HAMMILL INSTITUTE
ON DISABILITIES

Guided by Theory, Informed by Practice: Training and Support for the Good Behavior Game, a Classroom-Based Behavior Management Strategy

Jeanne M. Poduska, ScD[1] and Anja Kurki, PhD[1]

Abstract

Moving evidence-based practices for classroom behavior management into real-world settings is a high priority for education and public health. This article describes the development and use of a model of training and support for the Good Behavior Game (GBG), one of the few preventive interventions shown to have positive outcomes for elementary schoolchildren lasting through to young adulthood, ages 19–21, including reductions in the use of drugs and alcohol, school-based mental health services, and suicide ideation and attempts. We first describe the conceptual framework guiding the development of the model of training and support. Data on implementation of the model, from an ongoing trial of GBG being conducted in partnership with the Houston Independent School District, are then presented. We end with a discussion of the lessons learned and the implications for the next stage of research and practice.

Teachers rate support with classroom behavior management as a pressing need (Evertson & Weinstein, 2006; Reinke, Stormont, Herman, Puri, & Goel, 2011; Walter, Gouze, & Lim, 2006). Yet, teachers have limited tested tools to help children and youth master the behavioral skills they need to succeed as students and classroom management is not a priority of pre-service teacher training. In fact, the National Council for Accreditation of Teacher Education (NCATE; 2008) does not require proof of proficiency in this area for teachers to be certified. The good news is that a number of preventive interventions aimed at aggressive, disruptive behavior, including the Good Behavior Game (GBG; Barrish, Saunders, & Wolf, 1969), have shown both short- and long-term positive impact on epidemiologically based randomized field trials. The challenge is that the outcomes seen in highly controlled studies are often not replicated when programs are moved into general practice in schools (Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005; Elliott & Mihalic, 2004; Ringwalt, Vincus, Ennett, Johnson, & Rohrbach, 2004).

Moving programs into practice requires an understanding that the *programmatic intervention*, for example, GBG, and the *support system*, that is, training and support, for the intervention are independent, though interrelated, components of a whole (Chen, 1998; Greenberg, Domitrovich, & Bumbarger, 2001; Klein & Sorra, 1996). Indeed, research has shown that regardless of specific content or delivery mode, programs require some form of professional development or delivery strategy to be implemented effectively (Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005; Lochman et al., 2009; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997).

GBG: The Programmatic Base

GBG is a team-based classroom behavior management strategy that helps children master the role of student and be successful at the key demands of the classroom, including paying attention and working well with others. Through GBG, children work together to create a positive learning environment for all students by monitoring their own behavior as well as that of their classmates. As a universal preventive intervention, the focus of GBG is on strengthening the classroom environment and socializing children to the role of student.

[1]American Institutes for Research, Washington, DC, USA

Corresponding Author:
Jeanne M. Poduska, American Institutes for Research, 1000 Thomas Jefferson Street, NW, Washington, DC 20007, USA.
Email: jpoduska@air.org

In GBG classrooms, teams of children win when they meet behavioral expectations by not exceeding a criterion level of classroom rule infractions. There are four core elements: classroom rules/behavioral expectations, team membership, monitoring behavior, and reinforcement. The teacher and students work together to define classroom rules and the corresponding behavioral expectations. After observing students for several weeks, the teacher divides students into several teams, heterogeneous with regard to behavior, learning, and sex. During the game, students work together with their teammates to maintain good behavior. When a child does not meet a behavioral expectation/breaks a classroom rule, the team receives a point. At the end of the game, each team that has four or fewer points receives a reward. Teams do not compete against one another; all teams can win. Over the course of the year, the duration of the game increases and GBG is played at different times throughout the day, during different activities and instructional subjects, and in a variety of venues. Reinforcers shift from being tangible to intangible and are delivered on an intermittent schedule. Teachers may change teams during the school year to maintain balance or to provide students the opportunity to work with a broad range of peers. Of importance to the education sector is that as a strategy rather than a curriculum, GBG does not compete for instructional time.

GBG, tested through two highly controlled epidemiologically defined randomized field trials conducted in partnership with the Baltimore City Public Schools (Dolan et al., 1993; N. S. Ialongo et al., 1999) has demonstrated positive impacts from elementary school through young adulthood. At first and second grades, there were reductions in aggressive, disruptive behavior and increased time on task (Brown, 1993; Dolan et al., 1993; N. S. Ialongo et al., 1999). In later elementary school, there were reductions in rates of conduct disorder diagnoses in students (Brown et al., 2008; N. Ialongo, Poduska, Werthamer, & Kellam, 2001). And there was less aggressive/disruptive behavior through middle school (N. Ialongo et al., 2001; Kellam, Ling, Merisca, Brown, & Ialongo, 1998). Results to young adulthood, ages 19–21, include reductions in the rates of antisocial personality disorder; drug and alcohol abuse and dependence; tobacco use (Kellam et al., 2008); the use of school-based services for problems with emotions, behaviors, and drug and alcohol use (J. M. Poduska et al., 2008); violence (Petras et al., 2008); and suicidal ideation (Wilcox et al., 2008). Positive impact has been greatest for males who entered first grade exhibiting higher levels of aggressive, disruptive behavior.

Conceptualizing a Model for Training and Support

A third epidemiological randomized field trial conducted in Baltimore City Public Schools included an aim to design a model of training and support focused on developing a district's capacity to implement and sustain GBG (J. Poduska et al., 2009). We drew upon several literatures as we developed the model for GBG training and support. Mindful of the fact that GBG and the corresponding support system are independent, interrelated components, we focused on both individual and organizational change to understand the fit between the model of training and support, the implementer, and the broader context (Baldridge & Burnham, 1975; Giacquinta, 1973; Glasgow, Lichtenstein, & Marcus, 2003). The work of Leavitt (1965) and others (Owens, 2004; Robertson, Roberts, & Porras, 1993) provided a framework for understanding schools and districts as organizations. We also considered the extensive literatures on stages of program implementation and multi-level contextual factors that impede or enhance programs being moved into practice keeping in mind that different factors might be salient at different stages (Adelman & Taylor, 2003; Berman & McLaughlin, 1975, 1978; Burns & Hoagwood, 2005; Domitrovich et al., 2008; Dusenbury, Brannigan, Falco, & Hansen, 2003; Fixsen et el., 2005; Greenhalgh et al., 2005; McLaughlin, 1990; Rogers, 2003; Rohrbach, Grana, Sussman, & Valente, 2006). With these literatures in mind, we focused on three areas: (a) the model of training and ongoing support for teachers and local GBG coaches and trainers; (b) consideration of multi-level contextual factors; and (c) the use of data to monitor program implementation and inform professional development.

Theoretical Base for Coaching

Two strands of training and support were developed: one for teachers and one for local coaches. We considered the training teachers received through the prior randomized field trials and viewed it alongside the literature on professional development and coaching with the aim of developing a model that was feasible, relevant, and replicable. The literature on professional development for educators highlights the importance of providing opportunities for active learning through observation, meaningful discussion, practice, and reflection (Garet, Porter, Desimone, Birman, & Yoon, 2001; Lieberman, 1996; Loucks-Horsley, Hewson, Love, & Stiles, 1998). Research also suggests that effective professional development is an ongoing process rather than a single event; aligned with and incorporated into one's daily professional work; and of long enough duration to ensure a thorough and working knowledge of the core program principles and how they translate into practice (Hargreaves & Fullan, 1992; Leach & Conto, 1999; Little, 1993; McCormick, Steckler, & McLeroy, 1995; Rose & Church, 1998).

Central to our model of training and support is a coach who works directly with the teacher in the classroom. GBG coaches support teachers in bringing together the new knowledge of GBG learned in group-based trainings with their craft knowledge of classroom teaching through

relationship building (mentoring and providing feedback), data-driven decision making (observing, analyzing, and interpreting data), and supporting teachers in becoming self-reflective and self-directed implementers of GBG (American Institutes for Research [AIR], 2005; Spouse, 2001; Taylor, 2007). Han and Weiss's (2005) conceptual model of the process of teacher implementation and sustainability of practices places the coach at the center of "a self-sustaining feedback loop." The primary purpose of coaching is to support teachers' implementation of a program; this support is provided through observation, direct feedback, and modeling in keeping with Bandura's social cognitive theory (1986). The goal of the initial GBG trainings is to provide the necessary knowledge about the intervention; coaching provides the necessary support for teachers to become independent self-directed learners who can sustain the intervention in their classrooms after the active coaching support is phased out. The second function of coaching is to promote teachers' motivation to implement the new practice (Han & Weiss, 2005). For example, a coach can help a teacher with low self-efficacy with regard to classroom behavior management attribute small incremental successes to GBG. The self-sustaining loop occurs when teachers reach a high enough level of success and mastery that students' classroom behavior is attributed to the program, thereby motivating the teacher to use the program.

Current Study

In this mixed-methods article, we draw upon quantitative data and qualitative interviews to understand the dynamic between the model of training and ongoing support for GBG and implementation quality. The data were collected as part of a randomized field trial of GBG being conducted in partnership with the Houston Independent School District (HISD; J. Poduska, Gomez, Capo, & Holmes, 2012). An important feature of the trial is that we are implementing GBG under conditions that are not tightly controlled. Rather than research staff serving as GBG coaches who support teachers in the classroom, we trained staff of a local professional development provider as GBG coaches. As such, this study offers an important opportunity to examine how the training and support model operates outside of a highly controlled research trial in a large urban school setting. In the "Method" section, we present the details of the randomized field trail, the model of training and support, and the results. Lessons learned, areas for future research, and implications for practice are then discussed.

Method

Design

This article draws upon data from an ongoing, federally funded randomized field trial testing the impact of GBG under two conditions of professional development against a control condition. Children entering first grade in 18 schools were randomized across first-grade classrooms/teachers, balancing on gender, academic achievement, and behavior using procedures we have used in two previous trials (N. S. Ialongo et al., 1999; J. Poduska et al., 2009). First-grade teachers entered the study at the beginning of the 2010–2011 school year or at the beginning of the 2011–2012 school year and were randomized to one of three interventions conditions: GBG Basic, GBG with Coach, or Standard (control). Language of instruction was used as a blocking factor because in some classrooms, the language of instruction was Spanish and in others, it was English. The total sample includes 73 classrooms. The GBG Basic condition provided teachers with 3 days of group-based training: a 2-day initial training and a 1-day booster session, described below. In GBG with Coach, the group-based activities were enhanced by the presence of a coach who worked in the classroom with the teacher over the course of one school year, described below. Group-based training was conducted in English; training materials were in English. Of note is the fact that GBG with Coach is the model of training and support the AIR uses when working with schools and districts to develop local capacity for GBG.

Participants

The participants for this article were the 26 teachers randomized to the GBG with Coach condition who taught in the same classroom for the entire school year during their first year of GBG implementation, and four local GBG coaches. Twenty-three of the teachers were females. Three (14%) of the 22 teachers who completed a background survey were younger than 30 years; 13 (59%) were between 30 and 50 years. Four (18%) were older than 50 and 2 did not respond to the question. Fifteen (68%) teachers had completed a bachelor's degree, 5 (23%) had completed a master's degree, and 1 had a doctoral degree. Six (27%) had 5 or fewer years of teaching experience; 4 (18%) had 6 to 10 years; 6 (27%) had 11 to 15 years; and 5 (23%) greater than 15 years; 1 did not respond.

In keeping with our model of developing local capacity to implement and sustain GBG, AIR contracted with a local purveyor of professional development who provided staff to be trained and serve as local GBG coaches. As it was a time of leadership transition in the district, school district personnel and the teachers union suggested that we work with the purveyor to help ensure that local capacity for GBG would be sustained (J. Poduska et al., 2012). We interviewed and trained four local GBG coaches, all females. Each coach worked with between five and nine teachers. Two were fluent in Spanish and worked with the teachers in the bilingual classrooms. All had experience as classroom teachers. Three had a master's and one had a

doctorate. Three had special education backgrounds; one had a background in early childhood.

Training and Support for Teachers and Coaches

Teachers participated in 2 days of group-based initial training on GBG that included sessions on GBG core elements, demonstrations of strategies and procedures, and practice in the delivery of GBG. A 1-day group-based booster session was held after the mid-year break. During the school year, the teacher was supported in the classroom by a coach. The coaching process is cyclical—each cycle includes four bi-weekly visits, each 90 min in length. There were three coaching cycles across the school year. During the first visit in a cycle, the GBG coach conducts a classroom observation using the *Fidelity Checklist* and the *GBG Scoreboard* (described below). Together, these measures provide data on the teacher's practices and student behavior in the classroom. The teacher and coach then develop a *Professional Development Plan* (*PD Plan*) for the cycle. During the next three visits in a cycle (visits 2–4), the coach supports the teacher in GBG practices through activities such as modeling, guided practice, and visiting other classrooms to observe GBG implementation. Activities vary to meet the needs of individual teachers (AIR, 2005; Hasbrouck & Denton, 2005).

The coaching model combined a pre-determined amount of coaching support with support differentiated by teacher need. In the first semester, Cycle 1, coaches were expected to visit the classroom for 90 min every other week to support teachers as they first set up the conditions for GBG and then infused GBG into the classroom routine and the instructional day. During the second semester, Cycles 2 and 3, support could be differentiated. Teachers who were rated as "high" on the fidelity checklist and whose students demonstrated positive behavior could receive a check-in visit once a month rather than a bi-weekly visit. Teachers who rated low on the fidelity checklist and needed additional support could receive a greater number of visits from the coaches.

Local GBG coaches were trained in situ by training teachers and participating in professional development activities with AIR trainers over the school year. They participated in the group-based GBG trainings described above. They received additional training on using data to support effective GBG implementation and inform professional development; designing individual PD plans and providing ongoing support to teachers; working with adult learners; and providing professional development to teams of teachers, other school-level staff, and principals. AIR trainers conducted two implementation audits during which the trainer and coaches visited classrooms. Coaches participated in regularly scheduled phone calls with an AIR trainer with an agenda driven by the needs articulated by the coaches. E-mail support was provided to individual coaches and the group, as needed.

Measures

GBG practice and support. During the first visit of each coaching cycle, the GBG coach completed a *GBG Fidelity Checklist* that provides information about teacher practices in classroom behavior management. Behavioral indicators are rated for occurrence (yes/no) and level of quality (scale of 1 being *needs significant improvement* to 6 *exemplary*); a global rating of classroom behavior management is also made (α = .80). A companion rubric guides the quality ratings. Teachers use the *GBG Score Record* to record when the game was played (date and time, instructional subject), the number of minutes played, and the result for each team (the number of rule infractions). Teachers and coaches also used the score record to log information about "probes"—instances where the teacher keeps a tally of infractions but does not announce the game to the students. A probe provides information about the extent to which students are generalizing the skills. The coach and teacher develop a *PD Plan* for each cycle by the end of the second visit. The *Progress Report* is completed by the coach at visits 2–4 and records the date and time of the visit, the focus of coaching, and the activities undertaken.

Self-report of multi-level contextual factors. Teachers and coaches completed surveys in the spring on a set of multi-level contextual factors hypothesized to influence implementation. The *Evidence-Based Practice Attitude Scale* (EBPAS; Aarons, 2005) is a 15-item measure that assesses the attitudes of mental health professionals toward adopting evidence-based practices, which we adapted for use with teachers. Responses are made using a Likert-type scale of 1 (*not at all*) to 6 (*a very great extent*): (α = .77 for the full scale). EBPAS measures the extent to which a provider would adopt a program based on its intuitive appeal, its use by a colleague or that it is a requirement of the job; their willingness to try new interventions; and whether they see evidence-based practices as useful. The *Perception of GBG* scale measures perceived ease of implementation; ease with which GBG fits into daily practices; fit of GBG with one's teaching philosophy; and perceived effectiveness of GBG using a 6-point Likert-type scale ranging from 1 (*not at all*) to 6 (*very much*). The coefficient alpha for teacher reports' for the full measure is .92. The *Coach–Teacher Alliance* scale has 23 items that assess the quality of the working relationship between a coach and teacher (α = .92). The measure incorporates 10 items from a Teacher–Coach Alliance scale developed by members of the National Behavior Rating Coordination Center (Wehby, Maggin, Moore Partin, & Robertson, 2012). Corresponding versions are completed by the teacher and coach. *Principal Support* rated by teachers and GBG coaches on six items includes tangible resources and encouragement (α = .86).
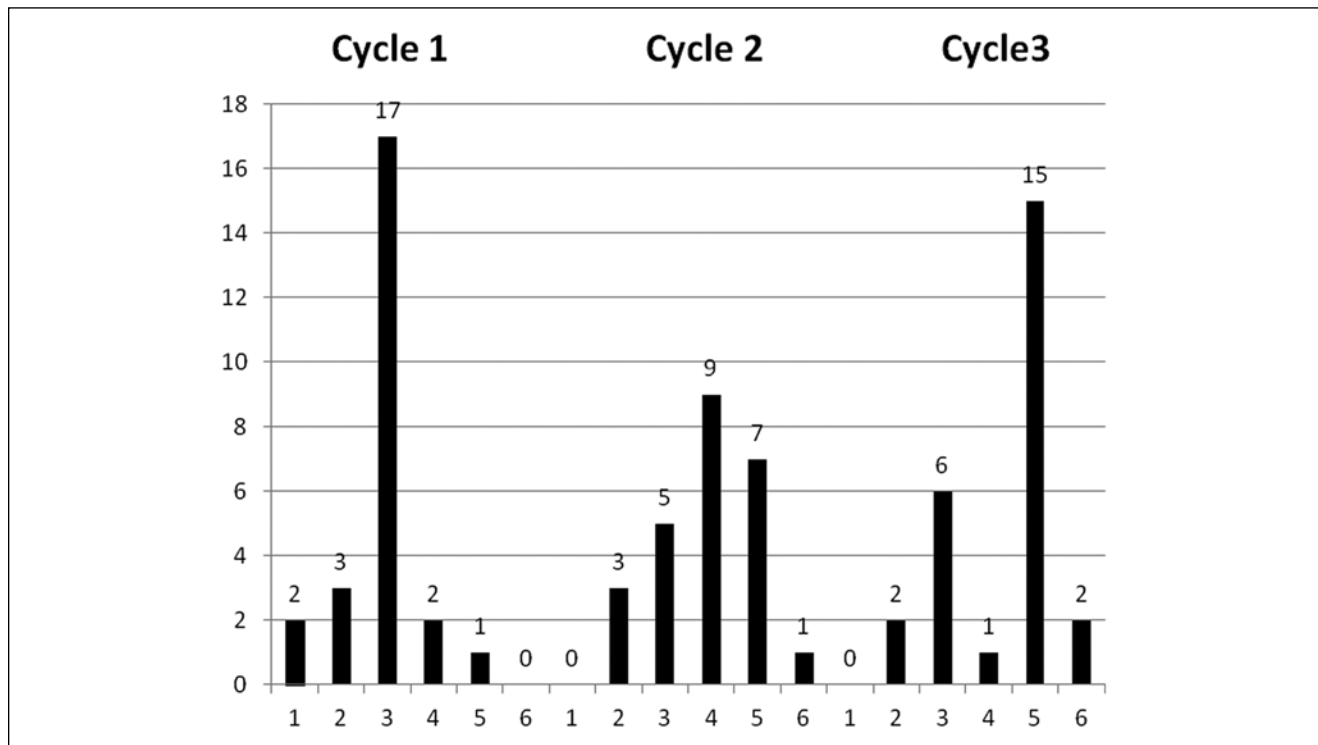
Figure 1. Global quality rating of teachers' GBG practice across the three coaching cycles.
Note. Global quality rating is on a scale from 1 (*poor*) to 6 (*excellent*). GBG = Good Behavior Game.

Qualitative interviews. To further inform our understanding of the coaching process, the coaches were interviewed about the support they provided to teachers and the training and support they themselves received. Thirty minute interviews were conducted by the AIR research team, guided by a detailed protocol with open-ended questions and probes. Interviews were audiotaped and transcribed; NViVo was used for coding.

Results

Analytic Approach

The results are organized by type of data and mode of measurement. We first describe implementation to the model of GBG training and support using the measures listed above under *GBG practice and support*, followed by exploration of multi-level contextual factors. We then present data from the qualitative interviews with the coaches.

Dose of Training and Support

All teachers and coaches received the group-based initial GBG training, which was held each year in late October after baseline data collection for the parent study was completed. They also attended the mid-year, group-based Booster session. Each

school year had three coaching cycles: one in the first semester and two in the second semester. Overall, teachers received less coaching support as the school year progressed. Coaches reported a total of 225 classroom visits: 90 visits occurred in Cycle 1, 74 visits in Cycle 2, and 61 visits in Cycle 3. Of the 26 teachers who stayed in the study the entire year, 12 (46%) received a total of between 10 and 12 coach visits; 6 (23%), 7 to 9 visits; and 8 (31%), 4 to 6 visits. In Cycle 1, 18 of the 26 teachers (69.2%) received the expected 4 visits, 2 (7.6%) 3 visits, and 6 (23.1%) 2 visits. In Cycle 2, 8 teachers (30.8%) received 4 visits, 7 (26.9%) 3 visits, 10 (38.4%) 2 visits, and 1 (4.5%) 1 visit. The trend of fewer visits continued in Cycle 3, 3 teachers (11.5%) received 4 visits, and 9 (34.6%), 8 (30.8%), and 6 (23.1%) received 3, 2, and 1 visits, respectively.

Dose and Proficiency of Practice

Consistent with our coaching model, the reduction in the number of coaching visits that occurred was associated with an overall increase in teachers' proficiency with GBG over the course of the school year as measured by the coaches' global rating of quality on the *GBG Fidelity Checklist* (see Figure 1). At the beginning of Cycle 1, 1 teacher out of 26 received a rating of 5 or 6 (the criterion for moving from a bi-weekly coaching support schedule to a monthly check-in by the coach). Eight (30.8%) and 17 (65.4%) teachers met

this criterion at the beginning of Cycles 2 and 3, respectively. To examine patterns of proficiency, we dichotomized the global rating into "high" for ratings of five or six and "low" for ratings of one to four. Three patterns emerged. One teacher (3.8%) was rated high in Cycle 1 and retained that rating across the year. The majority of teachers (16 teachers: 61.5%) started with a rating of low and upon reaching a high rating, maintained it. Within this group, half (8: 30.7%) were rated low in Cycle 1 and high in Cycles 2 and 3; half were rated low in Cycles 1 and 2 and high in Cycle 3. Nine teachers (34.6%) were rated low in all three cycles.

Adherence to the Coaching Cycle

Despite the fact that the majority of teachers were gaining proficiency with GBG over the course of the school year, there was a discrepancy between the number of classroom visits we would have expected to occur based on teacher need and the number of visits reported by the coaches. In Cycle 1, it was expected that all teachers would receive four classroom visits: 18 teachers (69.2%) received four visits. In Cycles 2 and 3, coaches were expected to conduct four classroom visits unless the teacher scored a 5 or 6/high on the fidelity checklist at the first visit in the cycle. At the beginning of Cycle 2, 7 out of 26 teachers received a rating of high on the implementation checklist. Of the remaining 19 teachers, only 9 (47.4%) received four visits by coach report. At the beginning of Cycle 3, 17 out of 26 teachers received a rating of high; coaches reported visiting 3 of the remaining 9 teachers (33.3%).

We reviewed the coaches' paperwork to examine their adherence to the coaching cycle. At the first visit of a cycle, the coach was expected to complete a *GBG Fidelity Checklist*; by visit 2 a *PD Plan* was to be developed; a *Progress Report* and a *GBG Score Report* recording a probe were to be completed for the second through fourth visits. As mentioned, probes are times when teachers "play" the game without announcing it to the students and are a way to see how well the students are internalizing new skills and generalizing them to a variety of situations and contexts. Overall, coaches were consistent in that they began each cycle by observing the classroom, completing a fidelity checklist, and developing a *PD Plan*. One teacher in each of Cycles 1 and 2 did not have a *Fidelity Checklist* on file; three teachers in Cycle 2 and one teacher in Cycle 3 did not have a *PD plan* on file. However, the number of progress reports on file for visits 2 to 4 of each cycle varied widely between coaches. The data (not shown due to the small number of coaches) showed that some coaches visited some teachers far less than would be expected or that they were underreporting their coaching visits. With regard to probes, the a priori expectation was that coaches would conduct a probe at each of visits 2 to 4 in a coaching cycle. Keeping in mind that the number of visits was already less than

expected, *GBG Score Records* for probes were recorded for 46 (71.8%) of visits 2 to 4 that took place in Cycle 1; and for 39 (81.2%) and 21 (60.0%) of visits 2 to 4 that took place in Cycles 2 and 3, respectively.

Follow-up discussions with the coaches revealed some potential influencing factors. First, we discovered that there may have been some confusion around the *Progress Report* because the local purveyor had GBG coaches complete a form called a progress report for their internal recordkeeping. Second, all GBG coaches reported difficulties in scheduling visits with teachers as the school year progressed. Scheduling class visits was the most challenging for coaches who were full-time employees with job responsibilities beyond GBG because their schedules were less flexible than their colleagues who were only GBG coaches or were part-time consultants.

Focus of Professional Development

The top four professional development needs identified in the *PD Plans* for each coaching cycle are presented in Table 1, followed by the number and percent of teachers identified as having each specific need. In Cycle 1, the focus is on teachers' implementation of GBG: posting class rules; teaching behavioral expectations; assigning tasks to team leaders; announcing winners; using check, comment, and redirect; and conducting probes. There is a broader range of professional development foci in Cycles 2 and 3. This is in keeping with a greater use of intangible and delayed reinforcement and the game becoming more challenging by playing longer and with greater variation as to length of time and when it is played during the day. Of note are the two areas for professional development that were identified in all three cycles: using check, comment, and redirect and conducting probes. In GBG, when a behavioral expectation is not met, the teacher gives the team a check on a wall poster, neutrally comments on the rule that has been broken, and provides verbal reinforcement that serves to redirect the students to the desired behavior. That the use of check, comment, and redirect was identified as an ongoing professional development need is a reflection of the fact that many teachers struggle to provide students with consistent, neutral feedback regarding behavior. Conducting probes was also identified across the school year as a professional development need for teachers by coaches despite that fact that coaches themselves were not conducting probes at the level expected. We discuss this finding in the "Discussion" section.

Teacher Report of Multi-Level Contextual Factors

Overall, teachers' perceptions of evidence-based practices and of GBG were favorable. Teachers' mean score on the EBPAS

Table 1. Focus of Professional Development by Cycle: Items Ranked Within Cycle.

Focus	Rank of item as focus of professional development: Number (%) of teachers identified as having the specific need		
	Cycle 1	Cycle 2	Cycle 3
Posting class rules	1: 16 (64%)	4: 5 (20%)	
Implementing the "check, comment, redirect" system when behavioral expectations not met	2: 15 (60%)	1: 9 (36%)	1: 10 (40%)
Assigning tasks to team leaders	3: 14 (56%)		
Teaching behavioral expectations	4: 10 (40%)	4: 5 (20%)	
Conducting probes	4: 10 (40%)	2: 7 (28%)	3: 7 (28%)
Announcing winners	4: 10 (40%)		
Limiting verbal and nonverbal feedback to "check, comment, redirect" during game		2: 7 (28%)	2: 9 (36%)
Use encouraging praise as positive reinforcement		2: 7 (28%)	1: 10 (40%)
Maintaining records of the game results		3: 6 (24%)	3: 7 (28%)
Regularly providing appropriate feedback on behavioral expectations		4: 5 (20%)	4: 6 (24%)
Applying behavioral expectations across all situations			3: 7 (28%)
Using positive reinforcers			4: 6 (24%)

measuring the attitudes of teachers toward adopting evidence-based practices was 4.50 (range = 3.83–5.30). The mean score of perceived effectiveness of GBG on students' skill development was 5.51 (range = 4.38–6.0) and 5.57 (range = 4.07–6.0) for GBG attributes: ease of implementation, ease with which GBG fits into daily practices, and fit of GBG with one's teaching philosophy. Although teachers tended to rate their relationship with their coach favorably (M = 5.82, range = 5.00–6.00), there was considerable variation in the coaches' ratings (M = 5.30, range = 2.50–6.00) of the teacher–coach relationship. Administrator support for GBG implementation was rated more favorably by teachers (M = 5.76, range = 4.81–6.00) than by coaches (M = 2.63, range = 1.06–5.44).

Qualitative Interviews With Coaches

In the qualitative interviews, coaches reported that the training and support for teachers and 90 min for a coaching visit were adequate. Coaches noted two main challenges. First, all four coaches reported that scheduling classroom visits was challenging, particularly in Cycle 3 when spring achievement testing takes place. The coaches who were used full-time by the third-party purveyor and had responsibilities other than being a GBG coach had more difficulty scheduling classroom visits than their colleagues who worked as part-time consultants and had greater flexibility with regard to their schedules. Coaches also reported that teachers were less motivated if their principal was not actively supportive of GBG, illustrated by showing knowledge about the game, inquiring about how the game was being integrated into classrooms, and expecting that the game was being played. Coaches

identified their colleagues and the AIR team as their primary resources for problem solving.

Coaches offered suggestions for the model of training and support. With regard to teachers' professional development, coaches noted that some teachers would have benefited from additional support for GBG implementation at the beginning of the school year. With regard to their own professional development, coaches liked that their initial training provided a comprehensive overview of their work. They suggested that the scope and sequence of the coach training be more finely aligned with the GBG implementation calendar for the year. A specific suggestion was that the coach/trainer telephone conferences focus on specific topics at times when they are particularly relevant. They suggested that the coach calls with AIR trainers be guided by a sequence of in-depth guided discussions over the course of the year. Each GBG coach mentioned the same two things when asked about the most rewarding aspects of their work: seeing the students' progress with regard to awareness, self-regulation, and socialization and watching teachers grow and improve as practitioners.

Discussion

For evidence-based programs to be used by teachers, they must be appealing and feasible. In this study, teachers reported favorable impressions of GBG with regard to implementation, fit into daily practice, and effectiveness on children's skill development. Teachers also rated the working relationship with their coach quite highly.

The challenges to implementing the model of training and support are instructive in that they provide us opportunity to

learn and improve. In the next section, we frame challenges in relation to our three foci: (a) the model of training and ongoing support for teachers and local GBG coaches; (b) consideration of multi-level contextual factors; and (c) the use of data to monitor program implementation and inform practice. We discuss how the lessons learned from each challenge can inform the next stage of work moving evidence-based programs, GBG specifically, into practice.

Challenges Related to the Model of Training and Support

Challenge 1. We reported a discrepancy between the number of classroom visits recorded by GBG coaches using the *Progress Report* and the visits that would be expected based on teacher proficiency with GBG in the second semester.

Response 1: Develop clear expectations for training and staffing supports. In response to this discrepancy, we now focus conversations in the pre-implementation readiness stage more specifically on the functions of the coach and the lessons we have learned about the daily life of a coach in contrast to providing only a job description. We have considered possible models for staffing additional GBG coaches to ensure both skill level and flexibility of schedule with stakeholders in Houston.

Challenge 2. We learned several lessons from our review of the coaches' log data and experiences. For example, we found that we set the expectations for probe administration (i.e., when the game is played by the teacher but is not announced to the students and the results are not displayed) too high. The purpose of the probe is to see how well students are generalizing the skills used in GBG, and thus it is an important tool for assessing application of the GBG skills by students.

Response 2: Reconsider expectations after discussion. Conversations with the GBG coaches in Houston and at other sites revealed that our a priori expectation with regard to probe administration was too high. During a 60-min classroom observation, coaches can reasonably watch a game or conduct a probe; it is rare that both occur in the same visit except as a directed professional development strategy. For example, in a classroom where the students are struggling to generalize behavioral skills beyond times when the game is being played, the teacher might play a game and a probe during the same coach classroom visit to provide data for reflection and discussion. More important than determining a set numerical expectation is developing a process to ensure that the teacher and the coaches are using probes and

other data about teacher practices and student outcomes to inform GBG implementation.

Challenge 3. The feedback of GBG coaches regarding their training was consistent with the feedback of coaches in other locales (e.g., Oxfordshire, UK; Nebraska; Washington, DC) and led to a refinement of their support.

Response 3: Adjust sequence and structure of trainer/coach support. The trainer/coach calls are now more highly specified. They begin with a review of *GBG Scoreboard* and *Fidelity Checklist* data, followed by discussion of specific topics aligned with the scope and sequence of GBG in the classroom. Discussion topics at the start of GBG implementation include *Aligning the game with students' individualized educational plans; Team membership*; and *Integrating GBG with classroom and school behavioral management strategies*. Other topics include *Using GBG Score Record and Fidelity Checklist data to inform practice; Moving from tangible to intangible reinforcers;* and *Strategies for generalizing the game into other venues*. Calls also include time for problem solving among the trainer/facilitator and coach peers based on coaches' immediate concerns in working with teachers. Coaches can e-mail the trainer in advance to suggest topics given their work with teachers.

Challenges Related to Multi-Level Contextual Factors

Challenge 4. Coaches reported lack of support by the school principal and leadership as a challenge to implementation in both surveys and interviews. We report elsewhere on developing a base for GBG in Houston (J. Poduska et al., 2012). Worth noting here is that during the pre-implementation/ readiness stage, the district was undergoing leadership change and re-structuring at multiple levels. In addition to a new superintendent, there were several changes within the top central office positions, and a re-structuring of the support and supervisory structure for schools. There were also many lay-offs in the principal and teacher workforce. In short, the context in Houston is similar to the context of many large urban systems across the country.

Response 4: Develop a broad as of community support. For teachers to implement programs in a systematic and systemic fashion requires a shift in how many program developers and purveyors, and education policy makers at the local, state, and federal levels think about professional development for teachers. Moving programs into practice will likely require more time for school and district staff and community members to participate in information-sharing and buy-in activities than is currently funded by most schools and districts or offered by most program developers.

Challenges Related to Using Data to Monitor Implementation and Inform Practice

Challenge 5. The follow-up conversations the research team had with Houston-based GBG coaches about "missing data" led to a series of conversations about the use of data to inform practice with a larger group that included AIR trainers, and GBG coaches and stakeholders from Houston as well as from other locales. Although factors such as scheduling difficulties are real impediments to implementation, it became clear that a main reason *GBG Score Records and Progress Reports* were not being completed is that the data were not seen as useful by coaches in their daily practice working with teachers. GBG is quite elegant in that the playing of the game or conducting a probe generates data about how well the class is doing behaviorally. The *GBG Score Record* provides the teacher data by team on the number of checks/rule infractions as well as data on which classroom rules have been broken. The *GBG Score Record* might show that although a team is winning the game consistently, the team continues to break a specific rule during games. Similarly, data collected during probe administrations may show that students are struggling to generalize new behavioral skills during specific times of the day or during specific situations, for example, transitions.

Response 5: Train to ensure that data informs teachers' practice. To strengthen the use of data, we have added activities to the training for teachers and coaches so there is ample opportunity to practice interpreting data, determining "next steps," and developing solutions in response to what the data reveal. Coach classroom visits and meetings with teachers are structured around a data review process.

Limitations

This study had two limitations. The small sample size of coaches, $n = 4$ makes it difficult to generalize to a larger population from this study alone. We were able to confirm many of the challenges reported in this study through conversations with coaches in other locales, for example, the United Kingdom, Baltimore, Colorado, Nebraska, and with AIR GBG trainers. In addition, the fact that some coaches did not make or underreported the expected number of classroom visits means that the information in the progress reports was biased toward the coaches who did conduct classroom room visits regularly.

Future Research

There are several lines of research that can guide us toward the goal of moving evidence-based programs into general practice so that we see positive outcomes over time and programs being sustained. As illustrated in this article, the collection of both qualitative and quantitative data provided insights for improving implementation and the coaching process. We are in the process of analyzing extensive data from qualitative interviews with the teachers as well as data from classroom observations of teacher practice and student outcomes. Taken together, these data will provide additional direction as we create training and support that are differentiated by and responsive to teachers' needs and learning styles.

We are also exploring ways in which web-based technology can help improve the implementation and practice for GBG (Brown et al., 2013). Although terms such as "data-driven decision making" are common parlance in educational setting, we found that teachers and coaches need training and support to use data to inform and improve their practice. We are developing data tools, easy-to-use reports, and training that will support teachers in becoming self-directed users of data regarding student behaviors and their own management practices. Web-based training and support may serve as an effective mode to increase the reach of GBG and similar programs by being accessible to a greater number of teachers than face-to-face coaching models, particularly in rural areas and dense urban locals.

Implications for Practice and Conclusion

It is a critical time in the field of school-based prevention and classroom behavior management. Increasingly, decision makers at local, state, and federal levels require that funds be used for evidence-based programs, programs found to be effective through testing in studies using scientifically rigorous designs. To ensure that positive impacts result and programs are sustained, it is essential that the field has effective models of training and support.

Developing structures to support the systemic and systematic use of evidence-based programs in schools requires an understanding of the multi-level organizational system of schools and districts as well as the broader social and political context. To that end, it is critical that researchers, practitioners, and policy makers work together to solve problems of practice. The field needs models of researcher/practitioner partnering that promote program design at all stages as a shared activity. Such partnerships can ensure that programs and strategies along with support systems are feasible, relevant, and replicable, and that schools are ready to adopt and scale up programs and strategies as they prove effective.

References

Aarons, G. A. (2005). Measuring provider attitudes toward evidence-based practice: Consideration of organizational context and individual differences. *Child and Adolescent Psychiatric Clinics of North America*, 14, 255–271.

Adelman, H. S., & Taylor, L. (2003). On sustainability of project innovations as systemic change. *Journal of Educational & Psychological Consultation*, 14, 1–25.

American Institutes for Research. (2005). *Conceptual overview of the coaching component of the reading professional development impact study*. Washington, DC: Author.

Baldridge, J. V., & Burnham, R. A. (1975). Organizational innovation: Individual, organizational, and environment impacts. *Administrative Science Quarterly*, 20, 165–176.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

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.

Berman, P., & McLaughlin, M. W. (1975). *Federal programs supporting educational change, vol. 4: The findings in review*. Santa Monica, CA: Rand Corp.

Berman, P., & McLaughlin, M. W. (1978). *Federal programs supporting educational change, vol. VIII: Implementing and sustaining innovations*. Santa Monica, CA: Rand Corp.

Brown, C. H. (1993). Analyzing preventive trials with generalized additive models. *American Journal of Community Psychology*, 21, 635–664.

Brown, C. H., Mohr, D. C., Gallo, C., Mader, C., Palinkas, L. A., Wingood, G., . . . Jacobs, C. (2013). A computational future for preventing HIV in minority communities: How advanced technology can improve implementation of effective programs. *Journal of Acquired Immune Deficiency Syndromes, 63*(Suppl. 1), 72-84.

Brown, C. H., Wang, W., Kellam, S. G., Muthen, B. O., Petras, H., Toyinbo, P., . . .Windham, A. (2008). Methods for testing theory and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug and Alcohol Dependence, 95*(Suppl. 1), 74–104.

Burns, B. J., & Hoagwood, K. (Eds.). (2005). Evidence-based practices Part II: Effecting change. *Child and Adolescent Psychiatric Clinics of North America*, 14(2), xv–xvii.

Chen, H. (1998). Theory-driven evaluations. In H. J. Walberg (Ed.), *Advances in educational productivity: Evaluation research for educational productivity* (Vol. 7, pp. 15–34). London, England: JAI Press.

Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., . . .Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology*, 14, 317–345.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J., Hoagwood, K., Buckley, J., Olin, S., . . .Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, 1, 6–28.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, 18, 237–256.

Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research*, 20, 308–313.

Elliott, D. S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science*, 5, 47–52.

Evertson, C. M., & Weinstein, C. S. (2006). Classroom management as a field of inquiry. In C. Evertson & C. Weinstein (Eds.), *The handbook of classroom management: Research, practice, & contemporary issues* (pp. 3–15). Mahwah, NJ: Lawrence Erlbaum

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Education Research Journal*, 38, 915–945.

Giacquinta, J. B. (1973). The process of organizational change in schools. In F. N. Kerlinger (Ed.), *Review of research in education* (pp. 178–208). Oxford, UK: F.E. Peacock.

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, 93, 1261–1267.

Greenberg, M. T., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention & Treatment*, 4(1).

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). *Diffusion of innovations in health service organizations: A systematic literature review*. Oxford, UK: Blackwell.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology*, *33*, 665–679.

Hargreaves, A., & Fullan, M. G. (1992). *Understanding teacher development*. London, England: Cassell.

Hasbrouck, J., & Denton, C. (2005). *The reading coach: A how-to manual for success*. Denver, CO: Sopris West.

Ialongo, N., Poduska, J., Werthamer, L., & Kellam, S. (2001). The distal impact of two first grade preventive interventions on conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders*, *9*, 146–160.

Ialongo, N. S., Werthamer, L., Kellam, S. G., Brown, C. H., Wang, S., & Lin, Y. (1999). Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. *American Journal of Community Psychology*, *27*, 599–641.

Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, *10*, 165–185.

Kellam, S. G., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., Petras, H., . . .Wilcox, H. C. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence*, *95*, S5–S28.

Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *Academy of Management Review*, *21*, 1055–1080.

Leach, D. J., & Conto, H. (1999). The additional effects of process and outcome feedback following brief in-service teacher training. *Educational Psychology*, *19*, 441–462.

Leavitt, H. J. (1965). Applied organizational change in industry. In J. G. March (Ed.), *Handbook of organizations* (pp. 1144–1170). New York, NY: Rand McNally.

Lieberman, A. (1996). Practices that support teacher development: Transforming conceptions of professional learning. In M. W. McLaughlin & I. Oberman (Eds.), *Teacher learning: New policies, new practices* (pp. 185–201). New York, NY: Teachers College Press.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Revaluation and Policy Analysis*, *15*, 129–151.

Lochman, J. E., Boxmeyer, C., Powell, N., Qu, L., Wells, K., & Windle, M. (2009). Dissemination of the coping power program: Importance of intensity of counselor training. *Journal of Consulting and Clinical Psychology*, *77*, 397–409.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

McCormick, L. K., Steckler, A. B., & McLeroy, K. R. (1995). Diffusion of innovations in schools: A study of adoption and implementation of school-based tobacco prevention curricula. *American Journal of Health Promotion*, *9*, 210–219.

McLaughlin, M. W. (1990). The rand change agent study revisited: Macro perspectives and micro realities. *Educational Researcher*, *19*(9), 11–16.

National Council for Accreditation of Teacher Education. (2008). *Professional standards for the accreditation of teacher preparation institutions*. Washington, DC: Author. Retrieved from http://www.ncate.org/LinkClick.aspx?fileticket=nX43fwKc4Ak%3D&;tabid=474

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly*, *12*, 77–88.

Owens, R. G. (2004). *Organizational behavior in education: Adaptive leadership and school reform* (8th ed.). Boston, MA: Pearson.

Petras, H., Kellam, S. G., Brown, C. H., Muthen, B. O., Ialongo, N. S., & Poduska, J. M. (2008). Developmental epidemiological courses leading to antisocial personality disorder and violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug and Alcohol Dependence*, *95*, S45–S59.

Poduska, J., Gomez, M., Capo, Z., & Holmes, V. (2012). Developing a collaboration with the Houston Independent School District: Testing the generalizability of a partnership model. *Administration and Policy in Mental Health and Mental Health Services Research*, *39*, 258–267.

Poduska, J., Kellam, S., Brown, C. H., Ford, C., Windham, A., Keegan, N., & Wang, W. (2009). Study protocol for a group randomized controlled trial of a classroom-based intervention aimed at preventing early risk factors for drug abuse: Integrating effectiveness and implementation research. *Implementation Science*, *4*, 56.

Poduska, J. M., Kellam, S. G., Wang, W., Brown, C. H., Ialongo, N. S., & Toyinbo, P. (2008). Impact of the good behavior game, a universal classroom-based behavior intervention, on young adult service use for problems with emotions, behavior, or drugs or alcohol. *Drug and Alcohol Dependence*, *95*, S29–S44.

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

Ringwalt, C. L., Vincus, A., Ennett, S., Johnson, R., & Rohrbach, L. A. (2004). Reasons for teacher's adaptation of substance use prevention curricula in schools with non-White student populations. *Prevention Science*, *5*, 61–67.

Robertson, P. J., Roberts, D. R., & Porras, J. I. (1993). Dynamics of planned organizational change: Assessing empirical support for a theoretical model. *Academic of Management Journal*, *36*, 619–634.

Rogers, E. M. (2003). *Diffusion of innovations* (2nd ed.). New York, NY: The Free Press.

Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation: Transporting prevention interventions from research to real-world settings. *Evaluation & the Health Profession*, *29*, 302–333.

Rose, D. J., & Church, R. J. (1998). Learning to teach: The acquisition and maintenance of teaching skills. *Journal of Behavioral Education*, *8*, 5–35.

Spouse, J. (2001). Bridging theory and practice in the supervisory relationship: A sociocultural perspective. *Journal of Advanced Nursing*, *33*, 512–522.

Taylor, J. E. (2007). Instructional coaching: The state of the art. In M. M. Mangin & S. R. Stoelinga (Eds.), *Effective teacher leadership: Using research to inform and reform* (pp. 10–35). New York, NY: Teachers College Press.

Walter, H. J., Gouze, K., & Lim, K. G. (2006). Teachers' beliefs about mental health needs in inner city elementary schools. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*, 61–68.

Wehby, J. H., Maggin, D. M., Moore Partin, T. C., & Robertson, R. (2012). The impact of working alliance, social validity, and teacher burnout on implementation fidelity of the good behavior game. *School Mental Health*, *4*, 22–33.

Wilcox, H. C., Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., & Anthony, J. C. (2008). The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempts. *Drug and Alcohol Dependence*, *95*, S60–S73.