J Behav Educ (2017) 26:75–100
DOI 10.1007/s10864-016-9252-5

ORIGINAL PAPER

The Effects of Class-Wide Function-Related Intervention Teams (CW-FIT) on Students' Prosocial Classroom Behaviors

Carl G. Conklin[1] · **Debra Kamps**[2] · **Howard Wills**[2]

Published online: 21 July 2016

Abstract Students with challenging, disruptive behavior have difficulty learning in school, and their behavior adversely impacts the learning of other students and the classroom teacher. Class-Wide Function-related Intervention Teams (CW-FIT) is an evidence-based approach that teachers can use to prevent and reduce problem behavior and increase prosocial classroom behaviors. Previous studies have demonstrated that CW-FIT produced improvements in student appropriate classroom behaviors which led to increased available instruction time. The purpose of this investigation was to systematically replicate CW-FIT adding to the empirical research base supporting it. A novel aspect compared to prior studies was measurement of the student behaviors related to skills taught during CW-FIT (compliance, hand-raising, out-of-seat, and talking out), showing a direct relationship to students' improvements. Students in four classes and their teachers participated in this study. An ABAB reversal design was used to demonstrate intervention effectiveness and experimental control. Implications for research and practice are discussed.

Keywords Classroom management for students with behavioral risks · Prosocial skills

✉ Debra Kamps
dkamps@ku.edu

[1] Behavioral Education Services and Training, 6224 Cloud Dance Dr, 88012 Las Cruces, NM, USA

[2] Juniper Gardens Children's Project, University of Kansas, 444 Minnesota Avenue, 3rd Floor, Kansas City, KS 66101, USA

Introduction

Research suggests that behavior problems and, by extension, discipline problems are more frequent in urban public schools where the majority of low-income and ethnic/minority students are educated (Greenwood et al. 2008; Jacob2007; see National Center on Educational Statistics, http://nces.ed.gov/surveys/urbaned/environment.asp). Teachers rank disruptive behaviors and conduct problems in the classroom as a major barrier to teaching their students (Harrison et al. 2012). However, effective management of the urban classroom is a skill in which teachers are often ill prepared (Graziano 2005). Additionally, when these skills are taught, repeated follow-up is often required for improvement (Sutherland et al. 2000). Therefore, it is critical to provide teachers with classroom management strategies that have not only demonstrated results such as increased on-task behavior and decreased disruptions, but are also time efficient and easy to implement (Hawken and Horner 2003; McComas et al. 2002; Skinner et al. 1996).

One classroom management strategy with substantial evidence and replications supporting its effectiveness is group-oriented contingencies (Litrow and Pumroy 1975; Stage and Quiroz 1997). There are three types of group contingencies: (1) independent method in which students earn incentives based on their individual behavior, (2) the dependent method when a group or team of students earns incentives based on the behavior of one student or a small subgroup within the team, and (3) interdependent method that requires all team members to behave in a specified way in order to reach a criterion to earn incentives (Tingstrom et al. 2006). Although evidence for the effectiveness of group-oriented contingencies has been widely demonstrated (Embry 2002; Maggin et al. 2012; Stage and Quiroz 1997), the majority of this evidence has been in support of the effects of interdependent group-oriented contingencies on negative social behaviors (i.e., disruptive behavior). Perhaps the most widely examined educational intervention using this contingency is the Good Behavior Game (GBG; Barrish et al. 1969). The Good Behavior Game procedures consist of (1) defining rules for the classroom, (2) use of a group contingency with class teams and "fouls" for rule infractions, (3) winning the game by having fewest fouls or below a criterion for multiple winning teams, and (4) brief activity prizes and posted scoreboards (Embry 2002). Tingstrom et al. (2006) provided a review of the GBG with variations and adaptations to original procedures conducted between the years of 1969 and 2002. Of the 28 total studies, 18 were implemented to diminish disruptive behavior (e.g., Davies and Witte 2000; Gresham and Gresham 1982), eight to increase academic behavior and diminish disruptive behavior (Darch and Thorpe 1977; Robertshaw and Hiebert 1973), one to increase prosocial behaviors (Patrick et al. 1998), and one to increase oral hygiene (Swain et al. 1982).

Although the GBG is a specific interdependent group-oriented contingency intervention, other group-oriented contingency programs have also been implemented to diminish disruptive behavior (e.g., Coogan et al. 2007; Salend and Lamb 1989; Schanding and Sterling-Turner 2010; Sharp and Skinner 2004; Thomas et al. 1987); increase academic performance (e.g., Lew et al. 1986; Lloyd et al. 1996;

Popkin and Skinner 2003; Trevino-Maack et al. 2014); increase prosocial behavior (Cashwell et al. 2001; Lewis et al. 2002); and increase on-task behavior (Babyak et al. 2000; Christ and Christ 2006; Williamson et al. 2009). Many studies using interdependent group contingencies as a primary component of intervention were conducted in the school environment (Kamps et al. 2010; Kelshaw-Levering et al. 2000; Theodore et al. 2001; Popkin and Skinner 2003).

Class-Wide Function-Related Intervention Teams (CW-FIT)

Class-Wide Function-related Intervention Teams (CW-FIT) is a class-wide intervention similar to the GBG (Barrish et al. 1969) in that it includes an interdependent group-oriented contingency component. However, CW-FIT differs in several important ways from the GBG. Whereas the GBG's primary focus has been on disruptive, negative behaviors, CW-FIT's primary component is focused on directly teaching and reinforcing student prosocial classroom behaviors via the group contingency within the classroom environment. Teaching these skills takes advantage of what is known about the common behavior functions as they relate to promoting behavior problems in the classroom (see review of functional assessment in school settings by Ervin et al. 2001; Hanley et al. 2003). These skills address attention—"How to get the teacher's attention," escape—requesting assistance or a break, and recruiting reinforcement. Finally, the modification of the classroom environment (e.g., creating working groups of students) and teaching and reinforcing appropriate prosocial classroom behaviors are two components with the strongest level of evidence for reducing problem behaviors in the classroom as recommended by the Institute of Education Sciences (Epstein et al. 2008).

CW-FIT research has explored the effects on increasing student on-task behavior, diminishing disruptive behavior, and increasing teacher praise primarily in elementary schools (Kamps et al. 2010). Kamps et al. (2010) reported that students' on-task behavior in six separate classes (kindergarten, first, fourth, and fifth grades) increased from a baseline mean of 43.6 % occurrence to an intervention mean of 79.7 % (+36 %). Additionally, the disruptive behaviors of eight target students decreased from a mean of 18.2 disruptive behaviors during baseline to 5.7 during CW-FIT intervention. These results are thought to have occurred because CW-FIT addresses the function of inappropriate behaviors at the class-wide level addressing the potential additional social consequences/attention when interdependent group-oriented contingencies are applied. For example, students are taught at the class-wide level to raise their hands as an appropriate means of getting teacher attention, rather than calling out. This study did not, however, measure students' use of the CW-FIT skills across conditions.

Wills et al. (2010) reported improvements in class-wide on-task behavior in 16 urban classrooms in three elementary schools. Descriptive data for individual high-risk students reported decreases in disruptive behaviors; however, these differences were not experimentally validated using reversal or multiple baseline designs, nor was use of CW-FIT skills measured. Several single-case design studies have been published showing the effects of CW-FIT for (a) improving class-wide on-task

behavior across multiple class periods in a first-grade class (Wills et al. 2014); (b) improving class-wide on-task and individual students' on-task behavior and decreasing disruptive behavior using self-management and help cards in a first-grade class (Kamps et al. 2015); and (c) improving on-task for students in a special education setting (Weeden et al., in press). Recently, a randomized controlled trial demonstrated significantly increased class-wide on-task behavior (a 31 % increase from baseline to intervention) and increased teacher praise following implementation of CW-FIT (Kamps et al. 2015). Although positive, limitations of the studies are that the use of CW-FIT skills was not measured across conditions. Therefore, improved behaviors cannot be directly linked to increased use of the CW-FIT skills. In addition, the oldest students in the prior studies were in fifth grade, limiting generalizability of the program to middle school grades.

Purpose

The purpose of the current study was to systematically replicate the prior studies of CW-FIT (Kamps et al. 2010; Wills et al. 2010) and effects for on-task behavior and teacher praise, and address several novel components and new research questions using a rigorous research design. This replication study investigated the generality of the CW-FIT effects on a wider range of behaviors, students, teachers, settings, and grade levels. A novel addition to the study design included measurement of the behaviors being taught within the protocol of CW-FIT (e.g., compliance, hand-raising, out-of-seat, and talking out). Prior studies had broadly measured "on-task" and "disruptive" behaviors. In addition, CW-FIT was implemented across a wide age range, kindergarten, second-grade, and two seventh-grade classes. Previous implementation was limited to elementary schools. Specifically, the research questions guiding this study were:

1. To what extent will the CW-FIT intervention have an effect on increasing students' appropriate classroom behaviors (i.e., on-task, compliance, hand-raising) and reducing inappropriate behaviors (i.e., talking out, out-of-seat)?
2. To what extent will the CW-FIT intervention have an effect on increasing teacher praise statements?
3. Are CW-FIT outcomes similar in middle school age classes to elementary aged classes?

Methods

Participants and Settings

Students

Four classes participated in the study. The kindergarten classroom served 22 students for the full day; the second-grade classroom, 20 students; the seventh-grade

a.m. (morning) classroom, 18 students, and the seventh grade p.m. (afternoon), 20 students. All students in each class participated in the study. Groups of children (seated in rows or at tables) were monitored for class-wide performance). Additionally, target students from these classes were nominated by the teachers as having behavioral problems: three target students from the second grade, six from kindergarten, two from seventh a.m., and two from seventh p.m. All but two participants were males (one female in each of the seventh-grade classes), and all were of Hispanic ethnicity.

A process similar to the *Systematic Screening for Behavior Disorders (SSBD)* (Walker et al. 1991) was used wherein the teacher rank ordered the target students for externalizing behaviors. That is, the highest risk student as number one, the next highest as number two, the next highest as number three, until all students had been rank ordered. Two to six at-risk students from each classroom were selected as target students (based on the severity of behavior and teacher agreement of the need to monitor the student) and parent consent secured. After selecting target students, the researcher observed the target students, gathering baseline data on them for the purpose of identifying the effects of CW-FIT on the target students specifically.

Teachers

All participating teachers had requested classroom management assistance as part of the protocol of a larger University research project investigating School-wide Positive Behavior Support (Sugai and Horner 2006). The teacher in the kindergarten classroom was a 36-year-old female with 5 years of teaching experience. The teacher in the second-grade classroom was a 28-year-old female with 4 years of teaching experience, and the teacher in the seventh-grade classroom was a 28-year-old male with no previous teaching experience. Additionally, when the second-grade teacher went on maternity leave, a 24-year-old female with student teaching experience took her place.

Setting

All participating students attended classes on a parochial school campus that included elementary and junior high school buildings in a large Midwestern city. The school campus had a total population of 347 students: 87 % received free or reduced lunch, 8 % had Individual Education Programs (IEPs), 87 % were racial or ethnic minorities, and 25 % were English language learners. Selected classrooms were representative of the school demographics.

Observations

Each observation session across all conditions lasted no more than 40 min. Data were collected on the three measures in random order within each observation: (1) on-task, praise, reprimands, and compliance (15 min); (2) out-of-seat and talking out (10 min); and (3) hand-raising (10 min). That is, the order for the dependent variables was randomly selected for observation at the start of each session.

Observations lasted approximately 40 min to include all three combined measures within the CW-FIT sessions which were 45–60 min. Observers used a small handheld timer to track intervals. Activities remained similar across all baseline and CW-FIT sessions during which observations occurred. Data collection during baseline began at the beginning of the session. Data collection during CW-FIT sessions began following the pre-corrects of the skills (see procedures).

Dependent Variables

The dependent variables consisted of five class-wide and individual student behaviors: on-task, compliance, hand-raising, out-of-seat, and talking out; and two teacher behaviors: praise and reprimands. Class-wide data were collected for the groups as seated in their teams (i.e., all students were seated in rows or tables, and these were the assigned groups for CW-FIT and data). These groups remained constant throughout the study (see Table 1 for operational definitions).

Dependent variables also included praise and reprimands. Praise statements were defined as positive feedback to students for their behavior. Examples include "Thank you for staying in your seat," or "I like the way Lorie raised her hand to get my attention." Reprimands were defined as statements for students to stop inappropriate behaviors and/or to provide corrective feedback. Examples included, "Stop talking out and raise your hands to talk!" or "It's getting loud in here!"

Data Collection

Data were collected by the researcher using paper and pencil data sheets. Four recording procedures were used: (a) frequency for teacher praise (only verbal praises were recorded) and reprimands; (b) momentary time sampling for on-task behavior ; (c) partial interval recording for out-of-seat, talking out, and hand-raising behavior ; and (d) compliance was computed as percent of opportunities to comply. Each behavior was coded for groups of children, and calculated as "class-wide," and for individual target students. Observers scanned from group to group in consistent order and noted behaviors of children within the groups (see definitions, Table 1). On-task, teacher praise and reprimands, and compliance were collected simultaneously for 15 min of the session. Out-of-seat and talking out were collected simultaneously during a separate 10-min block of the session. Hand-raising was collected during a separate 10-min block of the session. The number of data points collected varied across classes based on availability with the most data collected during second grade with the change in teachers, and fewest data points in the seventh-grade classes, as they began the study approximately 3 weeks after the primary grades.

Recording procedures for on-task occurred for 15 min during the session using a 30-s momentary time sampling procedure (Sulzer-Azeroff and Mayer 1991). Target behaviors were recorded at the end of each 30-s interval. For class-wide on-task, all students in the group had to be on-task at the moment. For each target student, he or she had to be on-task at that moment. The observer looked at group one, scored; then looked at group two, scored; etc. On-task was scored as a plus (+); off-task was

Table 1 Operational definitions of student behaviors

On-task	On-task is defined as attending to the lesson in the classroom (e.g., watching the teacher, reading along with the teacher silently or out loud, following directions of the teacher, writing assignments, answering/asking academic questions)
Out-of-seat	Out-of-seat is defined as being out of one's assigned seat without the teacher's permission. Conversely, in-seat CW-FIT skill was taught/defined as: (a) Sit and stay seated in our chair, (b) keep your feet underneath your desk, and the chair legs on the floor
Talking out	Talking out is defined as talking in a voice loud enough to be heard at the next table. This is addressed with the CW-FIT skill "Talk with a quiet voice," defined as: (a) Talk in a whisper and (b) Talk only to those at your table or those close to you
Hand-raising	Hand-raising: Hand-raising is defined as raising one's hand and waiting to be called on before talking. Hand-raising is addresses with the CW-FIT skill, "How to get the teacher's attention." This includes: (a) Look at the teacher, (b) Raise your hand, (c) Wait for the teacher to call on you, and (d) Ask your question or give your answer
Compliance	Follow directions the first time is defined in the CW-FIT skill as: (a) Look at the person and listen, (b) say OK, (c) do it, and (d) check back (if needed). For data collection, the class, group, or individual must be in the process of following the direction given within 5 *s* after the direction is given

scored as a minus (−) on the data sheet. Criteria for scoring a + for group on-task were that all students in the group had to be engaged at the 30-s moment. Class-wide means consisted of the sum of group data divided by the number of groups in the class. For example, if there were four groups in the class, and their percent of intervals of on-task was 87, 96, 74, and 92 %, our computation was as follows: (.87 + .96 + .74 + .92) = 3.49/4 = 87.25 % class-wide mean for that session.

During this same 15-min session, frequency counts of teacher praise and reprimands were tallied (points awarded were not counted as praise).

Recording procedures for compliance (i.e., following directions) were recorded as percent of opportunities to comply, also recorded during this 15-min period. For example, if 20 students were in the classroom and a direction was given (e.g., open your notebooks or move to the corner carpet) and 15 of the students responded within 5 s, it was scored as percent of compliance (i.e., 15/20, or 75 %).

Recording procedures for out-of-seat and talking-out behaviors were recorded during 10 min of the session using 20-s partial interval recording procedures (Sulzer-Azeroff and Mayer 1991). If the behavior occurred by any member in a group at any time during the 20-s interval, a plus (+) was recorded on the data sheet; a minus (−) was recorded if the behavior did not occur. Recording procedures for hand-raising were conducted during a separate 10-min portion of the class period using the same 20-s partial interval procedures.

Criterion for Changing Conditions

Baseline data were collected until the data were stable (i.e., visual inspection indicated data were not trending up or down) or moving in the opposite direction of that expected from intervention, and a minimum of three data points were collected. Additionally, on-task behavior was the behavior chosen as the indicator for a change

in conditions. The rationale for this decision was that on-task is an indicator of both appropriate behavior and disruptive behavior. That is, disruptive behaviors would be at lower levels as these behaviors are incompatible with on-task behavior. Reversals were data-based decisions, that is, five to six on-task data points at high, stable rates during CW-FIT conditions indicated readiness for reversal.

Procedural Fidelity

An 11-item checklist was used to measure CW-FIT procedural fidelity. The checklist was completed by the observer immediately following data collection. Measures were taken on the implementation of CW-FIT to identify the extent to which the intervention components were implemented as designed. The CW-FIT fidelity portion of the checklist contained items directly related to the CW-FIT intervention (i.e., pre-corrects of skills occur at beginning of session, skills are prominently displayed on posters, corrections for behavior match language of skills, timer is set for 1- to 3-min intervals, praise for use of skills is delivered at 4:1 ratio for praise-to-reprimands, teacher uses specific praise for use of skills, point goal determined, team point chart displayed, points are awarded to individuals or teams for use of the skills at set intervals). These items were scored as "yes" or "no." The procedural fidelity was assigned a score by dividing the total number of "yes" items by the total number of items on the checklist. For example, 9 "yes" items divided by the total of 11 items equals a score of 81.8 %. Procedural fidelity measures were taken on approximately 36 % of all intervention sessions. Procedural fidelity was at a mean level of 94 % (range 45–100 %). Only a few sessions indicated low fidelity for the kindergarten teacher. These were related to with not assigning points, not tallying points at the end of the game, and not providing frequent praise or immediate access to the reinforcer at the end of the game. Coaching improved fidelity back to 80 % or higher.

Inter-Observer Agreement (IOA)

A graduate student was available for IOA for 26 % of all baseline and 26 % of CW-FIT intervention sessions. Training was accomplished through verbal and written instructions and was continued until the second observer and the primary researcher obtained an agreement score of 85 % or higher for at least three data sessions. In addition, the researcher and the second observer reviewed the operational definitions for the behaviors being measured before each IOA session. IOA was computed by using the "point-by-point" agreement method (Kennedy 2005). Inter-observer agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements, then multiplying by 100 (Kennedy 2005; Repp et al. 1996). IOA averaged 93 % (range 89–100 %) for on-task behavior, 98 % (range 92–100 %) for compliance, 92 % (range 89–96 %) hand-raising, 94 % (range 90–98 %) for out-of-seat, 91 % (85–96 %) for talking-out behavior, and 97 % (82–100 %) for teacher praise.

Consumer Satisfaction

A consumer satisfaction survey was distributed to all teachers and students to assess social validity (Horner et al. 2005; Wolf 1978). The consumer satisfaction survey for the teachers consisted of 16 questions on a 5-point Likert scale. These questions related to the (a) ease of use, (b) acceptability, (c) difficulty, (d) effectiveness, and (e) student responsiveness. The survey for the kindergarten and second-grade students consisted of 15 questions with three-point response choices: (a) Yes/a lot with a smiley face above, (b) Okay/All right with a straight face above, and (c) No/Not at all with a frowning face above. Questions related to (a) likeability, (b) easy to learn or do, and (c) ease of performance. The survey for the seventh graders was the same as for kindergarten and second grade without the faces.

Experimental Design

A withdrawal design was selected to provide the desired level of control for the study (Kennedy 2005). This design provided analysis of functional relations across teachers and classes resulting in multiple points at which a functional relation was confirmed (see Horner et al. 2005). In the design, experimental conditions were manipulated as ABABABABAB in second A = Baseline and B = CW-FIT, and ABAB in kindergarten and seventh-grade classes.

Procedures

Teacher Preparation

A short initial conference with each teacher was conducted (approximately 30 min). The intended outcome of the conference was to obtain preliminary information about the classroom environment, the target students, and to discuss the students' classroom behaviors. Following the conference, classroom observations were conducted to obtain information on student behaviors, as well as other variables that could be affecting the student behaviors (e.g., seating arrangements, access to materials).

After the observations, a second teacher conference was conducted. In this conference, the information from the initial conference and the information from the observation of the classroom were discussed (e.g., identifying target students and student seating arrangements for classroom groups). This meeting lasted approximately 30 min. Seating arrangements for teams were agreed upon and remained constant throughout all conditions. Target students were initially identified by the teachers as students having difficulty staying engaged, and exhibiting out-of-seat, talking out, talking to peers, or other disruptive behaviors and with use of a screening and ranking procedure (see Participants section).

Baseline Procedures (A)

During baseline conditions, the classrooms were not altered other than assigning groups; thus, baseline was "business as usual." The common procedure in the classrooms included reprimands and a behavior chart. The behavior chart consisted of a matrix of pockets containing green, yellow, and red cards. Each student had their name on one of the pockets. The card facing out of the pocket each morning was green. If a student disrupted the class, he was asked to flip his card. This first occurrence usually indicated a yellow card meaning the loss of recess. The next occurrence moved to a red card, which precipitated a written report of the incident and a note home to the parent.

Baseline data for the class and individuals were collected during those times during the day when the teacher reported the most challenging student behaviors occurred. This same time period was also designated for the CW-FIT intervention. For example, the most challenging student behaviors occurred during morning instructional floor time and reading for kindergarten, math for the second-grade, and science for both seventh-grade classes. These times remained constant during the study conditions.

CW-FIT Intervention (B)

Intervention consisted of (a) teacher training including written procedural descriptions, (b) explanation of CW-FIT materials, (c) student training on the CW-FIT skill components, and (d) the group contingency token economy (earning points and getting rewards). Teacher training in CW-FIT started with a discussion of the baseline data and the need for CW-FIT. Next, the teachers were provided with training material.

CW-FIT training materials consisted of (a) a written description of CW- FIT, (b) teaching scripts for each skill, (c) brief written procedures for the group contingency game, and (d) procedures for assessing student reinforcer preferences. Other classroom materials included (a) five skill posters (11 × 17) that define the rules for the appropriate behaviors linked to the CW-FIT game and (b) a point chart (11 × 17) for keeping track of group points. A small timer was provided for the teacher to keep track of intervals for giving points on the point chart for group use of the target skills. Use of the teams and points in a game format and rewards for meeting point goals was intended to reinforce occurrences of the targeted student behaviors. Training consisted of a 1.5-h session reviewing materials, modeling and role-playing CW-FIT and two to three demonstrations in class.

The day immediately following teacher training, the CW-FIT intervention was implemented in the classrooms and consisted of (a) teaching students prosocial classroom behaviors, (b) teaching the teachers to reinforce the students' appropriate use of the behaviors (e.g., specific praise and points for use of skills), and (c) reinforcing teacher use of specific praise statements related to the students' use of the skills within the group contingency game format. Researchers modeled the main components of CW-FIT: They taught the first prosocial skill lesson with the class, set a goal for the first session, awarded points using specific praise, and

delivered a reward for meeting the goal. This was then monitored for two to three sessions as the teacher independently engaged in the behaviors, with prompts as needed. CW-FIT intervention data were not collected for at least 3 days, and after intervention, fidelity for teachers was at a minimum level of 80 % for one session. All of the classes met this requirement, and data collection began on the fourth day after training.

The first component of the CW-FIT intervention is to teach the students replacement behaviors for the inappropriate behaviors that currently function to (a) obtain attention (adult or peer); (b) escape demands; and (c) gain access to materials and activities. In this study, the following skills were taught to all classes: (a) How to get the teacher's attention, (b) Follow directions the first time, and (c) Ignore inappropriate behavior. These skills cover the range of behavior functions (i.e., attention, escape, and access; Kamps et al. 2006; Mitchem et al. 2001). With these skills, students are able to access attention appropriately by raising their hand and receive attention in the form of praise and points for appropriate use of all CW-FIT skills. Raising their hands to request assistance for difficult tasks was a replacement to engaging in inappropriate behaviors to escape a difficult task. Earning points and teacher praise was contingent on students using the skills, including following directions and ignoring inappropriate behaviors. Meeting the CW-FIT point goal for the session earned access to preferred items in the form of tangible rewards (i.e., items nominated by students during a reinforce assessment process), and 4–5 min of social activities (e.g., dancing, brain breaks, free time) providing brief escape from school work. Two additional skills were taught: (d) Staying in our seat, and (e) Talking in a quiet voice. Student training consisted of the researcher and the teacher using a direct instruction model for teaching skills to the students (i.e., defining the skill, modeling the skill, having students role-play examples and non-examples of the skills, and providing feedback on their performance). For the kindergarten and second-grade classes, the skills were identified and demonstrated one at a time, every other school day, across a two-week period. This skill training was spread over a series of days because many of the younger students did not have the skills that are taught in the CW-FIT protocol or the skills were not a fluent component of their repertoire. For the two seventh-grade classes, all of the skills were identified and demonstrated in one session.

The second component of CW-FIT is related to reducing attention to inappropriate behaviors through planned ignoring. Teachers and students were given direct instruction and opportunities to practice ignoring those student behaviors targeted for elimination. This second component was addressed by teaching students the skill, "Ignore other's inappropriate behavior," and by coaching teachers to reduce reprimands (attention) to minor inappropriate behaviors. Through the process of teaching and reinforcing appropriate behaviors and ignoring inappropriate behaviors, the teachers and students developed a new repertoire of classroom behaviors.

The final component of CW-FIT is *differential reinforcement*. Differential reinforcement for the study was defined as frequent attention to appropriate behaviors, and rewards in the form of praise, and other reinforcer preferences. For example, students were praised when raising their hands to get the teacher's

attention and not shouting out. Similarly, students are praised for staying seated as opposed to being out of their seats.

As stated earlier, CW-FIT functions within an interdependent group contingency game, token economy format. Each classroom was divided into four or five groups of three to five students per group. These groups were usually rows of students or groups of students' desks facing each other. During the game, the teacher set the timer to beep at 2- to 4-min intervals. This was determined by the teacher based on the predetermined goal, and opportunities to give frequent attention to use of skills. After each timer beep, the teacher rewarded the teams exhibiting the appropriate behaviors at the time of the beep with a point (the token economy aspect of the game). At the end of the game, for example, if the goal was 14 points, and three of five groups earned 14 or more points, they received the designated reward and the teams that earned less than 14 points did not receive the reward (Kamps et al. 2010).

CW-FIT sessions started with teaching the aforementioned classroom skills (e.g., How to get the teacher's attention, etc.), presented in a token economy, game format (e.g., earning points to get a reward). Subsequent sessions, for 1–2 weeks, began with a series of pre-corrects (i.e., brief reviews and reminders of the CW-FIT skills lasting for 2–3 min). The number of points earned during the CW-FIT game depended upon two factors: (a) the amount of class time available and (b) the time intervals set on the timer. When CW-FIT was first introduced into the classroom, timer intervals were short (1–2 min) and the point goal was large (e.g., 20 points). To insure that the game was highly reinforcing and all teams won, a maximum of 70–80 % of the possible points was required as a point goal. For example, during a game scheduled to run for 40 min with average timer duration of two min, the possible points would be ($40/2 = 20$). A reasonable point goal would be ($20 \times .70 = 14$ points). Each time the timer beeped, the teams were specifically praised for using their newly acquired skills and for their good behavior. As the students learned the skills and subsequently the rules of the game, the intervals became longer (3–5 min), the number of points required to win the game decreased, and students were required to earn a higher percentage of the possible points (85–90 %) in order to win. Lengthening the intervals resulted in thinning the schedule of reinforcement, as students were required to maintain the skills and behaviors for longer periods of time.

The length of CW-FIT sessions varied depending on the activity and the amount of class time available. In the kindergarten class, sessions were 45 min during centers or reading. In the second-grade math class, sessions were 50 min and in the two seventh-grade science classes, sessions were 45–60 min. Observations and data collection occurred only during these designated class times. These activities/routines and time period remained constant throughout the study and conditions.

Reversal Conditions

Reversals were initiated by having a short teacher meeting (10 min) in which the teacher was asked to remove all CW-FIT stimuli from the classroom, discontinue use of the timer, and conduct class sessions "business as usual" for the next few days as reversal data were collected.

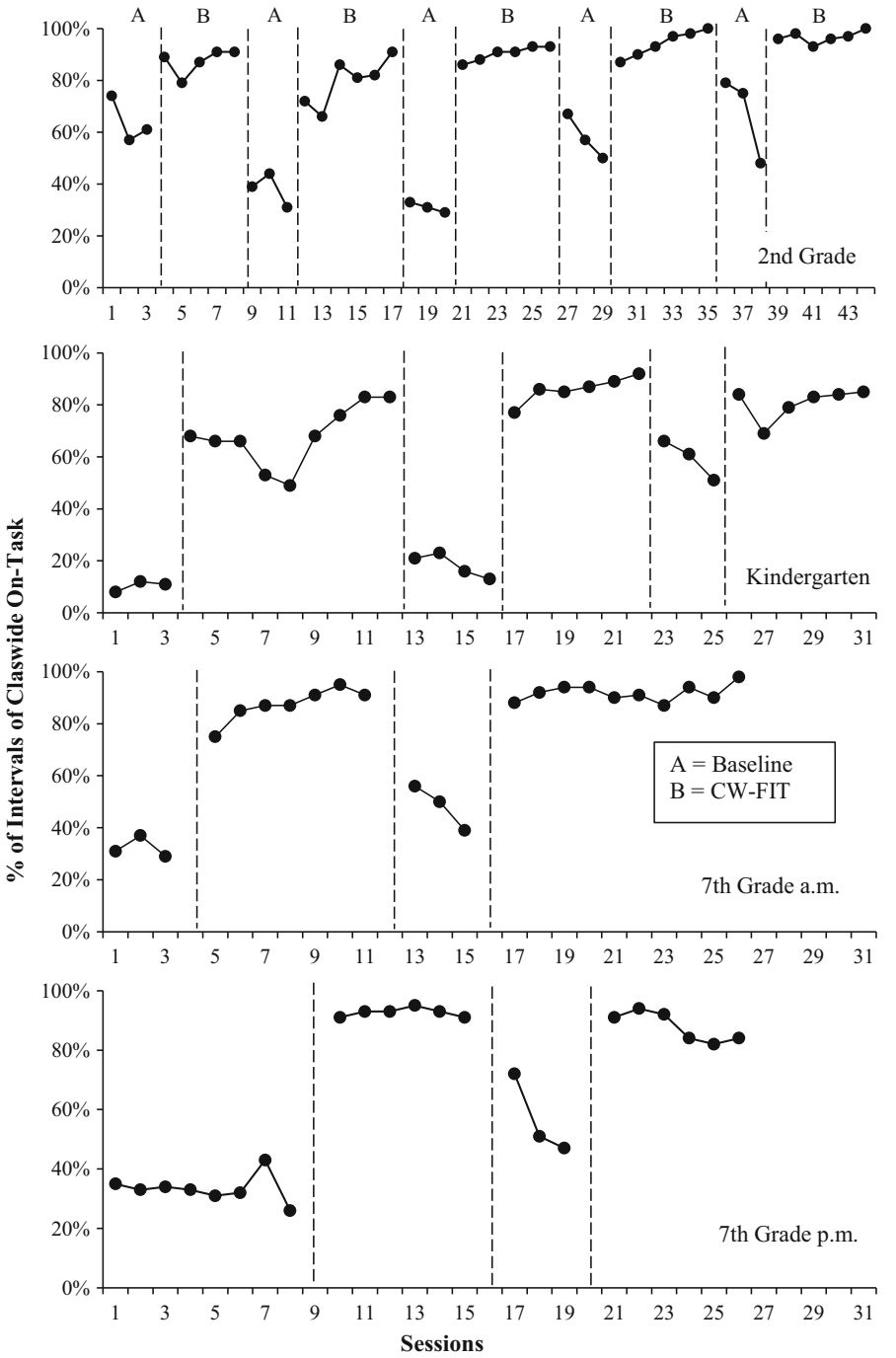

Fig. 1 Class-wide on-task behavior

Results

Class-Wide Behaviors

Research question one asked: To what extent did the CW-FIT intervention have an effect on increasing students' prosocial classroom behaviors? Results indicated that class-wide on-task behavior increased substantially over baseline levels in all classes. As depicted in Fig. 1 and Table 2, baseline levels of class-wide on-task

Table 2 Means, standard deviations, and change index across conditions for all classes and all of the six behaviors

	Mean		Standard deviation		Change index
	Baseline	CW-FIT	Baseline	CW-FIT	
Second grade					
Class-wide on-task	51.7 %	89.7 %	17.1 %	7.9 %	38.1 %
Class-wide compliance	43.7 %	92.6 %	14.4 %	6.0 %	48.9 %
Class-wide hand-raising	4.85 %	45.0 %	3.5 %	10.4 %	40.3 %
Class-wide out-of-seat	43.0 %	6.2 %	23.2 %	9.1 %	−36.8 %
Class-wide talking out	36.1 %	7.7 %	15.4 %	8.9 %	−28.4 %
Teacher praise	1.6	36.4	1.5	12.3	34.7
Kindergarten					
Class-wide on-task	28.2 %	76.8 %	22.2 %	11.7 %	48.6 %
Class-wide compliance	24.2 %	92.1 %	11.9 %	3.1 %	67.9 %
Class-wide hand-raising	8.2 %	38.3 %	2.4 %	8.7 %	30.1 %
Class-wide out-of-seat	58.5 %	9.5 %	13.9 %	3.0 %	−49.0 %
Class-wide talking out	59.3 %	10.3 %	13.2 %	4.3 %	−49.0 %
Teacher praise	3.1	32.3	3.1	11.6	29.2
Seventh-grade AM					
Class-wide on-task	40.3 %	89.9 %	10.7 %	5.2 %	49.2 %
Class-wide compliance	37.0 %	91.8 %	17.8 %	8.9 %	54.8 %
Class-wide hand-raising	3.8 %	30.2 %	3.9 %	10.4 %	26.3 %
Class-wide out-of-seat	22.0 %	1.29 %	8.0 %	1.57 %	−20.7 %
Class-wide talking out	44.0 %	8.1 %	13.2 %	2.57 %	−35.9 %
Teacher praise	0.7	36.5	1.03	9.8	35.8
Seventh-grade PM					
Class-wide on-task	39.7 %	90.3 %	13.1 %	4.4 %	50.5 %
Class-wide compliance	35.8 %	95.6 %	17.7 %	2.3 %	59.9 %
Class-wide hand-raising	8.8 %	40.3 %	3.1 %	6.7 %	31.4 %
Class-wide out-of-seat	28.5 %	1.8 %	6.9 %	1.9 %	−26.6 %
Class-wide talking out	39.8 %	7.3 %	11.3 %	4.2 %	−31.9 %
Teacher praise	.5	38.5	.52	5.6	36.8

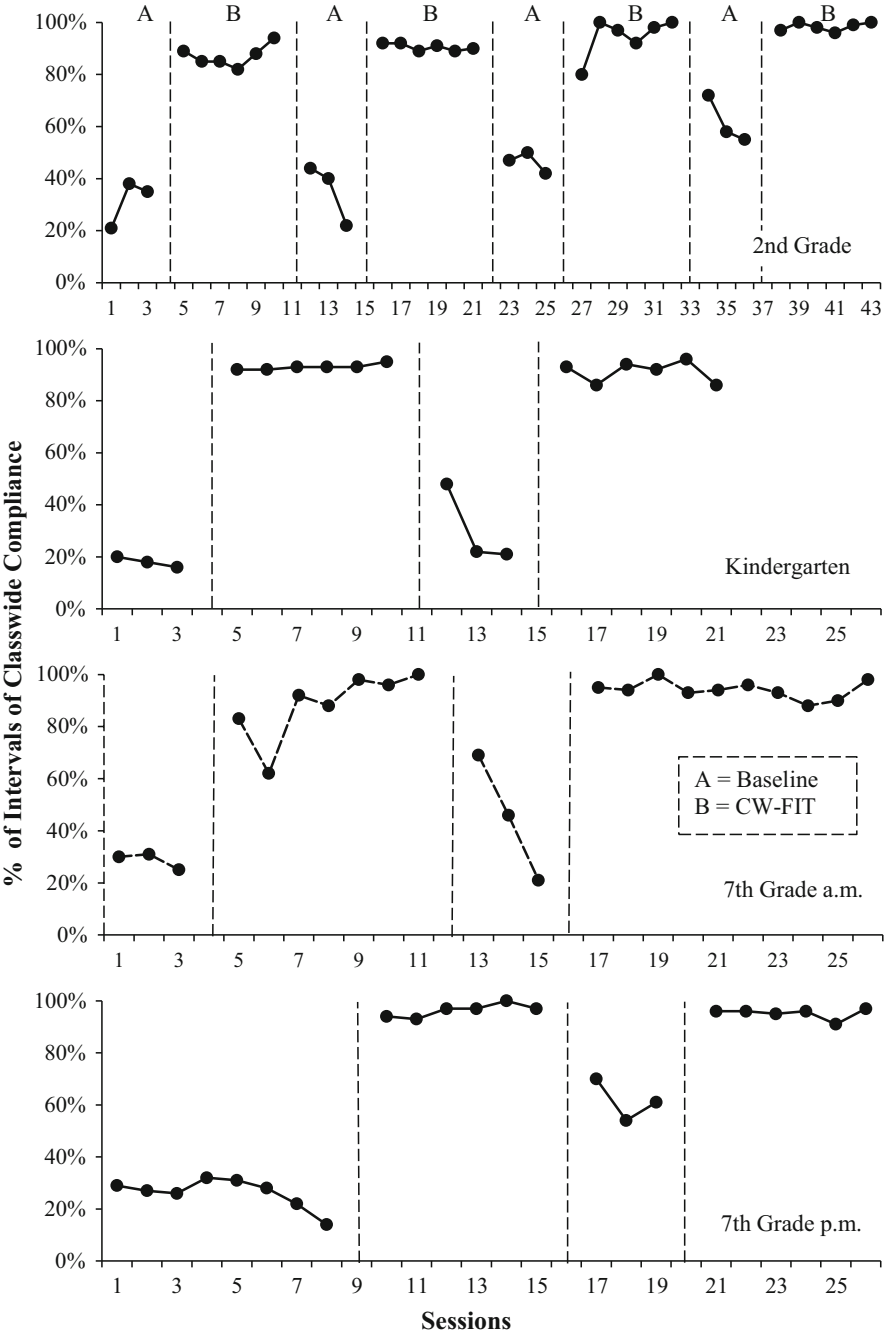

Fig. 2 Class-wide compliance

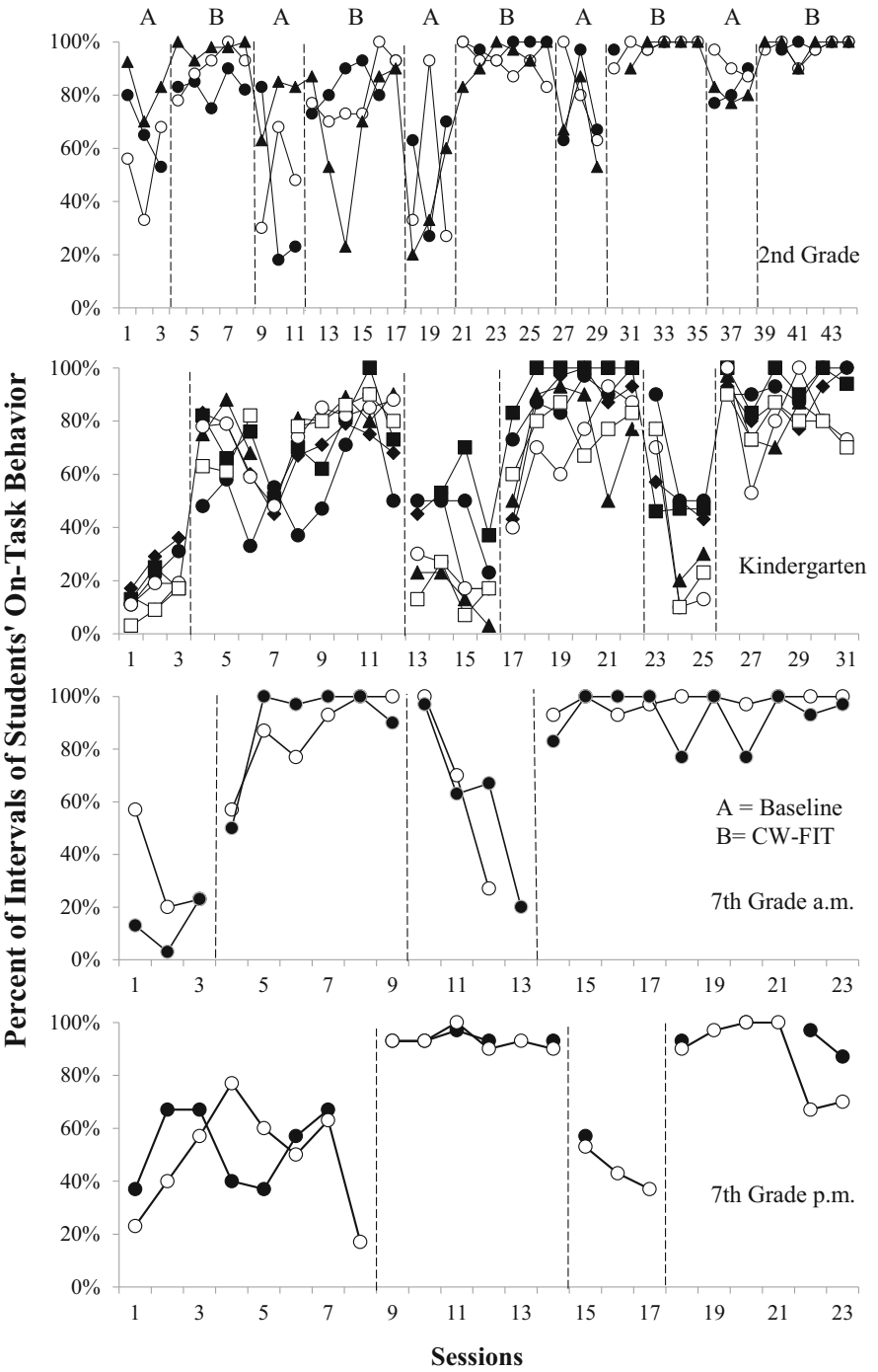

Fig. 3 Individual students' on-task behavior

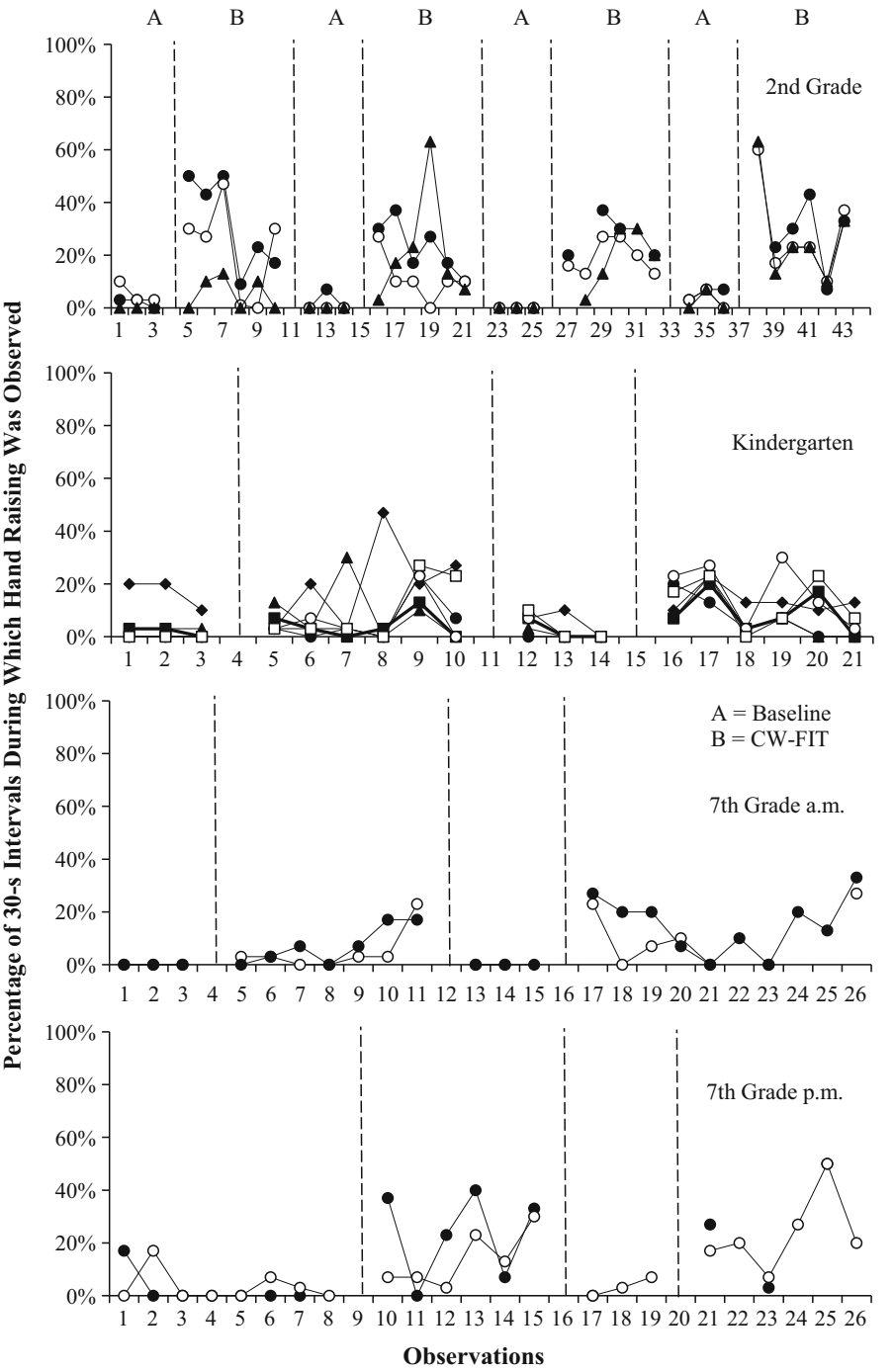

Fig. 4 Individual students' hand-raising behavior

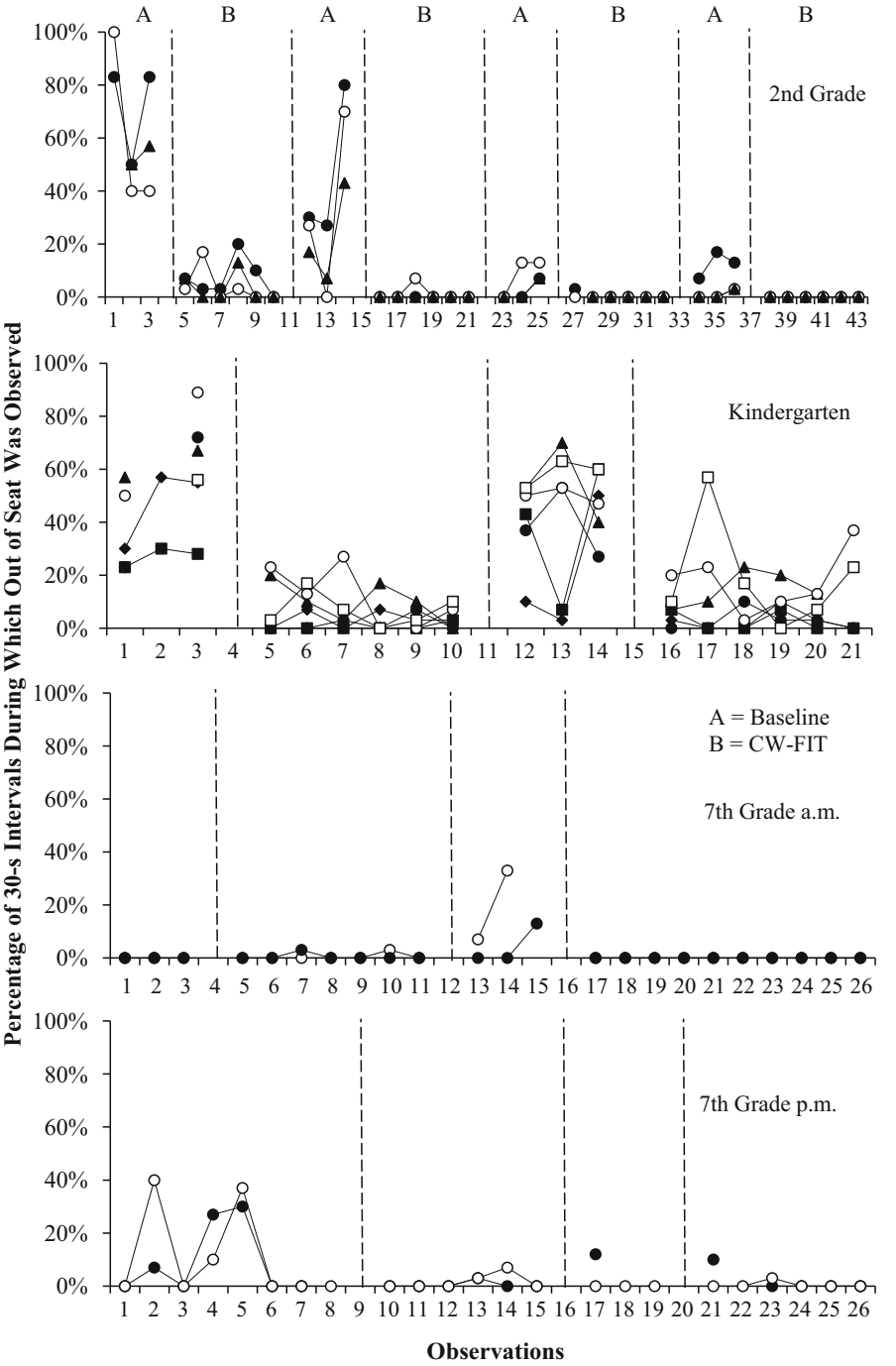

Fig. 5 Individual students' out-of-seat behavior

behavior pooled means across groups of students within classes ranged from moderate to low levels at 51.7 % (range 29–79 %), 28.2 % (8–66 %), 40.3 % (range 29–56 %), and 39.7 % (26–72 %) of intervals per session for the second-grade, kindergarten, seventh-grade a.m., and seventh-grade p.m. classes, respectively. During CW-FIT, the class-wide on-task behavior increased to much higher levels and with substantially less variability. CW-FIT intervention pooled means (see Table 2) increased to 89.7 % (72–100 %), 76.8 % (49–92 %), 89.9 % (75–98 %), and 90.3 % (82–95 %) for the classes, respectively. The exception was for the fourth and fifth data points of the first CW-FIT intervention condition in the kindergarten classroom (see Fig. 1, panel 2). At this time, a booster session was conducted. Table 1 shows the pooled means during baseline and CW-FIT conditions.

Effects of CW-FIT on Class-Wide Behaviors Taught During the Intervention

As part of the CW-FIT intervention, several appropriate classroom behaviors were operationally defined and taught to the students in each classroom with data collected to monitor skill use: (a) compliance, (b) hand-raising, (c) out-of-seat, and (d) talking-out behaviors. In Fig. 2, class-wide compliance behavior across conditions is depicted showing improvements during CW-FIT. Table 2 shows pooled means across all baseline conditions ranged from moderate to low levels: compliance at 43.7 % (range 21–72 %), 24.2 % (range 16–48 %), 37 % (range 21–69 %), and 35.8 % (range 14–70 %) of opportunities per session for the second-grade, kindergarten, seventh-grade a.m., and seventh-grade p.m. classes, respectively. CW-FIT intervention pooled means increased to 92.7 % (range 82–100 %), 92.1 % (range 86–95 %), 91.8 % (range 83–100 %), and 95.7 % (range 91–100 %), respectively. Hand-raising pooled means increased from 4.7 to 45 %, 8 to 38 %, 3 to 30 %, and 8 to 40 % across classes and conditions. Out-of-seat and talking-out behaviors decreased in all four classes during CW-FIT conditions (see Table 2; Fig. 2).

Effects of CW-FIT on Individual Target Students Behaviors

During baseline and the CW-FIT intervention, data were gathered on four behaviors for individual students: (a) on-task, (b) hand-raising, (c) out-of-seat, and (d) talking out. Figure 3 depicts on-task for individual target students across conditions and classes. On-task behavior improved for all 13 target students during CW-FIT conditions over baseline levels, with some reduction with the withdrawal of the intervention. During the later withdrawal conditions, the second-grade class reduced on-task but did not return to the low baseline levels. The overall baseline mean for all target students was 42.7 % (range 20.3–69.1 %) of intervals, and the mean CW-FIT rate was 86.4 % (range 75.2–96.7). Figure 4 shows hand-raising behavior for individual students across conditions and classes, with increases during the CW-FIT conditions. The overall baseline level for all target students was low at 2.3 % (range 0.3–11.2 %) of intervals and the mean CW-FIT rates increased somewhat, though it

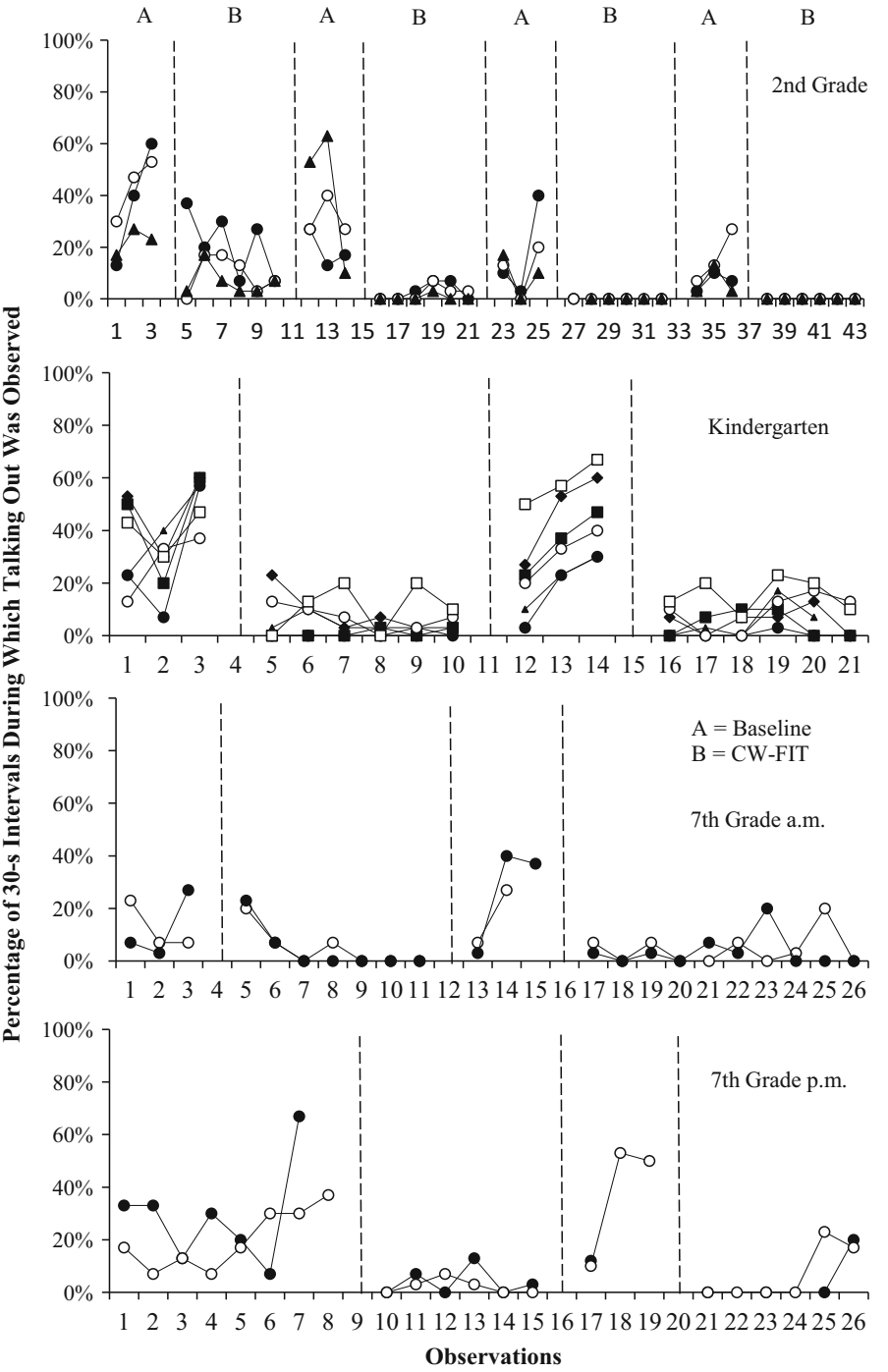

Fig. 6 Individual students' talking-out behavior

was not a high rate behavior, mean of 14.1 % (range 6.6–27.3 %). Percentages of out-of-seat behavior are depicted in Fig. 5 for individual target students across conditions and classes. Baseline for all students averaged 29.7 % (range 2.17–58.8 %) of intervals decreasing to 4.1 % (range 0.2–14.7 %) during CW-FIT. Differences were noted for classes. For second graders, out-of-seat decreased with the use of CW-FIT, but did not return to baseline levels in the later withdrawal conditions. Kindergarten at-risk students showed the highest baseline levels with decreases during CW-FIT and increases during withdrawal conditions. The seventh-grade morning and afternoon classes showed limited out-of-seat behavior across conditions. Figure 6 shows the percentage of intervals of talking-out behavior across students by conditions and across classes. All students showed clear reductions in talking-out behaviors during CW-FIT conditions from baseline and withdrawal phases. The kindergarten students showed the highest baseline levels and decreases during intervention. Overall baseline average for all students was 28.5 % (range 14.2–49 %) of intervals and 4.9 % (range 0.3–13 %) during CW-FIT.

Teacher Praise and Reprimands

Research question two was as follows: To what extent did the CW-FIT intervention have an effect on increasing teacher praise statements? Results indicated that teacher praise statements increased substantially in frequency over baseline levels for all four classroom teachers. As depicted in Table 2, the pooled means of praise statements across all baseline conditions were very low at 1.6 (range 0–5), 3.1 (range, 0–8), 0.7 (range 0–2), and 1.7 (range 0–6) per 15-min session for the second-grade, kindergarten, seventh-grade a.m., and seventh-grade p.m. classes. The frequency of praise statements during CW-FIT increased substantially to means of 36.3 (range 10–63), 32.3 (range 18–68), 36.4 (range 23–59), and 38.5 (range 26–49) for the second-grade, kindergarten, seventh-grade a.m., and seventh-grade p.m. classes, respectively. Small changes were noted between conditions for reprimands. Means changed from 3.3 to 2.6 for second grade across conditions with slight increases in the other classes: 4.8–5.1, 2.7–3.1, and 1.7–3.2, respectively.

Consumer Satisfaction

Consumer satisfaction surveys were distributed to all four teachers and all students who participated in the study. The teacher consumer satisfaction scores were 4.0 for the kindergarten teacher, 4.7 for the original second-grade teacher, 4.4 for the substitute second-grade teacher, and 3.96 for the seventh-grade science teacher. The overall mean for all four teachers was 4.3 out of a possible score of 5. Consumer satisfaction scores for the kindergarten student class averaged 2.8. Consumer satisfaction scores for the second-grade student class averaged 2.8. Consumer satisfaction scores for the seventh-grade a.m. student class averaged 1.96. Consumer satisfaction scores for the seventh-grade p.m. student class averaged 2.4, which was a mean score of 2.5 out of 3 across all four classrooms.

Discussion

The purpose of this study was to perform a systematic replication of the previous studies of CW-FIT (Kamps et al. 2010; Wills et al. 2010) and to address novel components and new research questions. This replication adds to the generality of the effects of CW-FIT on a wider range of behaviors, students, teachers, and grade levels. A novel addition was to measure the behaviors taught as CW-FIT skills/lessons during the intervention. Specifically, the impact of CW-FIT for talking out, out-of-seat, and hand-raising was measured, showing a link between skills taught in the CW-FIT lessons (see Table 1) and improved performance over baseline levels. These data had not been collected in prior studies. The study added to what is known about CW-FIT through implementation in a predominantly Hispanic school, with a high number of English language learners, and across a span of grade levels from kindergarten to seventh grade. This provides additional data for the literature reporting the effects of interdependent group-oriented contingencies (e.g., Kamps et al. 2010; Skinner et al. 2000).

Class-Wide and Individual Student Behavior

The general findings of this study were that CW-FIT had a desirable effect on all of the dependent variables directly manipulated by the intervention package. This study demonstrated increases in students' class-wide on-task, compliance, and hand-raising behaviors, and decreases for talking-out and out-of-seat behaviors. Individual target student data were gathered on 13 target students across the four classrooms with improvements noted on all four behaviors: increased on-task, hand-raising; decreased out-of-seat, and talking out (see Figs. 3 and 4). Findings are consistent with previous CW-FIT studies showing increases in on-task behavior and decreases for high-risk target students (Kamps et al. 2010; Wills et al. 2010). These findings in decreasing disruptive behaviors are similar to past research reports using interdependent group-oriented contingencies (e.g., Davies and Witte 2000; Kelshaw-Levering et al. 2000, Lohrmann and Talerico 2004; Theodore et al. 2001).

Additionally, although there was variability between teachers' and students' behaviors, all of the dependent variables demonstrated increases in appropriate behaviors and decreases in disruptive behaviors at levels that produced a more positive teaching environment. Finally, with these increases in appropriate classroom behavior and decreases in disruptive behavior, more time was available for teaching and learning.

Teacher Behavior

Teacher-specific praise statements during baseline were most often at levels equal with teacher reprimands. However, during the CW-FIT intervention, teacher praise levels often exceeded the ratio 4–1 as recommended by previous researchers (e.g., Rhode et al. 1992; Sutherland et al. 2000; Walker et al. 2004) and were equal to those found in previous CW-FIT studies (e.g., Kamps et al. 2010; Wills et al. 2010).

Previous researchers have demonstrated that teacher praise (i.e., social reinforcement and teacher attention to positive classroom behaviors) has been found to be related to positive changes in student behaviors (e.g., Walker et al. 2004). Teacher reprimands did not show meaningful changes during CW-FIT.

Consumer Satisfaction

Teacher satisfaction surveys indicated a moderate to high degree of satisfaction with an overall mean of 4.27 out of 5. Several positive comments were noted: "I really liked the support of the program." "It was motivating for the students." "It's a great program." and one negative comment, "Using the timer was annoying." The mean student scores across all four classrooms were 2.47 out of 3, indicating a moderate to high degree of satisfaction.

Limitations

Although this study produced important results in several areas, the following limitations should be considered. The study included a small number of classrooms and teachers ($n = 4$). While an experimental design was used and important findings produced, this study represents only one of several studies reporting the results of CW-FIT. It has been suggested that for an intervention to demonstrate generality and achieve external validity, multiple replications across researchers, participants, and geographic locations must be performed (see Horner et al. 2005). While the teachers implemented at a high level of fidelity during intervention, these data were not collected in baseline. Thus, no documentation was collected to ensure CW-FIT procedures were not in place during baseline. Anecdotally however, teachers were not observed to use point systems or group contingencies prior to the intervention. Finally, this was a convenience sample of teachers from an urban parochial school, and all of the teachers in this sample had requested classroom management assistance. A further limitation is that the researcher collected data for the study rather than have independent observers.

Implications for Research and Practice

The implications for research and practice are that teacher and student behaviors, individually as well as class-wide, are amenable to change using the CW-FIT intervention. Effects were demonstrated with students in kindergarten through seventh grade showing increased appropriate classroom behaviors including on-task, compliance, and hand-raising, and decreased disruptive behaviors (talking out, out-of-seat). Improvements assisted in creating a classroom environment conducive to teaching and learning. Findings suggest continued use of group contingency interventions in urban classroom settings, and particularly in classes with multiple students engaging in disruptive behavior, with different age groups, and in larger studies.

Acknowledgments 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 teachers, students, and families for their time and ongoing support.

References

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

Barrish, H. H., Saunders, M., & Wolf, 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.

Cashwell, T. H., Skinner, C. H., & Smith, E. S. (2001). Increasing second-grade students' reports of peers' prosocial behaviors via direct instruction, group reinforcement, and progress feedback: A replication and extension. *Education and Treatment of Children, 24*, 161–175.

Christ, T. J., & Christ, J. A. (2006). Application of an interdependent group contingency mediated by an automated feedback device: An intervention across three high school classrooms. *School Psychology Review, 35*, 78–90.

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.

Darch, C. B., & Thorpe, H. W. (1977). The principal game: A group consequence procedure to increase classroom on-task behavior. *Journal of School Psychology, 14*, 341–347.

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.

Embry, D. E. (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: A practice guide* (NCEE #2008–012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved February 15, 2010 from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

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

Graziano, C. (2005). School's out. *Edutopia,* February/March, 40–44.

Greenwood, C. R., Horner, R. H., & Kratochwill, T. R. (2008). In C. R. Greenwood, T. R. Kratochwill, & M. Clements (Eds.), *Schoolwide prevention models: Lessons learned in elementary schools* (pp. 3–30). New York, NY: Guilford Press.

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent and independent group contingencies for controlling disruptive behavior. *Journal of Special Education, 16*, 101–110.

Hanley, G., Iwata, B., & McCord, B. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185.

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.

Hawken, L. S., & Horner, R. H. (2003). Evaluation of a targeted intervention within a schoolwide system of behavior support. *Journal of Behavioral Education, 12*, 225–240.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165–179.

Jacob, B. (2007). The challenges of staffing urban schools with effective teachers. *Excellence in the Classroom, 17*(1), 129–153.

Kamps, D., Conklin, C., & Wills, H. (2015a). Use of self-management with the CW-FIT group contingency program. *Education and Treatment of Children, 38*(1), 1–32.

Kamps, D., Wendland, M., & Culpepper, M. (2006). Active teacher participation in functional behavior assessment for students with emotional and behavioral disorders risks in general education classrooms. *Behavioral Disorders, 31*, 128–146.

Kamps, D., Wills, H., Bannister, H., Heitzman-Powell, L., Kottwitz, E., Hansen, B., & Fleming, K. (2015b). Class-wide function-related intervention teams "CW-FIT" efficacy trial outcomes. *Journal of Positive Behavior Interventions*. doi:10.1177/1098300714565244.

Kamps, D., Wills, H., Heitzman-Powell, L., Laylin, J., Szoke, C., Hobohm, T., & Culey, A. (2010). Class-wide function-based intervention teams: 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.

Kennedy, C. H. (2005). *Single-case designs for educational research*. New York: Pearson.

Lew, M., Mesch, D., Johnson, D. W., & Johnson, R. (1986). Components of cooperative learning: Effects of collaborative skills and academic group contingencies on achievement and mainstreaming. *Contemporary Educational Psychology, 11*, 229–239.

Lewis, T. J., Powers, L. J., Kelk, M. J., & Newcomb, L. L. (2002). Reducing problem behavior on the playground: An investigation of the application of schoolwide positive supports. *Psychology in the Schools, 39*, 181–190.

Litrow, L., & Pumroy, D. K. (1975). Brief technical report: A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8*, 341–347.

Lloyd, J. W., Eberhardt, M. J., & Drake, G. P. (1996). Group versus individual reinforcement contingencies within the context of group study conditions. *Journal of Applied Behavior Analysis, 29*, 189–200.

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

Maggin, D., Johnson, A., Chafouleas, S., Ruberto, L., & 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.

McComas, J. J., Goddard, C., & Hoch, H. (2002). The effects of preferred activities during academic work breaks on-task engagement and negatively reinforced destructive behavior. *Education and Treatment of Children, 25*, 103–112.

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

Patrick, C. A., Ward, P., & Crouch, W. D. (1998). Effects of holding students accountable for social behaviors during volleyball games in elementary physical education. *Journal of Teaching in Physical Education, 17*, 143–156.

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.

Repp, A., Deitz, D., Boles, S., Deitz, S., & Repp, C. (1996). Differences among common methods for calculating interobserver agreement. *Journal of Applied Behavior Analysis, 9*, 109–113.

Rhode, G., Jensen, W. R., & Reavis, H. K. (1992). *The tough kid book: Practical classroom management strategies*. Longmont, CO: Sopris West.

Robertshaw, C. S., & Hiebert, H. D. (1973). The astronaut game: A group contingency applied to a first grade classroom. *School Applications of Learning Theory, 6*, 28–33.

Salend, S. J., & Lamb, E. A. (1989). Effectiveness of a group-managed interdependent contingency system. *Learning Disability Quarterly, 9*, 268–273.

Schanding, G. T., & Sterling-Turner, H. E. (2010). Use of the mystery motivator for a high school class. *Journal of Applied School Psychology, 26*, 38–53.

Sharp, S. R., & Skinner, C. H. (2004). Using interdependent group contingencies with randomized selected criteria and paired reading to enhance class-wide reading performance. *Journal of Applied School Psychology, 30*, 29–45.

Skinner, C. H., Cashwell, C. S., & Dunn, M. S. (1996). Independent and interdependent group contingencies: Smoothing the rough waters. *Special Services in the Schools, 12*, 61–78.

Skinner, C. H., Cashwell, T. H., & Skinner, A. L. (2000). Increasing tootling: The effects of a peer-monitored group contingency program on students' reports of peers' prosocial behaviors. *Psychology in the Schools, 37*, 264–270.

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. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245–259.

Sulzer-Azeroff, B., & Mayer, G. R. (1991). *Behavior analysis for lasting change*. Belmont: Wadsworth.

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8*, 2–8.

Swain, J. J., Allard, G. B., & Holborn, S. W. (1982). The good toothbrushing game: A school-based dental hygiene program for increasing the toothbrushing effectiveness of children. *Journal of Applied School Psychology, 15*, 171–176.

Theodore, L. A., Bray, M. B., Kehle, T. J., & Jenson, W. R. (2001). Randomization of group contingencies and reinforcers to reduce classroom disruptive behaviors. *Journal of School Psychology, 39*, 267–277.

Thomas, J. R., Lee, A. R., McGee, L., & Silverman, S. (1987). Effects of individual and group contingencies on disruptive playground behavior. *Journal of Research and Development in Education, 20*, 66–76.

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The good behavior game: 1969 to 2002. *Behavior Modification, 30*, 225–233.

Trevino-Maack, S., Kamps, D., & Wills, H. P. (2014). A group contingency plus self-management intervention targeting at-risk secondary students' class-work and active responding. *Remedial and Special Education*. doi:10.1177/0741932514561865.

Walker, H., Ramsey, E., & Gresham, F. (2004). *Antisocial behavior in school: Evidence-based practices* (2nd ed.). Belmont, CA: Thomson Wadsworth.

Walker, H. M., Severson, H. H., & Feil, E. G. (1991). *Systematic screening for behavior disorders: SSBD*. Longmont, CO: Sopris West.

Weeden, M., Wills, H., Kamps, D., & Kottwitz, E. (in press). 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*.

Williamson, B. D., Campbell- Whatley, G. D., & Lo, Y. (2009). Using a random dependent group contingency to increase on-task behaviors of high school students with high incidence disabilities. *Psychology in the Schools, 46*, 1074–1083.

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*, 154–171.

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

Wolf, M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.