 243–255. doi:
10.1177/1098300708318961
- Muscott, H. S., Mann, E. L., & Le Brun, M. R. (2008). Positive behavioral interventions and supports in New Hampshire. *Journal of Positive Behavior Interventions*, *10*, 190–205. doi:10.1177/1098300708316258
- Närhi, V., Kiiski, T., Peitso, S., & Hannu, S. (2015). Reducing disruptive behaviours and improving learning climates with class-wide positive behaviour support in middle schools. *European Journal of Special Needs Education*, *30*(2), 274–275. doi:10.1080/08856257.2014.986913
- Orr, R. K., Caldarella, P., Hansen, B. D., & Wills, H. P. (2019). Managing student behavior in a middle school special education classroom using CW-FIT tier 1. *Journal of Behavioral Education*. Advance online publication. doi:10.1007/s10864-019-09325-w
- Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS identifying areas for enhancement. *Journal of Positive Behavior Interventions*, *15*, 39–50. doi:10.1177/1098300712459079
- 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
- Sinclair, M. F., Christenson, S. L., Evelo, D. L., & Hurley, C. M. (1998). Dropout prevention for youth with disabilities: Efficacy of a sustained school engagement procedure. *Exceptional Children*, *65*, 7–21.
- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education*

- and Treatment of Children, 31*, 351–380.
- Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review, 26*, 333–368.
- Walker, H. M., & Severson, H. H. (1992). *Systematic Screening for Behavior Disorders* (2nd ed.). Longmont, CO: Sopris West.
- 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*(1), 285–293.
- Wills, H. P., Caldarella, P., Mason, B. A., Lappin, A. M., & Anderson, D. H. (in press). Improving student behavior in middle schools: Results of a Tier 1 classroom management intervention. *Journal of Positive Behavior Interventions*.
- 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., 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., & Nsubunga, B. (2010). The class-wide function-based intervention team program. *Preventing School Failure, 54*, 164–171. doi:10.1080/10459880903496230
- Young, E. L., Caldarella, P., Richardson, M. J., & Young, K. R. (2012). *Positive behavior support in secondary schools: A practical guide*. New York, NY: Guilford Press.

Table 1

Descriptive Data for Participating Schools

School	Student Total	Student Demographics (Percentage)							
		FRL ^a	American Indian ^b	Asian	Black	Hispanic	Pacific Islander ^c	White	Multi-Race
1	648	67.0	0.9	5.5	15.0	18.0	0.8	49.0	10.0
2	558	83.3	0.2	5.5	56.0	19.0	0.4	14.6	4.1
3	812	60.1	0.3	5.3	16.6	15.0	1.6	55.0	5.8
4	847	56.9	1.9	0.8	1.1	35.7	3.8	54.9	1.9
5	875	55.0	1.3	0.3	0.7	34.8	1.6	57.9	3.3

Note. ^a Free and reduced priced lunch. ^b American Indian/Alaskan Native. ^c Pacific Islander/Native Hawaiian

Table 2

Target Student Demographics

	Age	Gender	Race/ethnicity	IEP/504 Plan
Student 1	11	M	White	No
Student 2	11	M	White	No
Student 3	11	M	Black	No
Student 4	12	F	Black	No
Student 5	12	M	Hispanic	IEP
Student 6	12	M	White	IEP
Student 7	12	F	Black	No
Student 8	12	F	Black	No
Student 9	13	M	Black	No
Student 10	13	F	White	No
Student 11	14	M	White	No
Student 12	13	M	Black	IEP
Student 13	14	M	White	No
Student 14	13	F	Black	No
Student 15	14	M	Hispanic	No
Student 16	14	M	White	No
Student 17	13	M	White	No
Student 18	13	M	White	No
Student 19	13	M	White	No
Student 20	12	M	Hispanic	No
Student 21	12	F	White	No
Student 22	12	M	Hispanic	504 Plan
Student 23	12	M	White	No

Table 3

Changes in Target Students' On-Task Behaviors Across Conditions

Student	On-Task Behavior (%)			
	Baseline	CW-FIT	Withdrawal	CW-FIT
1	17.00	95.00	27.05	93.00
2	44.00	98.00	55.55	99.00
3	46.40	96.98	50.60	96.50
4	44.55	98.50	74.50	98.50
5	40.00	85.80	41.44	85.00
6	23.33	100.00	32.10	99.00
7	55.67	93.13	62.00	80.50
8	36.60	74.75	55.50	78.00
9	47.00	71.20	52.50	83.00
10	67.80	79.20	59.29	82.30
11	59.00	97.00	*	*
12	47.25	83.33	63.28	*
13	35.60	81.00	70.19	73.70
14	27.40	77.20	35.50	79.60
15	35.00	79.00	63.13	88.75
16	27.93	89.17	68.50	96.00
17	51.20	83.73	76.69	96.00
18	51.72	83.94	71.48	91.50
19	53.78	80.55	83.41	89.00
20	44.28	84.00	60.41	87.00
21	43.00	92.50	85.36	98.23
22	32.50	66.25	24.50	73.70
23	49.24	62.00	56.96	74.30
Average (<i>SD</i>)	42.62 (6.45)	84.88 (5.37)	57.72 (6.24)	87.74 (3.73)

Note. *Insufficient data due to students' absence or suspension

Table 4

Changes in Target Students' Disruptive Behaviors Across Conditions

Student	Disruptive Behavior (#)			
	Baseline	CW-FIT	Withdrawal	CW-FIT
1	18.20	0.80	8.60	0.20
2	8.00	0.20	4.60	0.00
3	11.60	0.60	3.75	0.40
4	8.40	0.00	3.60	0.00
5	10.42	3.00	5.80	2.00
6	15.33	0.00	9.00	0.40
7	1.83	0.75	4.60	5.33
8	14.60	7.00	11.60	5.60
9	15.20	9.80	15.80	6.80
10	8.60	7.20	14.40	5.40
11	13.00	1.00	*	*
12	8.00	4.33	5.60	*
13	7.60	5.00	8.20	6.60
14	21.60	5.80	15.60	4.40
15	13.00	0.80	6.00	1.25
16	9.00	1.33	6.40	0.80
17	9.20	1.00	2.50	0.60
18	9.75	0.25	4.75	1.50
19	8.60	1.75	2.40	2.20
20	2.60	0.00	3.60	2.00
21	9.20	0.20	0.40	0.00
22	8.00	4.50	0.00	0.00
23	2.40	3.00	5.20	2.40
Average (<i>SD</i>)	10.18 (1.93)	2.54 (1.80)	6.47(1.33)	2.28 (1.38)

Note. *Insufficient data due to target students' absence or suspension.

Table 5

Teacher Praise-to-Reprimand Ratios Across Classrooms

	Praise-to-Reprimand Ratio			
	Baseline	CW-FIT	Withdrawal	CW-FIT
Classroom 1	0.05:1	4.00:1	0.25:1	2.79:1
Classroom 2	0.57:1	24.00:1	2.00:1	14.00:1
Classroom 3	0.38:1	4.56:1	0.14:1	2.42:1
Classroom 4	0.33:1	1.35:1	0.12:1	2.14:1
Classroom 5	0.10:1	0.28:1	0.12:1	0.29:1
Classroom 6	0.18:1	14.00:1	0.80:1	27.20:1
Classroom 7	1.60:1	39.00:1	1.75:1	6.00:1
Classroom 8	0.00:1	4.57:1	1.00:1	4.67:1
Classroom 9	0.00:1	3.31:1	0.40:1	1.90:1
Average	0.36:1	10.56:1	0.73:1	6.82:1

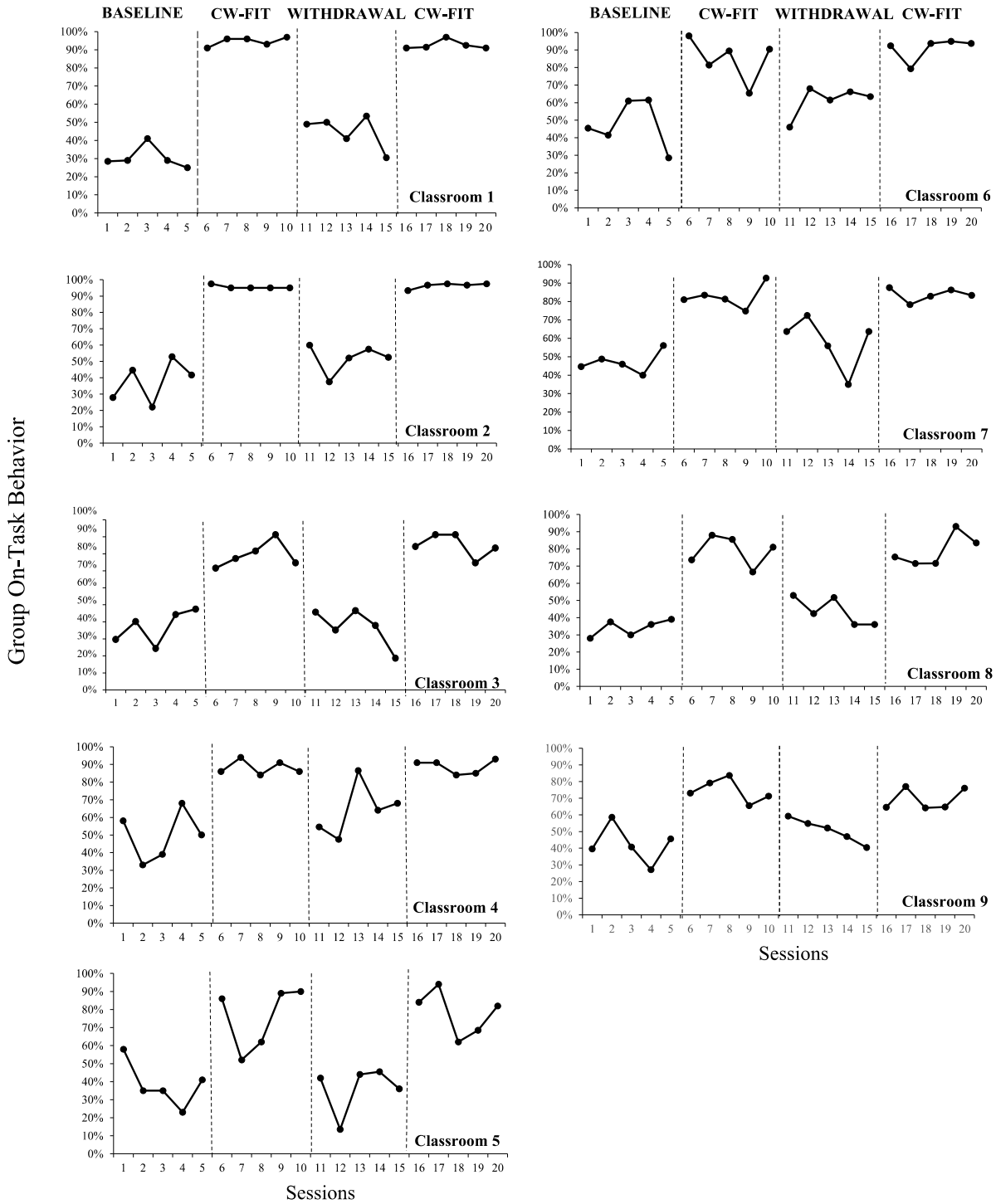


Figure 1. Group on-task behavior across classrooms and conditions.