EDUCATION AND TREATMENT OF CHILDREN Vol. 41, No. 1, 2018

The Effects of the CW-FIT Group Contingency on Class-wide and Individual Behavior in an Urban First Grade Classroom

Anna Schmidt Naylor

Lawrence Public Schools

Debra Kamps

Howard Wills

University of Kansas

Abstract

The current study examined the effects of the Class-wide Function-related Intervention Teams (CW-FIT), a class-wide group contingency, on the on-task behavior of all students in a first grade class and the on-task and disruptive behavior of three target students within that class who were nominated by their teacher through a behavioral screening. The classroom teacher used a systematic instruction model (i.e., direct instruction, model, and role-play) to teach three behavioral skills, including (a) gaining the teacher's attention appropriately, (b) following directions, and (c) ignoring inappropriate peer behavior. These skills were reinforced through use of an interdependent group contingency, the CW-FIT program, incorporating class teams, awarding points, and rewards for meeting point goals. The effects of the intervention on target student and class behavior were evaluated in a reversal design. Results showed the CW-FIT program was effective in increasing the on-task behavior of the class and target students. A reduction in disruptive behavior was also seen for each target student during intervention. Students and teacher reported a high degree of satisfaction with the intervention.

Keywords: group contingency, classroom management

Discipline and behavior problems have been identified by teachers and the general public as among the biggest problems facing public schools today (Chafouleas, Volpe, Gresham, & Cook, 2010; Rose &

Author Note: The research was funded by the Office of Special Education and Rehabilitation Services, Department of Education (H324X010011). Opinions expressed herein are those of the authors and do not necessarily reflect the position of the funding agency. We gratefully acknowledge the participating teacher, students, and families for their time and ongoing support.

Address correspondence to: Debra Kamps, 444 Minnesota Avenue, 3rd floor, Kansas City, KS 66101. E-mail: dkamps@ku.edu.

Gallup, 2005). Numerous organizations, including the American Psychological Association, Center for the Study and Prevention of School Violence, and Center for Positive Behavioral Interventions and Supports, advocate positive and preventative strategies (Horner, Sugai, & Anderson, 2010; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). Professionals in the field of special education research also advocate comprehensive strategies and preventative approaches to increase learning and improve behavior, focusing on early screening of children for emotional and behavioral disorders and the development of intervention programs designed to prevent these disorders in at-risk individuals (Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008; Walker et al., 1998).

Class-wide Function-related Intervention Teams (CW-FIT, Kamps, Wills, Bannister et al., 2015; Kamps et al., 2010; Wills et al., 2010) is a program developed in response to these recommendations and the concerns surrounding problem behavior in schools. CW-FIT is intended as a preventative approach for classrooms in which behavior problems may be more likely to occur, as well as a tool to address problem behaviors that may be ongoing in classrooms. Thus, CW-FIT was designed based on evidence- and research-based interventions to address disruptive student behavior. The program consists of several key components, including (a) teaching specific behavioral skills that may increase appropriate behavior, (b) increasing teacher attention (praise and points) to appropriate behavior, (c) reducing teacher and peer attention for inappropriate behavior, and (d) differentially reinforcing behavioral skills through a token system (points) within a group contingency format.

A primary component of CW-FIT is reinforcement of desirable classroom behaviors through a group contingency. The literature defines three different types of group contingency methods. They are (a) independent group contingency method in which students earn incentives based on their individual behavior, (b) dependent group contingency when a group or team of students earns incentives based on the behavior from one student or a small subgroup within the team, and (c) interdependent group contingency method that requires all team members to behave in a specified way in order to reach a criterion to earn incentives (Tingstrom, Sterling-Turner, & Wilczynski, 2006). CW-FIT uses an interdependent group contingency method.

Group contingencies have been used in classrooms to decrease disruptive behavior (Harris & Sherman, 1973; Lea, Bray, Kehle, & DioGuardi, 2004; Salend, Reynolds, & Coyle, 1989), and to increase on-task and academic behaviors (Lannie & McCurdy, 2007; Packard, 1970;

Popkin & Skinner, 2003; Thorne & Kamps, 2008). The "Good Behavior Game" (GBG) developed by Barrish, Saunders, and Wolf (1969), was the first interdependent group contingency approach to show significant class-wide improvements in students' on-task behavior (Embry, 2002). In the GBG, the class is divided in teams that (a) received points against the team for problem behaviors, and (b) earned incentives for meeting the goal of less than the specified number of points (Barrish et al., 1969). Several meta-analysis studies have found group contingencies including the GBG to be effective interventions in public school settings (Lea et al., 2004; Maggin, Johnson, Chafouleas, Ruberto, & Berggren, 2012; Stage & Quiroz, 1997). In the Maggin et al. analysis, over 20 studies showed the effects of the GBG for improving on-task behavior and following rules (e.g., Lannie & McCurdy, 2007; Lohrmann & Talerico, 2004).

One important aspect of group contingencies is that reinforcement can be applied differentially to behaviors for small groups of children. Differential reinforcement is the reinforcement of some members of a response class and not others (Catania, 1998). Interdependent group contingencies, such as the GBG and CW-FIT, require all team members to behave in a specified way in order to reach a criterion to earn incentives, thus embedding differential reinforcement within the group contingency (Kamps, Wills, Bannister et al., 2015; Tingstrom et al., 2006; Wills et al., 2010). Several group contingency studies have demonstrated the robust effects of differential reinforcement (Kamps, Conklin, & Wills, 2015) including randomization of reinforcers (Coogan, Kehle, Bray, & Chafouleas, 2007; Kamps et al., 2008; Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Thorne & Kamps, 2008). Group contingencies also typically incorporate use of points or tokens through which desired behaviors are reinforced (Kelshaw-Levering et al., 2000). In a token economy, a medium of exchange (e.g., poker chip, tally mark, point) is selected that can be used to purchase a "backup" reinforcer (e.g., candy, free-time). The token becomes a conditioned reinforcer due to its pairing with the backup reinforcer and it bridges the time between the occurrence of a behavior and delivery of the backup reinforcer (Ayllon & Azrin, 1968; Maggin, Johnson, Chafouleas, Ruberto, & Berggren, 2011).

One important benefit of the use of a group contingency to manage behaviors within a classroom is that it is possible to change the behavior of a large number of students using only one intervention. Group or class-wide interventions are often preferred over individual contingencies for multiple students because of the management difficulties for teachers associated with individual behavior plans (Skinner,

Skinner, Skinner, & Cashwell, 1999). Thus, interventions that address the needs of the greatest number of students may be beneficial for teachers and students. One method to increase the number of individuals affected by a class-wide intervention would be to target a common function of problem behavior. A review of literature on functional assessment in schools (Anderson, Rodriguez, & Campbell, 2015; Ervin et al., 2001) reported that in studies in which functional analyses were conducted with students without disabilities, problem behavior was most frequently found to be sensitive to adult attention as a reinforcer. One hypothesis based on these findings is that interventions that reduce teacher attention for problem behavior and employ it as a reinforcer for desirable behavior could affect the behavior of the greatest number of students. In CW-FIT, teacher praise (attention) for use of skills was embedded in the procedures including teacher praise to accompany the delivery of points for use of skills. Similarly, functioning as a member of a team to earn points and rewards may reduce engaging in inappropriate behavior maintained by peer attention.

CW-FIT Research

The CW-FIT intervention has been demonstrated in several studies to date showing improved on-task behavior in preschool and elementary classrooms (Caldarella, Williams, Hansen, & Wills, 2015; Jolstead, Caldarella, Hansen, Korth, Williams, & Kamps, 2017; Kamps et al., 2010; Kamps, Wills, Bannister, et al., 2015) and in a classroom serving students with behavioral disorders (Weeden, Wills, Kamps, & Kottwitz, 2016). Improvements in students' disruptive behaviors was shown in a recent study using a self-management component concurrently with CW-FIT (Kamps, Conklin, & Wills, 2015). The benefits of using CW-FIT across multiple time periods were also found in a first grade classroom (Wills, Iwaszuk, Kamps, & Shumate, 2014). Recently, improvement in skills directly linked to the CW-FIT skills (hand raising, compliance to directions) was found across four classrooms in an urban culturally diverse school (Conklin, Kamps, & Wills, 2017). Only one study has demonstrated the effects of CW-FIT in classes serving several students with English as a second language (Kamps et al., 2010). A limitation in the study is that effects were not experimentally validated for the specific individuals with English language learners (ELL), but rather as members of the class. Additional demonstration of the effects of CW-FIT with students with behavioral risk as well as concerns related to instruction for children learning English is warranted.

Purpose

The purpose of the current study was to examine the effects of the CW-FIT multi-component intervention on the behavior of first grade students in an urban elementary charter school serving a high number of students who are ELL. This expands the evidence of CW-FIT for a novel group of students, and provides additional evidence for individual students' decreases in disruptive behaviors in a classroom with high levels of off-task behavior. In this setting, several factors contributed to significant challenges in managing behaviors and ensuring effective instruction: (a) a high percent of children immigrating from outside the United States with very limited English; (b) high levels of poverty; and (c) high rates of disruptive behaviors (Hart & Risley, 1995; Jensen, 2009; Klingner, & Soltero-Gonzalez, 2009; Stage & Quiroz, 1998). Specifically, the CW-FIT intervention was implemented to address the following research questions: (a) What are the effects of the intervention on disruptive and on-task behavior for target students who are ELL and identified with ongoing problem behavior? (b) What are the effects of the intervention on class-wide on-task behavior? and (c) What are the effects of the intervention on teacher behavior?

Method

Participants and Setting

Participants. Three children were chosen as target students by the teacher and researcher (i.e., Schmidt Naylor). All students were from families who had recently emigrated from Africa. Lily, a 6-year-old girl, had been nominated by her teacher, Ms. D., during a routine behavior screening conducted by university researchers 4 months prior to the current study. Lily was identified as an "externalizer" based on her scores on the *Systematic Screening for Behavior Disorders* (*SSBD*) (Walker & Severson, 1992). The criteria for identification as an externalizer were either less than 31 points on the adaptive scale, more than 33 points on the maladaptive scale, or two or more critical events (e.g., damages others' property, sets fires). Lily had an adaptive score of 36, a maladaptive score of 28, and four critical events. Ms. D. identified stealing, ignoring teacher warnings and reprimands, requiring punishment to terminate an inappropriate behavior, challenging classroom rules, and running around the classroom as behaviors frequently exhibited by Lily. Albert, a 6-year-old boy, was

also identified as an externalizer through the *SSBD* with an adaptive score of 30, a maladaptive score of 28, and two critical events. On Albert's *SSBD*, Ms. D. noted that Albert had tantrums and exhibited a sad affect to the extent that it interfered with normal peer and classroom activities. Other behaviors identified in Albert's screening were frequent disturbances in the classroom and lack of compliance with teacher requests. Malachi, a 6-year-old boy, was identified as an externalizer, but was expelled from school the week following the start of the study and was therefore unable to participate. Faith, a 6-year-old girl, was then chosen as a third target student. Although Faith had not been nominated during the behavior screening, Ms. D. reported Faith's behavior was a concern, as she frequently disrupted class and often refused to follow directions. All of these children were reading at benchmark for their grade level, as determined by the *Dynamic Indicators of Basic Early Literacy Skills* (*DIBELS*) (Kaminski & Good, 1998). All three students were ELL and had parents who moved to the United States from Africa.

The teacher involved in this study was recruited by university researchers, and was paid $300.00 for her participation. Ms. D. was a first year teacher who had previously requested help with this class from the school leadership team. Teaching strategies employed by Ms. D. during the core curriculum period were primarily lecture and read-aloud. Students were typically required to actively listen, read academic material to themselves or with a partner, and complete worksheets related to lectures with partial independence.

Setting. A class of 11 first grade students and their teacher in an urban core, charter elementary school participated in this study. The student body of 423 students comprised 36% Africans, 34% African Americans, 21% Vietnamese students, 4% Caucasians, 4% Hispanics, and less than 1% Chinese students and Native Americans. A majority of these students (62%) were ELL, with limited English proficiency. The school qualified for Title I status, as 94% of the student body was eligible for free and/or reduced lunch.

The study was conducted within the classroom, during the final hour of the day. The Core Knowledge Curriculum taught during this time encompassed all major content areas, but focused on science and social studies. The group of students who participated in this study was together only during this class period due to the Direct Instruction curriculum used in the school. This curriculum requires children be placed within their skill level for each content area, and therefore with different groups of peers for each class.

A school-wide system of positive behavior supports (SWPBS) was in place in the school in which the study was conducted. Five

school-wide behavioral expectations had been established, and behaviors that exemplified these expectations were reinforced through the school-wide reward system. Teachers would deliver hole-punches to students' reward cards for use of the expected appropriate behaviors taught as part of the SWPBS system (e.g., following directions, completing work, being kind). Pre-determined numbers of hole-punches were then redeemed for tangible items (e.g., candy). This was used as a simple way to acknowledge students' appropriate behaviors more frequently in that all students kept the reward cards at their desks and the teacher had easy access to the hole-puncher. Ms. D., the teacher who participated in the current study, used this system of reinforcement in her classroom throughout the day prior to the study and during baseline and CW-FIT conditions.

Procedures

Baseline. Baseline conditions consisted of "business as usual." Students were taught the SWPBS expectations during the first month of school and during both baseline phases, Ms. D. reinforced appropriate student behavior with hole-punches on individual cards that students kept at the desk. School and classroom expectations were posted in the classroom. The use of hole-punches was the agreed upon token system to reward appropriate behaviors for the first grade classes. Tickets and the number of hole-punches were reviewed at the end of the day and could be exchanged for candy.

CW-FIT intervention. The CW-FIT intervention consisted of three primary components of (a) teaching three appropriate classroom behaviors, (b) using a group contingency, and (c) differentially reinforcing those behaviors through a class-wide program using teams and awarding points. Descriptions of the three components are as below.

Teaching appropriate behaviors. Three primary skills were identified based on best practices for teaching prosocial behaviors as published in prior curricula and studies (e.g., Mitchem, Young, West, & Benyo, 2001; *Tough Kid Social Skills,* Sheridan, 2010; *Utah's B.E.S.T Project,* Reavis, Jenson, Kukic, & Morgan, 1988; *Skillstreaming,* McGinnis, 2010). Skills were taught using scripted lessons consisting of defining the skill, linking to the school expectations, modeling, role playing, and feedback during the lesson and then when implementing CW-FIT. The first, "How to Get the Teacher's Attention," was designed to be a functional replacement for undesirable behaviors that may have been maintained by teacher attention (e.g., calling-out, getting out of seat). This skill was defined by four behavioral components: (a) Look at the teacher, (b) Raise your hand, (c) Wait quietly for the

teacher to call on you, and (d) Ask your question or give your answer. The second skill, "Follow Directions the First Time," was designed to broadly address compliance behaviors. This included behaviors commonly expected from students in the classroom, such as working quietly at a desk and attending to the teacher during instruction. This skill consisted of: (a) Look at the person (teacher) and listen, (b) Say OK, (c) Do it, and (d) Check back if needed. The final skill, "Ignore Inappropriate Behavior," was designed to minimize peer attention for inappropriate behaviors (e.g., touching peers, using inappropriate language) that might be sensitive to attention from peers as a reinforcer. The behavioral components of this skill were: (a) Keep a pleasant face, (b) Look away from the person, (c) Keep a quiet mouth, (d) Pretend you are not listening, and (e) Follow directions—do your work. The CW-FIT skills aligned with the SWPBS skills of Be Peaceful (raise hand to talk), Be Respectful (follow directions), and Be Responsible (complete your work and accept outcomes of your behavior).

Preceding intervention, Ms. D. was given training literature that explained the teaching components of CW-FIT and that offered scripts for teaching the three primary skills of the intervention. A 45-min meeting was then conducted with Ms. D. to review the literature and discuss the instructional methods to be used. Following baseline, the three identified skills were systematically taught to the class. Each skill lesson took 10 to 15 min and was repeated the following 3 days. Ms. D. taught the remaining two skills using the same methods. Three to four days were spent teaching each new skill, and previously taught skills were reviewed daily.

Group contingency and differential reinforcement. The CW-FIT game was introduced to the students after the third and final skill had been taught. The researcher explained the game to the class as follows. Each day, students would have the opportunity to earn a reward. The class would be divided into groups (teams) of three or four students each. If each student within the group displayed the skills that had been taught, they would earn points toward a daily goal. Ms. D. would set a point goal at the beginning of the class period and intermittently a timer would beep. At that time, Ms. D. would look around the room and determine if each student was displaying one or more of the skills. If all students in a group were doing so, Ms. D. would mark a tally in their group's column on a Team Chart and deliver behavior specific praise, for example, "Super job, Group A. You are all remembering to raise your hand to get my attention." If one or more members of a group were not displaying the skills, an X would be placed in the column and Ms. D. would provide corrective feedback

and a verbal reminder of behaviors that would allow them to earn points (e.g., "Group B, when the timer beeped you were not following directions. Your group needs to follow directions the first time to earn a point toward today's goal."). Groups that met the point goal by the end of the class period would receive a reward.

Initially, a dense schedule of reinforcement was established in which the interval between the deliveries of points was 1 to 2 min in length. This schedule was thinned as students become more engaged and less disruptive ranging up to intervals of 1 to 4 min. In order to increase the probability students would contact reinforcement, Ms. D. was advised to set a point goal that she thought the groups could reach. There was a maximum of 40 min to implement CW-FIT each day; however, this varied due to classroom routines and activities. Point goals ranged from 10–20, and were based on time available, Ms. D.'s assessment of classroom behaviors, and previous goals. Ms. D. set the timer for 1 to 4 min based on the goal and time available to play, and to thin the schedule of reinforcement.

Reinforcer assessment. A reinforcer assessment was conducted by the researcher and teacher prior to the introduction of the CW-FIT game. Students were told they would be playing a game in which they could earn rewards. The students suggested items (e.g., jewelry, stickers, temporary tattoos, gel pens, candy) and the teacher wrote them on the chalkboard. From this list, the researcher and teacher helped the students discriminate which items would be both appropriate and available. Additional rewards that were similar to items identified through the assessment were chosen (e.g., key chains, small tablets, pencils, extra recess, 5-min dance party) by the researcher to decrease the probability of satiation.

Intervention start up. During the first week, the researcher assisted Ms. D. by setting the timer and providing minimal verbal prompts (e.g., reminders to post Team Chart, assess group behavior quickly, and deliver verbal praise) to assist in her implementation of CW-FIT. Ms. D. tallied points and gave behavior specific verbal praise when a group earned points and corrective feedback when groups failed to earn points. Ms. D. gave frequent reminders, or "pre-corrects," of behaviors she would be looking for from student groups. Pre-corrects were related to the behavioral skills taught in CW-FIT. Following the initial week of intervention, Ms. D. ran CW-FIT independently for the remainder of the study. At three times during the intervention phases of the study, the researcher met with Ms. D. to give feedback on her implementation of CW-FIT and review behavior data for the class and target students.

Materials

Tangible items were used as rewards for students who met the daily CW-FIT game goal (stickers, small trinkets, erasers, pencils, tablets, key chains, bracelets). These items ranged in price from less than $0.05 to $0.35. Candy had been identified as a potential reinforcer, but, as it was used daily by the teacher as a reward within the SWPBS system, the researcher and teacher decided to eliminate it as a reward in CW-FIT to lessen the possibility of satiation.

Three 11" × 17" laminated posters were displayed in the classroom throughout the intervention (not in baseline condition). Each poster was of a different color and listed one of the behavioral skills that had been taught to the students and examples of that skill. Icons were printed by each example in order to serve as a visual reference for students. An 11" × 17" white Team Chart for points was posted on the blackboard each day CW-FIT was conducted. The Team Chart included spaces for teacher name, daily point goal, and columns for team member names and point tallies. The three classroom skills were printed individually inside rows of the matrix to allow for an X to be marked by a skill if students had failed to display the skill at the time of assessment (i.e., at the end of each 1- to 4-min interval). Ms. D. marked points or Xs on the Team Chart using a magic marker in order to maximize visibility for all students in the class. A digital kitchen timer was used to alert Ms. D. to the passing of intervals for point delivery opportunities.

Dependent Variables

Target students' behaviors. On-task behavior and disruptive behavior were measured for the three identified target students using the *Multiple Option Observation System for Experimental Studies* (*MOOSES*) (Tapp, Wehby, & Ellis, 1995). Each MOOSES observation occurred for 10 min and allowed continuous collection of frequency of disruptive behavior and duration of on-task behavior. On-task behavior (coded as a duration in seconds using a toggle key for on or off) was defined as a student being within the instructional area, complying with instructions for academic tasks or other activities, and attending to appropriate materials or the teacher. The percent of on-task was calculated by dividing the number of seconds on-task by the total number of seconds for the observation. Disruptive behaviors (coded as frequencies by tapping the variable name on the screen for each occurrence) were divided into three categories: verbal disruptive, physical disruptive, and general disruptive. Verbal disruptive was any verbal statement towards others that was argumentative, taunt-

ing, name-calling, or provocative in nature, including verbal refusal to comply with directions. This category included talking out (without raising hands if required during class discussion) or talking to peers, laughing, and making noises during times not allowed by the teacher. Physical disruptions included any throwing of objects, hitting, kicking, pushing, punching, scratching, biting, spitting, pulling or other physical contact with a peer or adult. General disruptive was defined as any instance of getting out of seat, tipping or rocking chair, putting feet on desk, playing with materials inappropriately, and any inappropriate behavior that did not fall into the verbal or physical categories. Students could display disruptive behaviors while remaining on-task. For example, a student could be tipping back in their chair (a general disruptive behavior) while writing answers on a worksheet (on-task behavior).

Class-wide on-task behavior. Class-wide on-task data were collected during 20-min observations using paper-and-pencil measures and a 30-s momentary time sample procedure. On-task behavior was measured using the preceding definition but for groups of children rather than for individuals. To do so, the class was divided into groups of three to four students each, based on proximity. To score on-task, all students in the group needed to be on-task at the 30-s time point. The observer would look at group 1 and score a "plus" if all students were on-task or "minus" if one or more group members were off-task. Each target student served as a group of their own. In this manner, their behavior contributed to the overall class mean, however, the behavior of other children (as different groups) was able to be monitored to show improvements as well. Disruptive behaviors were not collected for all students in the class as the target students identified were the primary students exhibiting classroom problem behaviors. All groups remained constant throughout the study. The percent on-task was calculated for each group and averaged for the class-wide percentage on-task.

Teacher behavior. Teacher delivery of praise/points and reprimands to individuals and groups of students were measured using frequency counts. Observers tallied each occurrence of teacher praise or reprimand during the 20-min class-wide on-task observation. Praise/points was defined as any verbal statement, physical gesture, point tallied, or hole-punch given as social attention in response to appropriate behaviors exhibited by individuals or groups of students. Examples include "Great job following directions." "Super work during math class." "I like that everyone on Teams A and B are ready." or gestures such as pats on the back or thumbs up. Reprimands included any verbal statement or physical gesture, or loss of point during CW-FIT following inappropriate behaviors of individual students or groups.

Examples include "Everyone needs to go back to their seats right now!" "Stop talking, this is quiet work time!" "Raise your hand before speaking, don't shout out the answer." and "Teams need to remember to do your own work and ignore inappropriate behavior."

Observation Procedures

A minimum of two observers were present daily; one collected class-wide data, the other collected target student data using the *MOOSES* coding system. The sequence of target students' *MOOSES* observations was randomly selected each day, with the exception of irregular classroom scheduling (e.g., the student selected for the third observation was leaving early, so they would be moved up to the first observation).

Interobserver agreement (IOA). An additional observer recorded class on-task data simultaneously but independently in 29% of baseline observations and 24% of intervention observations, for a total of 10 observations. An additional observer collected *MOOSES* data during nine intervention observations for each target student, for a total of 21% of observations. Observers were trained to a criterion of 90% for two observations for both data collection systems (i.e., class-wide on-task and *MOOSES*).

During class on-task data collection, the primary and secondary observers were in close proximity to each other. The primary observer operated a digital timer and either gestured towards individual groups or called-out the letter assigned to the group for data collection purposes, in sequence, to the second observer at the end of each 30-s interval. Agreement percentages were calculated by comparing each observer's records on a trial-by-trial basis. An agreement was defined as both observers recording the same behavior (i.e., on-task or off task) for a group. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. IOA for class on-task data was 95% (range, 89% to 99%). Each observer independently coded the frequency of praise and reprimands (in boxes labeled such on the data collection sheet) from Ms. D. concurrently with the 20-min class-wide on-task observation. IOA for both praise/points and reprimands was calculated as total agreement by dividing the number of agreements by the total number of agreements and disagreements and multiplying by 100. IOA was 90% (range, 77% to 100%) for praise/points and 88% (range, 63% to 100%) for reprimands.

IOA for *MOOSES* data was calculated using the software's interobserver agreement file analysis in which event codes, within a

time window of 10 s, were analyzed between the primary and secondary observer's data files. IOA for disruptive behavior was 86% (range, 63% to 100%) and on-task behavior was 99% (range, 87% to 100%) across target students.

Procedural Fidelity of Implementation

Procedural fidelity data were taken in 25% of observations during the two intervention phases of the study and are reported as percent of items implemented. A fidelity of implementation checklist covered 16 aspects of the CW-FIT procedures (e.g., skill posters displayed, review of behavior skills at beginning of session, display of point chart, behavior specific praise for use of skills, timer used to signal points, points awarded for use of skills, winners announced for meeting goal and reward delivered). An additional 10 aspects of general classroom management and student behavior were evaluated (e.g., classroom materials ready, directions provided for assignments, level of compliance, transitions are short, teacher ignores minor inappropriate behavior, praise to reprimand ratio is 4:1). Observers recorded the presence or absence of these items at the end of each observation. Thus, the direct observation of the CW-FIT helped the observers score procedural fidelity more reliably.

The data indicated high procedural fidelity overall with a mean of 89% (range, 20% to 100%). Ms. D. consistently implemented the majority of the components of the CW-FIT game ($M = 91\%$; range, 20% to 100%). She displayed CW-FIT skill posters, awarded points for students displaying the skills, and delivered rewards to students who had met the daily goal with 100% fidelity. However, she gave verbal reminders of the behaviors students should display during the game in only 20% of opportunities.

Social Validity Measures

Students and teachers completed social validity questionnaires following the completion of the final CW-FIT condition. Students were given a 12-item survey that included statements related to the CW-FIT procedures (e.g., "It was easy to learn the rules of the game," "It was easy to play the game," "I liked playing the game in my classroom," "I liked earning points on a team," and "I liked earning prizes.") and what they learned from the intervention (e.g., "I learned the right way to get the teacher's attention," "I learned what to do when a classmate is acting inappropriately," "I followed directions more during the game," and "I raised my hand more during the game."). Students were asked to circle a smiley face if they agreed

with the statement, circle a face with a flat affect if they were indifferent, or circle a sad face if they disagreed with the statement. Due to the age and varied reading ability of the students, as well as their unfamiliarity with the measurement tool, each statement was read to the class for completion.

Ms. D. completed a 16-item questionnaire in the month following the final CW-FIT condition. This survey incorporated a 5-point Likert-type scale, as well as three open-ended questions to assess what the teacher did and did not like about CW-FIT. Statements were relative to the effectiveness of intervention, satisfaction with training and support, time required for CW-FIT implementation, and ease of implementation. The survey also included questions such as "To what extent did the students gain teacher attention appropriately during the game?" and "To what extent did the students follow directions during the game?" to assess the teacher's opinion of the extent to which students' behavior was affected by CW-FIT.

Experimental Design

A reversal (ABAB) design (Baer, Wolf, & Risley, 1968) was used to assess the effects of CW-FIT on class on-task behavior, frequency of teacher praise/points, frequency of teacher reprimands, frequency of disruptive behaviors of target students, and on-task behavior of target students. Baseline data were collected for 7 days. During 3 days of the initial baseline, two class-wide on-task and two *MOOSES* observations per target student were conducted. This was done to ensure enough baseline data were collected to show stable trends without delaying the start of intervention. During both baseline phases, Ms. D. reinforced appropriate student behavior with hole-punches (the universal SWPBS method in the school) that could be exchanged for candy at the end of each day. Data collection was stopped during behavioral skill instruction (approximately 2 weeks). Following completion of skill instruction, the CW-FIT game was introduced and remained in place for 18 school days. A 3-day reversal, during which CW-FIT was not implemented, was followed by a reintroduction of CW-FIT for 6 days. Due to absences, the number of observations in baseline and intervention varied for target students.

Results

Overall, CW-FIT resulted in a reduction in disruptive behavior for each of the three target students and an increase in on-task behavior for the target students and the class. In addition, increased levels

of teacher praise/points to groups of students were observed when the intervention was in place.

Target Students' On-Task and Disruptive Behaviors

Disruption data are depicted in Figure 1. Verbal, physical, and general disruptions have been combined for total disruption frequency. During the two baseline conditions, Lily had a mean of 21–23 disruptions, with two low points. The number of disruptions immediately dropped well below baseline levels and became stable when CW-FIT was implemented ($M = 7$; 4, respectively). Albert's data reveal trends much like those in Lily's data. Albert's disruptions, before CW-FIT, averaged 28 and were quite variable during the initial baseline. Mean disruptions dropped dramatically when intervention was introduced, to a mean of 4; increased during the return to baseline and dropped again during the final CW-FIT condition, but with an overall increasing trend in the reintroduction of CW-FIT. During baseline conditions, Faith's disruptive behavior averaged 16 to 18, with a downward trend in the initial baseline. During CW-FIT conditions, Faith's disruptive behavior dropped to averages of 2 and 3, respectively.

Figure 2 represents on-task behavior for each target student. Preceding CW-FIT, Lily's on-task was fairly high during baseline averaging 77%, with a slight increase and high variability during the initial CW-FIT condition. On-task behavior decreased during baseline 2 and stabilized at a mean of 99% during the final intervention condition. Albert's on-task behavior was variable during initial baseline with a mean of 61%, and increased during the CW-FIT. His on-task showed only a slight decrease with the return to baseline, and remained high during the final implementation of CW-FIT. Variability characterized Faith's on-task behavior during baseline, similarly to Albert. Her mean on-task behavior was 68% and 76% during the two baseline conditions, with an increase during the two CW-FIT phases (means, 97% and 94%, respectively). Data for both Albert and Faith showed higher stability during CW-FIT conditions over baseline.

Class-Wide On-Task Behavior

An effect was also seen in class-wide on-task behavior when CW-FIT was implemented (Figure 3). During baseline, on-task behavior for the class (including target students) was variable, with an overall increasing trend and a mean of 57% (range 40% to 82%). Although variability continued following the implementation of CW-FIT, the average on-task behavior rose ($M = 87\%$). The removal of intervention resulted in a drop in class-wide on-task behavior to a mean of 61%. The reintroduction of CW-FIT resulted in a mean of 84%.

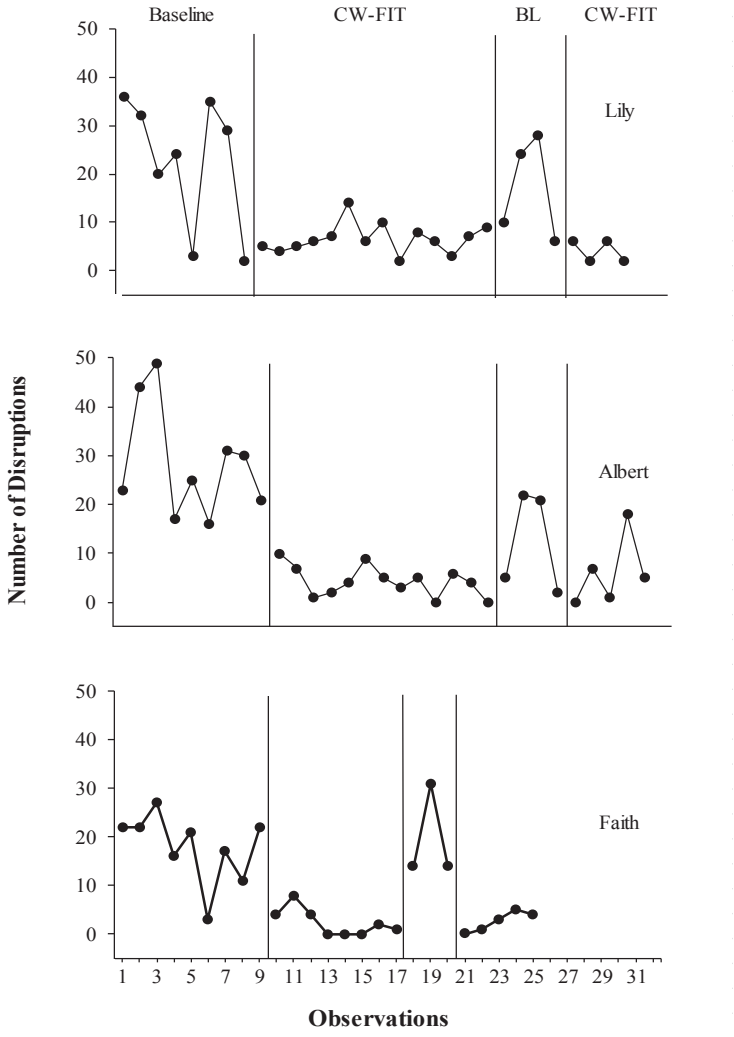

Figure 1. Number of Disruptive Behaviors for Lily, Albert, and Faith per 10-minute Observation

Teacher Praise and Reprimand Behavior

Table 1 shows the number of praise and reprimands delivered by Ms. D. to the target students during 10-min observations. In general, limited differences were noted for praise or reprimands across conditions for the target students. Praise was low overall and averaged 0 to 3 across baseline and from 0 to 4 during CW-FIT conditions and across

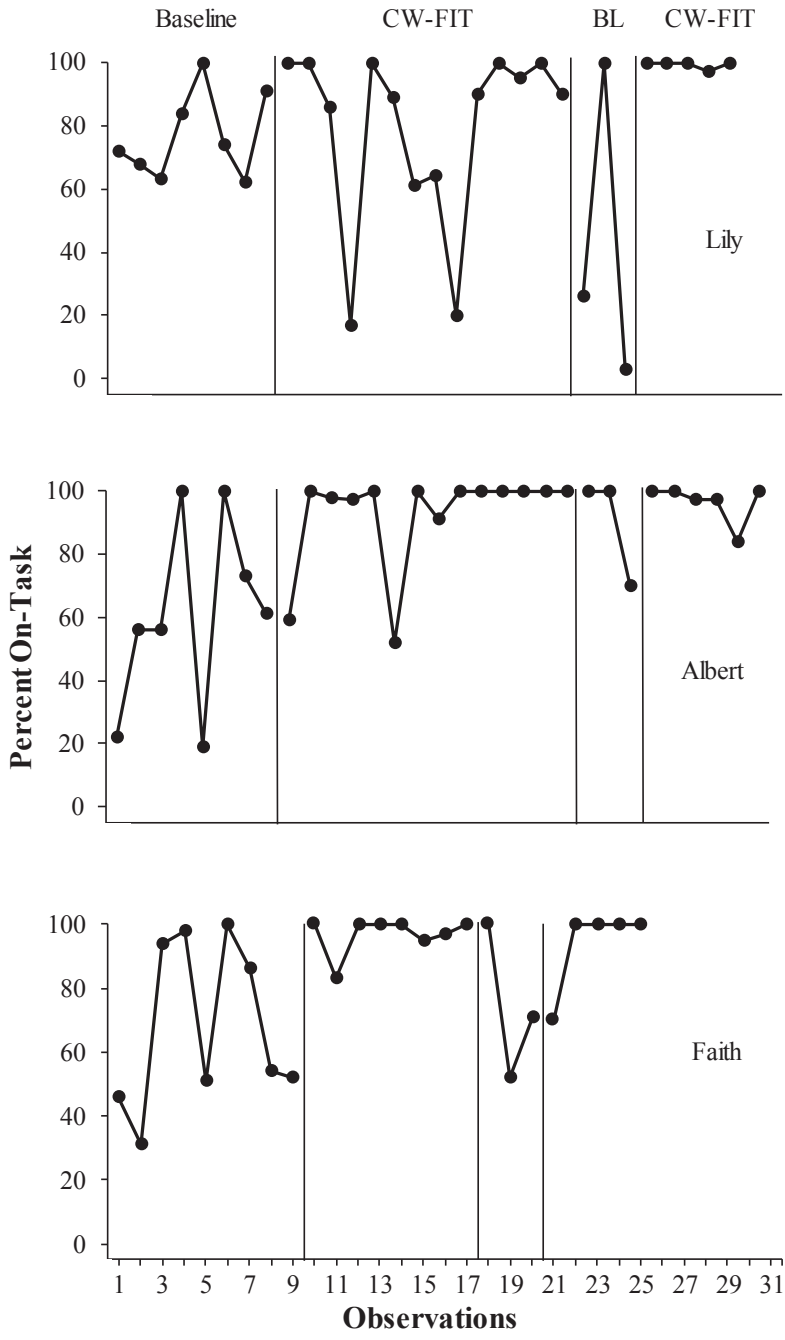

Figure 2. Percent On-Task Behavior for Students per 10-minute Observation

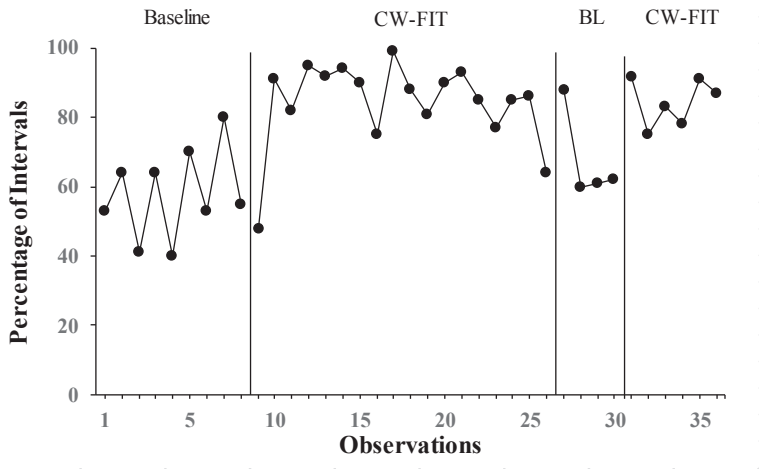

Figure 3. Percent of Intervals with On-Task Behavior for the Class per 20-minute Observation

Table 1
Mean and Range for Praise and Reprimands across Students and Conditions

Student	Baseline	CW-FIT	Baseline 2	CW-FIT 2
Lily				
Praise	1.0	1.4	.7	.8
	0–2	0–4	0–2	0–3
Reprimand	1.6	.9	2.0	1.6
	0–8	0–3	1–3	0–4
Albert				
Praise	.4	.3	2.0	.5
	0–3	0–1	1–3	0–1
Reprimand	1.7	0	1.0	.7
	0–4		0–3	0–2
Faith				
Praise	.7	1.1	.3	.8
	0–2	0–4	0–1	0–1
Reprimand	.7	.4	1.0	.4
	0–2	0–1	0–2	0–1

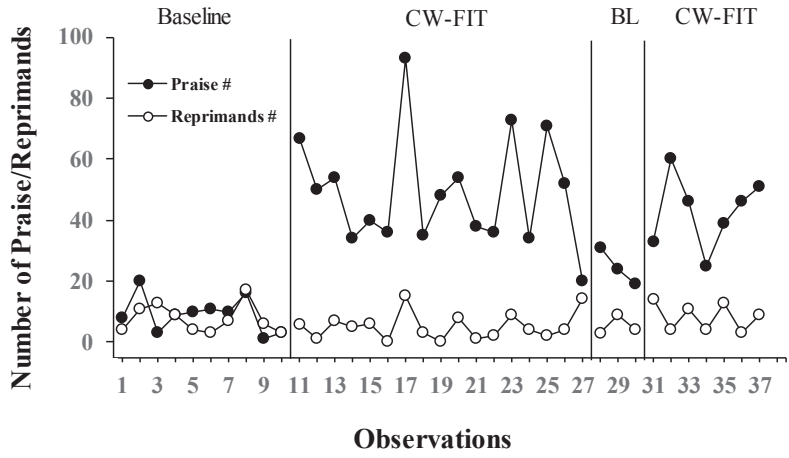

Figure 4. Number of Praise Including Points and Reprimands to the Class

students. Reprimands were also low averaging 0 to 8 during baseline with slight decreases and averages ranging from 0 to 4 during CW-FIT conditions.

Figure 4 depicts the number of praise/points and reprimands to the class. The level of praise/points was greatly increased when CW-FIT was in place; however, the level of reprimands remained stable during each condition of the study. During baseline, Ms. D.'s mean number of praise/points was 10, with an increase to 48 during CW-FIT, a reduction to 25 during baseline 2 and an increase to a mean of 45 during the final intervention condition. Reprimands averaged 5 to 9 across conditions.

Social Validity

Ten students completed the student social validity questionnaire regarding CW-FIT. Overall, students responded very favorably to all questions. The majority of students liked using CW-FIT and found it easy to learn and play. In addition, 90% of students reported they would tell friends that CW-FIT was fun to use. When questioned more specifically about reinforcement components, all students responded that they liked earning points and prizes. Nearly 100% of students replied they had learned the three skills taught in CW-FIT and displayed those skills more when CW-FIT was used in the classroom.

Ms. D. gave favorable ratings regarding CW-FIT to the majority of questions in the teacher questionnaire with a mean rating of 4.7. She

found the training and support she received for CW-FIT to be highly acceptable. Ms D. reported that CW-FIT was easy to implement and required very little time for preparation. In addition, she reported that implementing CW-FIT consumed very little instruction time. When questioned about the extent to which students displayed the skills taught in CW-FIT, Ms D. replied that students very frequently gained teacher attention appropriately and followed directions when CW-FIT was implemented. The lowest rating (2) was that students ignored inappropriate behavior of peers only a little. Ms. D. reported that she liked CW-FIT very much and would be very likely to use it in the future, and would also recommend it to a colleague.

Discussion

The present study demonstrated the effects of the CW-FIT intervention on improving the on-task behavior of a first grade class, as well as decreasing the disruptive behavior of three individual students within that class. These findings are similar to those obtained by others using group contingencies in primary grade classes (e.g., Davies & Witte, 2000; Kelshaw-Levering et al., 2000); and in prior CW-FIT studies (Kamps et al., 2010; Wills et al., 2014). Findings advance the literature on CW-FIT by demonstrating its effects with ELLs served in an urban, culturally diverse and economically impoverished community. This supports a prior CW-FIT study in which students who were ELLs participated but with measures collected without experimental demonstration for individual students (Kamps et al., 2010).

On-task behavior for the three target students increased when CW-FIT was implemented. Each student had highly variable on-task behavior during baseline, yet each responded immediately with improvements in behavior to the implementation of CW-FIT. Albert and Faith both displayed very high levels of on-task behavior during CW-FIT, with relatively more stable data paths when compared to those during baseline. Although the level of Lily's on-task behavior rose when CW-FIT was implemented, the mean was virtually unchanged (from 77% to 79%) and her on-task behavior dropped well below baseline during two of the observations. These data do not reveal what variables may have affected her behavior; however, we observed anecdotally that on these days Lily had frequent interactions with a classmate who displayed high rates of inappropriate behavior which seemed to affect her behavior. During the final implementation of CW-FIT, however, Lily was on-task in 100% of intervals on 4 out of 5 days.

CW-FIT also resulted in a considerable decrease in each target student's disruptive behavior. The frequency of these students' disruptions preceding intervention was at a level unacceptable to the classroom teacher and had contributed to the identification of Lily and Albert as at risk for behavioral disorders. Lily showed the greatest overall decrease in disruptive behavior during CW-FIT. Albert's and Faith's mean number of disruptive behaviors was also reduced considerably during intervention phases.

The similar effect of CW-FIT on both class-wide on-task behavior and target students' behavior shows the group contingency was effective with the large group of students and with high-risk individual students. Many group contingency studies targeting non-academic behaviors have shown the effect of an intervention on the group as a whole, but not on individuals within that group (e.g., Babyak, Luze, & Kamps, 2000; Harris & Sherman, 1973; Kelshaw-Levering et al., 2000; Lohrman & Talerico, 2004); whereas other studies have reported the effect of a group contingency on individuals only (e.g., Reitman, Murphy, Hupp, & O'Callaghan, 2004). In the current study, the effect of CW-FIT was shown for the large group (i.e., the class) and for individual target students. A single intervention (i.e., CW-FIT) was able to concurrently affect the behavior of more highly disruptive and off-task students and the class as a whole, and in this respect, it adds to the literature on group contingencies.

The SWPBS incentive program (i.e., hole punches exchangeable for rewards) used by Ms. D. in her classroom before CW-FIT was introduced did not result in a behavior change to levels that were acceptable to the teacher. It is unclear why these supports were ineffective; however, several possibilities exist. The SWPBS literature estimates 80% of students will be supported by primary level behavior supports in a school-wide PBS system (Sugai & Horner, 2002). It is possible that primary supports may have failed to reduce or maintain behavior at acceptable levels due to the number of high-risk students in the class. PBS literature reports 15% of students will need a secondary level of support (Sugai & Horner, 2002). However, in areas where increased risk factors (e.g., poverty, minority status, ELL) affect a majority of students, such as for students in this study, this percentage may be larger (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006; Horner et al., 2010; Wills et al., 2010). Although three students were chosen as target students because of their frequent disruptive behavior, anecdotally we observed high rates of problem behavior from several other students in the class. The higher level of reinforcement in CW-FIT compared to that of the primary school supports may have resulted in the increased on-task behavior of the class and target students.

Another possible explanation for the low levels of on-task behavior and high levels of disruptive behavior before intervention may be the frequency with which Ms. D. delivered praise. Flora (2000) and others (Sutherland, Wehby, & Yoder, 2002) recommended a minimum ratio of teacher praise to reprimands of 4–5:1. Ms. D.'s ratio of praise to reprimands to the class during initial baseline was approximately 1:1. During intervention phases, Ms. D.'s praise (and delivery of points) to reprimand ratio rose to 7.5:1. The increased ratio of praise to reprimand during intervention phases could have affected student behavior. However, it is more probable that the contingent nature of the praise during CW-FIT contributed to the increase of on-task behavior and decrease of disruptive behavior. During CW-FIT, Ms. D. delivered praise/points contingent on students' display of specific skills. The establishment of this contingency may have led to the subsequent increase in on-task behavior.

The levels of individual praise and reprimands delivered to Lily, Albert, and Faith were low in each condition of this study, although praise/points to groups increased substantially during intervention phases of this study. In an interdependent group contingency such as CW-FIT, the reinforcement of the behavior of the group is dependent on the behavior of the individuals within that group. Consequently, both praise and reprimands may be directed at the group rather than at an individual. As such, it is possible that Ms. D's delivery of individual praise and reprimands to the target students remained relatively low because it was delivered to the groups rather than to individual students (e.g., "Nice job following directions, teams!").

The positive effects of CW-FIT are further supported by Ms. D.'s and students' high satisfaction. Students enjoyed using CW-FIT and felt that it improved their behavior in class. Skinner et al. (1999) reported that students may complain when peers are receiving rewards if they are not. The implementation of a group contingency like CW-FIT in which all students work for the same reward may reduce this problem associated with individualized protocols. Elliot, Witt, Galvin, and Peterson (1984) reported that teachers prefer interventions that are simple and that require few resources and time to implement. Teacher social validity ratings in this study support those findings. Ms. D. reported that CW-FIT took very little time away from instruction and was easy to learn and implement. The managerial problems associated with individual contingencies and behavior plans make them less preferred by teachers than group contingencies (Skinner et al., 1999). In this study, the behavior of target students who exhibited high rates of problem behavior was positively affected by the intervention,

as was the overall on-task behavior of the class. These effects may have forestalled the need for individualized behavior plans for the target students, thereby reducing the teacher's burden and increasing the social validity of the intervention.

Limitations and Future Research

A limitation to the current study is the small number of participants and that the occurrence of the behavioral skills taught in CW-FIT (i.e., gaining teacher attention appropriately, following directions the first time, and ignoring inappropriate behavior of peers) was not measured. As such, we were unable to analyze the extent to which the systematic instruction resulted in the occurrence of these behaviors. Future studies may choose to measure those responses in order to provide a more comprehensive analysis of the effects of the intervention on student behavior. However, teaching the broad skill such as "Following Directions" appeared to increase general on-task behavior and decrease disruptive behaviors (talking out, being out of seat).

A further limitation of this study is that we were unable to evaluate maintenance of on-task and disruptive behavior under leaner schedules of reinforcement, due to the end of the school year. Future research might attempt to thin the schedule of reinforcement in the CW-FIT implementation and assess the maintenance of target behaviors. A leaner schedule of reinforcement would reduce both the costs associated with rewards and the time teachers would have to devote to the delivery of points, thus lessening the requirements of the teacher and potentially increasing the social validity of the intervention.

An additional limitation is that the teacher's praise and reprimands to individual students remained unchanged across experimental conditions. Future research might address enhancements to CW-FIT and other group contingencies to increase specific praise to individual students. Studies are also needed to develop additional intervention for teachers regarding understanding of the functional relation of reprimands as social attention and maintaining inappropriate behaviors in classroom settings. In this study, reprimands continued at similar levels to praise for the target students even with the CW-FIT condition. Also, the teacher's fidelity was good overall at 89%; however, she scored low (20%) on providing verbal reminders of the behaviors required to earn points during the game. This suggests that the use of "precorrects" or error correction related to specific skills is not as critical a component in reducing disruptive behaviors/increasing on-task perhaps as consistent use of points to give attention to the actual occurrence of the behaviors.

Two final limitations, trends in the disruptive data (e.g., decreasing trend in baseline for Faith, and slight increasing trend in the second CW-FIT phase for Albert) affect the degree of experimental control. Also, IOA were only collected during intervention with the MOOSES and not during baseline, and the range for disruptive behavior agreement was low (63% to 100%). Overall, however, 86% IOA for disruptive and 100% for on-task MOOSES suggest the viability of these measures.

An extension of this research that would contribute to the generality of the CW-FIT intervention is a large-scale follow-up study that could evaluate the effect of CW-FIT over time for individual students, in multiple schools and classrooms with comparison to a control group of high risk students who did not receive the intervention. Evidence that the CW-FIT program could reduce and prevent problem behavior and increase desirable behavior of students across multiple environments would be beneficial to schools seeking to incorporate evidence-based practices in their discipline policies and SWPBS procedures.

Practical Implications

Practical implications for the study are that the use of the CW-FIT group contingency, an easy to implement intervention, can affect the behavior of both groups of students and individuals within classrooms. Although a small number of studies have reported effects for both (Kamps et al., 2010), this has not been frequently demonstrated in recent group contingency literature. CW-FIT resulted in decreased disruptions for high-risk students and increased on-task behavior for those students and their classmates. This factor may be significant for schools and teachers seeking to address the needs of a large number of students with challenging behaviors. It is much more efficient to implement a class-wide intervention that will improve behaviors for individuals with disruptive behaviors than it is to develop multiple individualized behavior plans. The study also demonstrates that CW-FIT may be a practical and effective intervention to support classroom level implementation of SWPBS in urban, diverse schools. As calls for preventative approaches to discipline in schools continue to rise (www.pbis.org/school/tier1supports), future research may be directed towards programs that incorporate practices that address the needs of the greatest number of students. The findings of the current study may be beneficial to teachers and schools as they search for applications that are simple, effective, and evidence-based (Briesch & Chafouleas, 2009; Maggin et al., 2012; Stage & Quiroz, 1997).

References

Allyon, T., & Azrin, N. (1968). *The token economy: A motivational system for therapy rehabilitation.* New York: Meredith Corporation.

Anderson, C., Rodriguez, B., & Campbell, A. (2015). Functional behavior assessment in schools: Current status and future directions. *Journal of Behavioral Education, 24*, 338–371.

Babyak, A. E., Luze, G. J., & Kamps, D. M. (2000). The good student game: Behavior management for diverse classrooms. *Intervention in School and Clinic, 35*, 216–223.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91–97.

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*, 119–124.

Briesch, A., & Chafouleas, S. (2009). A review and analysis of the literature on self-management interventions to promote appropriate classroom behaviors (1988–2008). *School Psychology Quarterly, 24*, 106–118.

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.

Catania, A. C. (1998). *Learning* (4th Ed.). New Jersey: Prentice Hall, Inc.

Chafouleas, S. M., Volpe, R. J., Gresham, F. M., & Cook, C. R. (2010). School-based behavioral assessment within problem-solving models: Current status and future directions. *School Psychology Review, 39*, 343–349.

Conklin, C., Kamps, D., & Wills, H. (2017). The effects of Class-Wide Function-related Intervention Teams (CW-FIT) on students' prosocial classroom behaviors. *Journal of Behavioral Education, 26*, 75–100.

Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *School Psychology Quarterly, 22,* 540–556.

Davies, S., & Witte, R. (2000). Self-management and peer-monitoring within a group contingency to decrease uncontrolled verbalizations of children with attention-deficit/hyperactivity disorder. *Psychology in the Schools, 37*, 135–147.

Elliot, S. N., Witt, J. C., Galvin, G. A. & Peterson, R. (1984). Acceptability of positive and reductive behavioral interventions: Factors that influence teachers' decisions. *Journal of School Psychology, 22*, 353–360.

Embry, D. D. (2002). The good behavior game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*, 273–297.

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). *Reducing behavior problems in the elementary school classroom. An IES practice guide*. Washington, DC: Institute of Education Sciences.

Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, A. L. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review, 30*, 193–210.

Flora, S. R. (2000). Praise's magic reinforcement ratio: Five to one gets the job done. *The Behavior Analyst Today, 1*, 64–69.

Francis, D., Rivera, M., Lesaux, N. K., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions*. Available from http://www.centeroninstruction.org/files/ELL1-Interventions.pdf

Harris, V. W., & Sherman, J. A. (1973). Use and analysis of the "good behavior game" to reduce disruptive classroom behavior. *Journal of Applied Behavior Analysis, 6*, 405–417.

Hart, B., & Risley, T. (1995). *Meaningful Differences in the Everyday Experiences of Young American Children*. Baltimore: Brookes Publishing, Co.

Horner, R., Sugai, G., & Anderson, C. (2010). Examining the evidence for school-wide positive behavior support. *Focus on Exceptional Children, 42,* 1–14.

Jensen, E. (2009). *Teaching with Poverty in Mind: What Being Poor Does to Kids Brains and What Schools Can Do About It*. Alexandria, VA: ASCD.

Jolstead, K. A., Caldarella, P., Hansen, B. D., Korth, B. B., Williams, L. & Kamps, D. (2017). Implementing positive behavior support

in preschools: An exploratory study of CW-FIT tier 1. *Journal of Positive Behavior Interventions*, *19*, 48–60.

Kaminski, R., & Good, R. H., III (1998). Assessing early literacy skills in a problem-solving model: Dynamic indicators of basic early literacy skills. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement.* New York: Guildford.

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.

Kamps, D., Greenwood, C., Arreaga-Mayer, C., Veerkamp, M., Utley, C., Tapia, Y., Bowman-Perrott, L., & Bannister, H. (2008). The efficacy of ClassWide Peer Tutoring in middle schools. *Education and Treatment of Children, 31*, 1–34.

Kamps, D., Wills, H., 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 Intervention, 17,* 134–145.

Kamps, D., Wills, H., Heitzman-Powell, L., Laylin, J., Szoke, C., Hobohm, T., & Culey, A. (2010). *Class-wide F*unction-related *I*ntervention *T*eams: Effects of group contingency programs in urban classrooms. *Journal of Positive Behavior Interventions, 13*, 154–167.

Kelshaw-Levering, K., Sterling-Turner, H. E., Henry, J. R., & Skinner, C. H. (2000). Randomized interdependent group contingencies: group reinforcement with a twist. *Psychology in the Schools, 37*, 523–533.

Klinger, J., & Soltero-Gonzalez, L. (2009). Culturally and linguistically responsive literacy instruction for English language learners with learning disabilities. *Multiple Voices for Ethnically Diverse Exceptional Learners, 12*, 4–20.

Lannie, A. L., & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the Good Behavior Game on student and teacher behavior. *Education and Treatment of Children, 30*, 85–98.

Lea, T., Bray, M., Kehle, T., & DioGuardi, R. (2004). Contemporary review of group-oriented contingencies for disruptive behavior. *Journal of Applied School Psychology*, 20, 79–101.

Lohrmann, S., & Talerico, J. (2004). Anchor the boat. *Journal of Positive Behavior Interventions, 6*, 113–120.

Maggin, D., Chafouleas, S., Goddard, K., & Johnson, A. (2011). A systematic evaluation of token economies as a classroom man-

agement tool for students with challenging behavior. *Journal of School Psychology, 49*, 529–554.

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology, 50*, 625–654.

McGinnis, E., (2010). *Skillstreaming the Elementary School Child: A Guide for Teaching Prosocial Skills.* Champaign, IL: Research Press.

Mitchem, K. J., Young, K. R., West, R., P., & Benyo, J. (2001). CWPASM: A class-wide peer-assisted self-management program for general education classrooms. *Education and Treatment of Children, 24*, 111–140.

Packard, R. G. (1970). The control of "classroom attention": A group contingency for complex behavior. *Journal of Applied Behavior Analysis, 3*, 13–28.

Popkin, J., & Skinner, C. H. (2003). Enhancing academic performance in a classroom serving students with serious emotional disturbance: Interdependent group contingencies with randomly selected components. *School Psychology Review, 32*, 271–284.

Positive Behavioral Interventions & Supports. *Tier 1 Supports*. (2017). OSEP Technical Assistance Center. Retrieved from: www.pbis.org/school/tier1supports

Reavis, K., Jenson, W., Kukic, S., & Morgan, D. (1988). Utah's *Best Project: Behavioral and Educational Strategies for Teachers. Technical Assistance Manuals*. Salt Lake City, UT: Utah State Office of Education.

Reitman, D., Murphy, M. A., Hupp, S. D. A., & O'Callaghan, P. M. (2004). Behavior change and perceptions of change: Evaluating the effectiveness of a token economy. *Child & Family Behavior Therapy, 26*, 17–36.

Rose, L. C., & Gallup, A. M. (2005). The 37th annual Phi Delta Kappa/gallup poll of the public's attitudes toward the public schools. *Kappan, 87*, 41–57. Salend, S. J., Reynolds, C. J., & Coyle, E. M. (1989). Individualizing the good behavior game across type and frequency of behavior with emotionally disturbed adolescents. *Behavior Modification, 13*, 108–126.

Sheridan, S. (2010). *The Tough Kid Social Skills Book*. Champaign, IL: Research Press.

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 for Children, 31*, 351–380. doi:10.1353/etc.0.0007

Skinner, C. H., Skinner, C. F., Skinner, A. L., & Cashwell, T. H. (1999). Using interdependent contingencies with groups of students: Why the principal kissed the pig. *Educational Administration Quarterly, 35*, 806–820.

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *The School Psychology Review, 26*, 333–368.

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

Sutherland, K., Wehby, J., & Yoder, P. (2002). Examination of the relationship between teacher praise and opportunities for students with EBD to respond to academic requests. *Journal of Emotional and Behavioral Disorders, 10*, 5–13.

Tapp, J., Wehby, J. H., & Ellis, D. N. (1995). A multiple option observation system for experimental studies: M.O.O.S.E.S. *Behavior Research Methods, Instruments & Computers.*

Thorne, S., & Kamps, D. (2008). The effects of a group contingency intervention on academic engagement and problem behavior of at-risk students. *Behavior Analysis in Practice, 1*, 12–18.

Tingstrom, D. H., Sterling- Turner, H. E., & Wilczynski, S. M. (2006). The good behavior game: 1969–2002. *Behavior Modification, 30,* 225–253.

Walker, H. M., Forness, S. R., Kauffman, J. M., Epstein, M. H., Gresham, F. M., Nelson, C. M., & Strain, P. S. (1998). Macro-social validation: Referencing outcomes in behavioral disorders to societal issues and problems. *Behavioral Disorders, 24*, 7–18.

Walker, H. M., & Severson, H. H. (1992). Systematic screening for behavior Disorders: SSBD. Longmont, CO: Sopris West.

Weeden, M., Wills, H., Kamps, D., & Kottwitz, E. (2016). The effects of the Class-Wide Function-related Intervention Team (CW-FIT) program on the on-task behavior of children with Emotional Behavioral Disorders. *Behavior Disorders*, *42*, 285–293.

Wills, H., Iwaszuk, W., & Kamps, D., & Shumate, E (2014). CW-FIT: Group contingency effects across the day. *Education and Treatment of Children, 37*, 191–210.

Wills, H. P., Kamps, D., Hansen, B. D., Conklin, C., Bellinger, S., Neaderhiser, J. & Nsubuga, B. (2010). The Class-wide Function-based Intervention Team (CW-FIT) Program. *Preventing School Failure. 54,* 164–171.