CrossMark

Classroom Management in an Urban, Alternative School: a Comparison of Mindfulness and Behavioral Approaches

Anna C. J. Long[1] · Tyler L. Renshaw[1] · Devon Camarota[2]

Abstract

Managing classroom behavior is an important prerequisite to effective teaching and a salient need in alternative schools. Unfortunately, students from these schools are often underrepresented in the intervention literature. The primary aim of this study was to compare the effectiveness of two different theoretical approaches to classroom management, one behavioral (i.e., the good behavior game) and the other mindfulness-based (i.e., mindfulness skills training), with a sample of fifth-grade, predominantly African American students from an urban, high-poverty alternative school. The study examined the effectiveness of the two interventions in comparison to each other and a treatment-as-usual control using a quasi-experimental group design with blocked random assignment. Results revealed that neither intervention led to significant improvements in student internalizing behavior, externalizing behavior, or wellbeing. Though, some practically meaningful treatment effects were found through examination of effect sizes. Mindfulness skills training was the only condition to yield meaningful pre–post change in student outcomes, including a moderate therapeutic effect for externalizing behavior and an iatrogenic effect with respect to student wellbeing. These findings provide preliminary evidence that mindfulness skills training might have differential effects on student mental health outcomes, compared with education as usual and a traditional classwide behavioral intervention. Additionally, study findings make clear the importance of careful deliberation when transporting evidence-based interventions to unique student populations and intervention contexts.

Keywords Classroom management · Good behavior game · Mindfulness · Universal prevention

One of the greatest challenges of teaching is managing classroom behavior. Student off-task and disruptive behaviors are commonplace (Richards et al. 2010; Snider et al. 2002). Problem behaviors in classrooms result in a variety of negative outcomes for students (e.g., Beebe-Frankenberger et al. 2004; Broidy et al. 2003; Reinke et al. 2011): reduced instructional time, diminished academic achievement, and increased risk for grade retention and short- and long-term conduct problems. Furthermore, students' disruptive behaviors result in greater teacher stress, burnout, job dissatisfaction, and attrition (Hastings and Bham 2003; U.S. Department of Education, National Center for Education Statistics 2005).

The critical need to manage classroom behavior is especially salient in alternative and high-poverty schools given the higher rates of problem behavior and the fact that their student populations have an elevated number of risk factors (Lehr et al. 2009; Simonsen et al. 2010; Thomas et al. 2008). Alternative education can be broadly defined as public "schools or programs designed to address the needs of students that typically cannot be met in regular schools" (Carver and Lewis 2010, p. 1). Students enrolled in alternative education have higher rates of disability diagnosis and special education placement (Lehr et al. 2009). As well, they are at significant risk for school failure due to a host of wide-ranging factors including, but not limited to, poor academic achievement, truancy, and significant behavioral/conduct problems (Carver and Lewis 2010). High-poverty schools are defined as those having 75% or more of their students who qualify for free or reduced lunch (Aud et al. 2013). Like alternative schools, students within these schools are at elevated risk for low academic achievement and other behavioral challenges that relate to school dropout/withdrawal (Aud et al. 2013; McFarland et al. 2017; Thomas et al. 2008).

✉ Anna C. J. Long
along@lsu.edu

[1] Department of Psychology, Louisiana State University, 236 Audubon Hall, Baton Rouge, LA 70803, USA

[2] E.J. Ourso College of Business, Business Education Complex, Louisiana State University, Baton Rouge, LA 70803, USA

☶ Springer

Unfortunately, the enrollment of students in alternative and high-poverty schools has been increasing in recent years (Aud et al. 2013; Lehr et al. 2009). According to the National Center for Education Statistics (NCES), the number of students enrolled in alternative schools increased by more than 5% between 2002 and 2008 (Carver and Lewis 2010; Kleiner et al. 2002). Currently, about one in five schools are considered high-poverty nationally, up from one in eight schools (Aud et al. 2013). Schools categorized as both alternative and high-poverty also seem to be on the rise. Recent evidence suggests that at least 35% of alternative schools have a sizeable concentration of students living in poverty (Carver and Lewis 2010). Taken together, high-poverty alternative schools serve a population of students with significant academic and behavioral needs. This student population seems to be rising. Consequently, these schools play a vital role in ensuring that a growing number of our nation's most at-risk students are given every opportunity for educational success. Effective classroom management plays a central role in ensuring this success, given that students return to traditional community schools is largely based on the improvement shown in their behavior (Kleiner et al. 2002) The present study extends the literature on classroom behavior management by examining the effectiveness of two different theoretical approaches to classroom management: one based on behavioral principles and the other on mindfulness-based practices, with a sample of fifth-grade, predominantly African American students from an urban, high-poverty alternative school.

Effective Classroom Management

Classroom management involves teacher classwide implementation of purposeful procedures to create an environment conducive to learning (Ysseldyke et al. 2004). Effective classroom management includes both the regulation of student problem behaviors as well as the promotion of prosocial behaviors such as active engagement (Mitchell and Brashaw 2013). Limitations and differing theoretical orientations remain in research examining the effectiveness of behavior management interventions targeting the whole classroom (Emmer and Sabornie 2015; Richards et al. 2010). Despite this, a great deal of research has revealed three broad principles of effective classroom management (Good and Brophy 2008; Simonsen et al. 2015; Simonsen et al. 2008). First, teachers should focus on proactive strategies that promote desired behaviors, rather than reactionary, punitive ones that target undesirable behaviors. Second, teachers should provide a predictable structure and have a clear set of classroom expectations and routines. Third, teachers should stimulate the development of self-regulation skills through consistent responses to student appropriate and inappropriate behaviors.

Interventions included in meta-analytic reviews on classwide behavior interventions overwhelmingly align with the three principles outlined above and have demonstrated overall positive, large effect sizes for student prosocial and problem behaviors (Oliver et al. 2011; Chaffee et al. 2017). Unfortunately, to date, one major weakness of meta-analytic studies on classroom management is the limited investigation of potential moderating variables, especially variables that might inform recommendations for intervention with specialized student populations in unique school contexts (e.g., proportion of students in sample who are in special education, from racial/ethnic minority backgrounds, or eligible for free or reduced lunch).

Behavioral Approach to Classroom Management

A wealth of empirical evidence demonstrates the effectiveness of interventions for children derived from behavioral principles (Akin-Little et al. 2009). Lipsey and Wilson (1993) have provided one of the most comprehensive reviews of this research. In their review, they examine 300 meta-analytic studies on the efficacy of behavioral interventions (and other treatments) with children. These meta-analyses reported medium to large treatment effects across a broad range of child outcome domains. Meta-analytic methods have also shown behavioral interventions to be superior to non-behavioral interventions (Casey and Berman 1985; Weisz et al. 1987). Moreover, behavioral interventions conducted in schools have been found to significantly reduce student disruptive behaviors and often yield similar treatment effect sizes as child psychotherapy (ES ≥ 80; see Stage and Quiroz 1997 for review).

Good Behavior Game A commonly recommended classwide management intervention is the good behavior game (GBG). The GBG promotes positive classroom behavior by making reinforcers accessible to students' if their behavior, together with the behavior of others in their group, meets a preset criterion. In the original empirical evaluation by Barrish et al. (1969), a fourth-grade, general education classroom was divided into two teams. Next, explicit behavioral rules were created, along with contingencies for following those rules. Disruptive behavior by any member of a team resulted in a mark on the board. These marks accumulated toward a possible loss of privileges for the whole team. If a team had the fewest marks, or if neither team received more than five marks during an intervention session, the team(s) would earn a privilege. Barrish et al.'s study found that implementation of the game resulted in a significant reduction in out-of-seat and talking-out behaviors.

The GBG has a well-established line of research supporting its efficacy (Embry 2002; Tingstrom et al. 2006). Studies have demonstrated the effectiveness of the intervention across varying grade levels (e.g., preschool, elementary and high school), student demographic characteristics (e.g., socioeconomic status, urbanicity, racial/ethnic background, nationality), and disabilities or exceptionality statuses (e.g., general education, special education, intellectual/developmental disability, emotional or behavioral disorder; cf. Bowman-Perrott et al. 2016; Embry 2002; Flower et al. 2014a, b; Tingstrom et al. 2006). The GBG has been shown to improve a wide variety of target behaviors (e.g., on-task/off-task, appropriate and inappropriate social interactions, impulsive, aggressive, disruptive, swearing or negative comments, talking out, substance use, antisocial; Flower et al. 2014a, b; Tingstrom et al. 2006). In addition, beyond the immediate positive effects of the game, research suggests the intervention produces long-term positive outcomes for problem behavior (Kellam et al. 2008). Some researchers have adapted the game procedures to focus on the positive by awarding points to teams for meeting expectations, as opposed to counting rule violations (Flower et al. 2014a, b); although, research is lacking on the effectiveness of this modification in intervention conditions like that under investigation in this study.

Given the strength and breadth of research on the GBG, many behavioral experts consider it a gold standard intervention for classroom management (Embry 2002). Despite this, there remains a dearth of research regarding the effectiveness of the intervention among students in alternative school settings (Sy et al. 2016). To date, only three studies have examined the GBG in this context (Joslyn et al. 2014; Salend et al. 1989; Sy et al. 2016). All three studies implemented variations of the game with students who displayed severe behavioral challenges and found significant reductions in their problem behavior. Studied classrooms spanned from mid-elementary to high school. Unfortunately, none of the studies provided data indicating the game had been investigated in a high-poverty alternative school. Additionally, only one study examined the game with a racially/ethnically diverse sample (i.e., Joslyn et al. 2014). Therefore, more research is needed to delineate under what conditions and for whom the GBG maintains (or does not maintain) effectiveness as originally designed.

Mindfulness Approach to Classroom Management

Within the last decade, MBIs have gained widespread recognition in education. In general, MBIs are any interventions that seek to activate the practice of mindfulness toward improving individuals' functioning. Mindfulness practice originates from Eastern religious traditions, most notably Buddhism (Kabat-Zinn 2003), and may be described as an intentional practice of regulating one's attention on present stimuli, coupled with a nonjudgmental positive response to those stimuli (Renshaw et al. 2015a). There is a rapidly emerging literature base on the effectiveness of MBIs with children and adolescents (Burke 2010; Renshaw et al. 2015a). Meta-analytic studies have found that MBIs for youth are often delivered in schools and consistently demonstrate small to moderate positive treatment effects across wide-ranging outcomes (e.g., attentional, cognitive, academic, mental health, social, behavioral, and physiological; Kallapiran et al. 2015; Klingbeil et al. 2017a, b; Zenner et al. 2014; Zoogman et al. 2015). MBIs have been found to demonstrate effects comparable to or better than other active youth interventions (Kallapiran et al. 2015). Moreover, some data suggest the positive effects of MBIs remain stable or further improve at follow-up (Kallapiran et al. 2015; Klingbeil et al. 2017a, b). The effectiveness of these interventions has also shown generalizability across a diverse range of clients (with respect to age, gender, clinical/disability status, and race/ethnicity), implementers (e.g., parents, clinicians, school personnel, researchers), and settings (e.g., school, clinical), as well as when administered via indirect versus direct approaches (i.e., intervening with the caregiver as opposed to directly with the youth).

Despite these research advancements, several factors remain unclear about MBIs for youth. Namely, it is unknown what specific mechanism(s) drive the change in outcomes observed or what variables moderate treatment effects? Though, there appears to be a consensus that effect sizes are significantly larger for clinical versus non-clinical populations (Zoogman et al. 2015). Similar to the GBG, there is a limited research base on the effectiveness of MBIs in alternative school settings. To date, three studies have been conducted in alternative schools (Bluth et al. 2016; Fishbein et al. 2016; Wisner and Norton 2013). Although they revealed promising results, none of these interventions were classroom management approaches.

Mindfulness Skills Training To the authors' knowledge, there are no MBIs designed and tested specifically as a classroom management intervention. However, several MBIs for youth have been administered in a group or classwide format and have shown benefits for improving classroom behavior (cf. Felver et al. 2016), including building self-regulatory skills and increasing active engagement (e.g., Black and Fernando 2014). Although there is no consensus regarding the theoretical mechanisms of MBIs, experts suggest that by enhancing individuals' ability to acknowledge present moment events without becoming overly attached to them, these interventions help students respond in an adaptive manner (Baer 2003; Bishop et al. 2004; Burke 2010; Shapiro et al. 2006). Consistent with this, the present study based the design of

the mindfulness approach to classroom management on a widely recognized two-component model put forth by Bishop et al. (2004). Bishop and colleagues defined mindfulness as "the self-regulation of attention so that it is maintained on immediate experience" and "adopting a particular orientation toward one's experiences in the present moment, an orientation characterized by curiosity, openness, and acceptance" (p. 232). The cultivation of this skill set is a central component of all MBIs for youth and is commonly referred to as mindfulness skills training (MST).

To date, MST has been developed through the application of meditation-based techniques. These techniques directly align with Bishop et al. (2004) model by having the individual practice focusing his/her mind on the present moment experience, typically by directing attention to a specific anchor (e.g., breath), and then helping the individual accept this experience so that reactive or maladaptive responding can be avoided (Meiklejohn et al. 2012). This MST may result in the desensitization of conditioned responses, especially when practiced repeatedly in vivo (Baer 2003; Shapiro et al. 2006). Based on this literature, Renshaw et al. (2015a) have recommended a simple mindful STOP procedure for use in schools that is analogous to a common MST exercise, informal mindfulness practice. Informal mindfulness practice refers to MST that is integrated into one's everyday activities (Meiklejohn et al. 2012). The mindful STOP procedure outlines basic steps students can follow to be mindful in the classroom including stopping (S), taking three deep breaths (T), observing what is happening right now (O), and proceeding positively (P).

Study Purpose

The purpose of this study was to compare the effectiveness of two different theoretical approaches to classroom management, one behavioral (i.e., the GBG) and the other mindfulness-based (i.e., MST), with a sample of high-risk students underrepresented in the school intervention literature. Specifically, this study examined the effectiveness of the GBG and MST in comparison to each other and a treatment-as-usual control with a sample of fifth-grade, predominantly African American students from an urban, high-poverty alternative school. To date, there have been no studies of the GBG and only one investigation of a MBI in this unique school setting. Thus, this study was designed to shed light on the generalizability of previous findings regarding these interventions to students from high-poverty alternative schools who have a combination of salient risk factors (e.g., significant academic delays, heightened frequency and intensity of disruptive behaviors, low socioeconomic status). Given previous research, we hypothesized that compared to students in the control condition, those in both the GBG and MST conditions would evidence positive therapeutic changes in behavior across all

outcomes of interest: internalizing problems, externalizing problems, and wellbeing. We had no data to inform a hypothesis for differential effects between the two treatments. A secondary aim of this study was to address a gap in the existing literature on MBIs for youth by singling out a common core component of these interventions for testing in isolation (i.e., the mindful STOP procedure). In so doing, practitioners and scientists may gain a better understanding of one potential active ingredient in MBIs that is responsible for changes in treatment outcomes.

Method

Setting and Participants

Participants were 73 students from an urban, alternative, elementary school located in a large city in the Southeastern part of the USA. Large city is defined as a "territory inside an urbanized area and inside a principal city with population of 250,000 or more" (National Center for Education Statistics 2006). The alternative school is a public educational setting that provides programming for students who are well behind grade level (two or more years behind) and have significant disciplinary problems. Student participants were in six, fifth-grade classrooms reflecting 90-min social studies blocks (A–F). Their teacher, Ms. Cooper (name replaced to maintain confidentiality), was a White female in her second year of teaching. Each block of students was taught two times per week. Students ranged in age from 10 to 14 years old ($M = 11.69$). The majority of students were male (65.8%) and identified as African American (98.6%; see Table 1). With regard to school diversity characteristics, 100% of the students were eligible for free or reduced lunch and no students in the school had English as a second language.

Data included in this study represent analysis of a de-identified secondary dataset, following provision of consultation for classroom management support. As part of the consultative relationship, Ms. Cooper expressed interest in evaluating the effectiveness of two separate theoretical approaches to classroom management, one based on behavioral principles and the other based on mindfulness-based practices. The consultative relationship was structured as a collaborative action research project aimed at generating practice-based evidence for the field. Thus, the study served the dual-purpose of addressing practice in the short-term as well as informing the field about the generalizability of the GBG and MST (Fraenkel et al. 2012, p. 590). Consultation was provided related to the selection of evidence-based strategies representing each theoretical orientation. Consultation services also informed the selection of measures and methodological design which would permit evaluation of the two

Table 1 Sample demographics by class block

	Block A[a] (n = 12)	Block B[b] (n = 12)	Block C[b] (n = 13)	Block D[a] (n = 14)	Block E[c] (n = 13)	Block F[c] (n = 9)
Gender						
Male	12	0	13	14	0	9
Female	0	12	0	0	13	0
Race						
African American	12	12	13	14	13	8
White	0	0	0	0	0	1

[a] treatment-as-usual control

[b] good behavior game treatment condition

[c] mindfulness skills training treatment condition

interventions within the confines of Ms. Cooper's context and responsibilities.

Prior to commencing the new classroom management approaches, Ms. Cooper sent out a notification letter to parents/guardians which (a) overviewed the nature and extent of the consultative relationship and (b) provided an overview of the classroom improvement project. Also included in this parent notification letter was information about the data that would be collected and how students' confidentiality would be maintained. Ms. Cooper provided an opportunity for all parents to opt-out of having their children's data included in intervention evaluation analyses.

Procedure

A quasi-experimental group design was used to evaluate the effectiveness of each intervention independently, and then in comparison to the alternate classroom management approach. First, the six blocks (or classes) of students were divided into two groups according to the days they were taught on (either Mondays and Tuesdays or Thursdays and Fridays). This was done to increase feasibility of implementation, reduce the chance of cross-condition contamination, and ensure nearly equivalent numbers of males and females across intervention conditions. Next, the two groups (i.e., Monday/Tuesday classes and Thursday/Friday classes) were randomly assigned to one of two intervention conditions: GBG or MST. Within each intervention condition, one block was randomly selected to serve as the treatment-as-usual control. As a result of this process, blocks A, B, and C comprised group one and were assigned to the GBG condition, with block A randomly selected as the control. Blocks D, E, and F comprised group two and were assigned to the MST condition, with block D randomly selected as the control. Although Ms. Cooper could not be blind to the study conditions, she was blind to the researchers' a priori hypotheses to minimize threats to the internal validity of the study.

Pretest and posttest data were collected via teacher-report and student-report, with a four-week intervention period between each data collection point. Immediately following pretest data collection, Ms. Cooper underwent direct training. The training was conducted over one 60-min session by two doctoral students in school psychology. The training included (1) didactic instruction on the theory and procedures of each intervention, (2) directions on preparing for the first day of implementation, and (3) directions on how to continue intervention implementation after day one. Additionally, the training provided an opportunity to practice intervention components or have them modeled, as needed. The training concluded when Ms. Copper expressed comfort with intervention implementation.

Since the IRB did not permit direction observations for this study, Ms. Cooper was provided with intervention materials and trained on how to track intervention fidelity reliably and validly. Recent research suggests that the provision of direct training on self-monitoring of one's implementation results in reliable self-assessment and increases the likelihood of teachers maintaining appropriate levels of intervention fidelity (Mouzakitis et al. 2015; Simonsen et al. 2013). Throughout the study, investigators followed up with Ms. Cooper weekly via telephone to check-in regarding implementation and provide support. These check-ins followed a semi-structured guide which included (a) gaining information about how the interventions went over the past week, (b) reviewing intervention steps and inquiring about implementation, (c) addressing teacher questions or concerns, and (d) establishing a time for the next check-in as well as confirming the teacher's commitment to intervention fidelity.

Interventions

The GBG and MST interventions were implemented simultaneously in designated classrooms across a four-and-a-half-week period leading up to the winter school break. Given

the high level of student needs at the school, classrooms were small, highly-structured environments with fewer than 15 students per class. The GBG was implemented with group one treatment classrooms (blocks B and C), while the MST intervention was implemented with group two treatment classrooms (blocks E and F). Each intervention session lasted the duration of the class block (i.e., 90 min) and occurred each time during the week that the class met for social studies with Ms. Cooper. Most often, this resulted in treatment classes receiving intervention two times per week. In contrast, throughout the intervention phase, control classrooms continued to receive treatment as usual. Specially, Ms. Cooper continued to apply a "three strikes" procedure which involved students receiving a warning after the first infraction, a seat change after the second infraction, and finally a phone call home and/or office discipline referral following the third infraction.

Good Behavior Game The GBG is an interdependent group contingency that targets the reduction of classroom disruptions through application of behavioral principles. With interdependent group contingencies, all individuals who comprise the group receive access to the reinforcing consequence based on the group's behavior meeting a preset criterion (Skinner et al. 2009). Procedures for the GBG in the present study approximated those outlined by Barrish et al. (1969). Based on this original study and subsequent investigations of the GBG (i.e., Medland and Stachnik 1972), the core components of the GBG have been identified and include the following: (a) identifying, defining, teaching, and posting behavioral expectations or rules; (b) defining what constitutes rule violations; (c) designating an interval during which the game will be played; (d) dividing the classroom of students into equal teams (typically two); (e) explaining to the teams how a team may win the game by committing the fewest rule violations (or by remaining below a preset criterion of rule violations, enabling multiple teams to win); (f) identifying and administering a daily privilege or reward to the team (or teams) that win the game; and (g) maintaining a visible scoreboard that logs team rule violations (Embry 2002; Flower et al. 2014a, b).

Consistent with past use of the GBG, each block of students (B and C) receiving this intervention were divided into two teams with a roughly equal proportion of students who were disruptive and better behaved. This helped enhance the chance that both teams would have an opportunity of being successful and winning the game. To introduce the game, Ms. Cooper explained that (a) they would be implementing a new procedure in the classroom to support students' engagement in appropriate behaviors, (b) the class would be divided into two teams and each team would be assigned a rotating team leader, and (c) teams would earn rewards for appropriate behavior by limiting their number of class rule violations to below a pre-established goal (e.g., 5). Therefore, for each intervention

session, both teams in the classroom had the opportunity to win the game and receive a reinforcer. A list of reinforcers was created using a student reinforcer preference assessment. When a team (or teams) won, the teacher randomly selected one of the reinforcers from the list.

Prior to starting the game, Ms. Cooper posted and taught the classroom rules as well as the winner scoreboard (i.e., board that tracks rule violations). Classroom rules included following directions, maintaining one's personal space, being responsible and ready to learn, and being respectful. Each of these rules were operationalized for the class and taught in a tell-show-do format. Prior to each intervention session, Ms. Cooper was responsible for reviewing the class rules and reminding students of the goal and reward their teams were working toward. Throughout the intervention sessions, Ms. Cooper provided specific praise to students who were engaging in appropriate behavior and responded to rule violations by labeling the behavior, marking a tally on the scoreboard, and briefly redirecting the student to engage in desired behavior. At the end of each intervention session, Ms. Cooper gave teams feedback regarding their performance and provided a reward, if the goal was met.

The GBG was comprised of four intervention components. Following all intervention sessions, each intervention component was rated by Ms. Cooper along a three-point scale reflecting her adherence to the component ($0 = did\ not\ implement$, $1 = partially\ implemented$, $2 = fully\ implemented$). To enhance the reliability of teacher self-ratings, each rating (0, 1, 2) corresponded to a detailed behavioral anchor for the intervention component (e.g., for component two, $2 = fully\ implemented$, *set goal*, *and reviewed contingency to earn reward*). Classrooms in the GBG condition received a total of 9 intervention sessions. The mean level of intervention fidelity for all sessions was 93% (SD = 13%).

Mindfulness Skills Training The MST is an intervention that targets the reduction of student disruptions through building of skills in mindfulness. Procedures for the MST intervention were developed to parallel those for the GBG with respect to the number of intervention components (i.e., 4), the specific rules for the class (see above), and the organizational structure of delivery. This parallel structure was used so that the focus of the comparison between the two interventions was on differences in effectiveness related to the underlying theoretical approach, as opposed to differences related to other factors such as dosage or intervention complexity.

The MST intervention for the present study was structured as an informal mindfulness practice that targeted the development of mindful awareness and responsivity emphasized in Bishop et al. (2004) two-component model (cf. Renshaw 2012; Renshaw and O'Malley 2014). The mindful STOP strategy was taught to students to help them regulate their behavioral-emotional responses. Specifically, when a student

or students were engaging in inappropriate behavior, they were asked to (1) "stop," (2) "take three deep breaths," (3) "observe self and others," and (4) "proceed positively." This process was intended to help students disrupt their pattern of reactive responding by first getting them to focus their attention, and then pause to identify what thoughts and feelings might prevent them from proceeding appropriately. To introduce the strategy, Ms. Cooper explained that (a) they would be implementing a new procedure in the classroom to help students' build skills to manage their behavior and emotions and (b) they would discuss common barriers to being "at their best" and learn about how they can apply mindfulness skills to be resilient and overcome these barriers. Prior to starting the intervention, Ms. Cooper posted and taught the four classroom rules. Consistent with the GBG condition, each of these rules were operationalized for the class and taught in a tell-show-do format. Prior to each intervention session, Ms. Cooper was responsible for reviewing the class rules and reminding students that the goal of the intervention was for them to build mindfulness skills to effectively manage their behavior and emotions. Throughout the intervention sessions, Ms. Cooper provided specific praise to students who were engaging in appropriate behavior and responded to rule violations by labeling the behavior, marking a tally on the board, instructing the student (or students) who were engaging in problem behavior to use mindful STOP. At the end of each intervention session, Ms. Cooper gave students feedback regarding their performance and encouraged them to continue to use the mindfulness skills they were learning in and outside of the classroom.

Consistent with the GBG condition, following each intervention session, Ms. Cooper rated her adherence to the MST intervention components along a three-point scale for each component (0 = *did not implement*, 1 = *partially implemented*, 2 = *fully implemented*). Each rating corresponded to a specific behavioral anchor for the intervention component to enhance the reliability of teacher self-ratings. Classrooms in the MST condition received a total of 8 intervention sessions. There was one less intervention session for this condition due to an unanticipated school event. The mean level of intervention fidelity for all sessions was 80% (SD = 19%).

Measures

Student Internalizing Behavior Screener and Student Externalizing Behavior Screener

The student internalizing behavior screener (SIBS) and student externalizing behavior screener (SEBS) (Cook 2013; Cook et al. 2011) are teacher-report measures of students' internalizing and externalizing symptoms, respectively. Both measures are comprised of seven items reflecting problematic student behaviors at school. Items are rated along a four-point response scale that represents the relative frequency of behavior (0 = *never*, 1 = *rarely/*

seldom, 2 = *occasionally/ moderately*, 3 = *frequently/almost always*). Thus, higher scores suggest higher rates of problematic behavior. Previous research has indicated that the SIBS and SEBS have adequate-to-strong internal consistency reliability (Cronbach's α range = .79–.89) and good test–retest reliability (r range = .74–.92). Furthermore, concurrent validity has been demonstrated by strong associations between the SIBS and the internalizing scale of the *Achenbach System of Empirically Based Assessment* (ASEBA; Achenbach and Rescorla 2001) Teacher Report Form (TRF; $r = .82$) and the SEBS and the externalizing scale of the ASEBA–TRF ($r = .87$; Cook 2013). Analyses of the pretest data from the present sample indicated that responses to the SIBS and SEBS were relatively normally distributed (skewness and kurtosis $< |1|$), evidenced at least adequate internal consistency reliability ($\alpha = .94$ and .73, respectively), and were characterized by homogeneity of variance across conditions (Levene's statistics $p > .05$). The correlation between the pretest SIBS and SEBS scores for the total sample was negative and small ($r = - .10, p > .05$).

Student Subjective Wellbeing Questionnaire

The Student Subjective Wellbeing Questionnaire (SSWQ) (Renshaw et al. 2015b) is a 16-item, self-report measure of students' subjective wellbeing at school. The measure is comprised of four, four-item subscales: Joy of Learning (one's experience of positive emotions and cognitions when engaged in academic tasks), School Connectedness (one's feeling of being cared for and relating well to others at school), Educational Purpose (one's appraisal of school and academic tasks as important and meaningful), and Academic Efficacy (one's perception of his/her academic behaviors effectively meeting school demands). All items are directly phrased and rated along a four-point response scale that represents the relative frequency of subjective experiences (1 = *almost never*, 2 = *sometimes*, 3 = *often*, 4 = *almost always*). The subscales can be combined to form an overall Student Wellbeing composite scale. Previous research has demonstrated that the SSWQ has a psychometrically sound measurement model with adequate internal consistency reliability for all scales (Cronbach's α range = .72–.88), and that responses to both the subscales and the overall composite scale show meaningful associations with a variety of self-reported behavioral health outcomes and school-reported academic indicators (Arslan and Renshaw 2017; Renshaw 2015; Renshaw and Arslan 2016; Renshaw and Chenier 2016). Although the intention of the present study was to examine intervention effects at the level of the constructs represented by the subscales, analyses on the pretest data from the present sample indicated that internal consistency estimates were generally poor across all subscales (joy of learning $\alpha = .65$, school connectedness $\alpha = .50$, educational purpose $\alpha = .39$, academic efficacy $\alpha = .65$). However, pretest responses to the composite scale yielded strong internal consistency reliability ($\alpha = .81$),

and so this overall measure was selected to represent the student wellbeing outcome. Further analysis of pretest data indicated that responses to the wellbeing composite scale were relatively normally distributed (skewness and kurtosis < |1|), were characterized by homogeneity of variance across conditions (Levene's statistics $p > .05$), and had negligible correlations with the pretest SIBS scores ($r = -.03$, $p > .05$) and SEBS scores ($r = .01$, $p > .05$).

Missing Data

Processing the self-reports and teacher-reports into an electronic database revealed a substantial amount of missing data at both pretest and posttest. For the pretest data, one student from the GBG condition was missing all SIBS and SEBS data (1.4% overall missing), whereas 19 students were missing all SSWQ data (GBG $n = 3$, MST $n = 8$, EAU $n = 8$; 26% overall missing). For the posttest data, 9 students were missing all SIBS and SEBS data (GBG $n = 2$, MST $n = 2$, EAU $n = 5$; 12.3% overall missing), whereas 18 students were missing all SSWQ data (GBG $n = 3$, MST $n = 7$, EAU $n = 8$; 24.7% overall missing). According to the What Works Clearinghouse (2017), overall rates of missingness and attrition within the present study fall within the range of a "tolerable threat of bias" ($\leq 55\%$); although, the differential rates of missingness and attrition between groups was suboptimal. That said, all obtained data were deemed to be complete, with no non-responses observed for any items on either teacher-reports or self-reports. Upon further inspection, students with missing data at pretest were deemed to be distinct from those with missing data and posttest, suggesting that the missing data was not random. Follow-up with Ms. Cooper indicated that this was indeed the case, as she confirmed that missing data on the teacher-report measures was the result of teacher error ($n = 7$) or students transferring to another school ($n = 3$). Ms. Cooper also confirmed that all missing self-report data were the results of students being absent during the administration of the SSWQ at pretest or posttest, due to tardiness, absenteeism, suspensions, or being pulled from the classroom for some other purpose. The teacher also noted that student attendance and retention was a challenge for the school more broadly. Given that all students had largely shared demographic characteristics (i.e., from the same school, in the same grade level, taught by the same teacher, of similar racial/ethnic identity and economic background, and experiencing substantial academic and behavioral challenges), data missingness could not have varied as a function of these variables. Unfortunately, the anonymous nature of the study did not permit follow-up on these missing data, and prevented against the collection of individual attendance data that may have been used to evaluate intervention exposure and dosage.

Missingness on self-report measures was examined in relation to scores from teacher-report measures, and vice versa,

for both time points. Specifically, an independent samples t test comparing students with complete and missing SSWQ data at pretest indicated non-significant differences in pretest SIBS scores ($t[69] = -1.74$, $p = .09$, $M_{\text{Diff.}} = -1.93$, $d = 0.37$) and SEBS scores ($t[69] = -0.14$, $p = .89$, $M_{\text{Diff.}} = -0.23$, $d = 0.04$). Similarly, independent samples t tests comparing students with complete and missing SSWQ data at posttest indicated non-significant differences in posttest SIBS scores ($t[61] = 1.45$, $p = .15$, $M_{\text{Diff.}} = 1.55$, $d = 0.44$) and SEBS scores ($t[61] = -1.08$, $p = .29$, $M_{\text{Diff.}} = -2.06$, $d = 0.32$). Given only one student was missing SIBS and SEBS data at pretest, between-group differences were not examined for this case. However, an independent samples t test comparing students with complete and missing SIBS and SEBS data at posttest did indicate a non-significant difference in posttest SSWQ composite scores ($t[53] = 0.53$, $p = .60$, $M_{\text{Diff.}} = 2.14$, $d = .23$). Consideration of the trend of resulting effect sizes from these comparisons indicates a range of small-negative to small-positive differences, suggesting no consistent behavioral profile differences for students with missing data.

To optimize the data and maximize the statistical power of the planned analyses, missing data were ultimately handled using a maximum likelihood data imputation method, which estimates plausible responses for missing values based on observed response values (see Schlomer et al. 2010). This procedure was conducted using the data imputation and stochastic regression functions within AMOS version 23.

Data Analytic Approach

Prior to conducting the primary analyses, descriptive statistics were first calculated for the pretest and posttest SIBS, SEBS, and SSWQ scores, for both the total sample and by condition. Furthermore, given that participants were assigned to an experimental condition (i.e., EAU, GBG, or MST) by class, not individually, and that Ms. Cooper reported preexisting behavioral differences between classes prior to beginning the study, analyses of variances (ANOVA) were conducted to investigate potential between-condition differences for each of the pretest outcome measures (i.e., externalizing problems, internalizing problems, and student wellbeing). Following, the primary analyses were carried out using a series of repeated-measures ANOVA to test between-condition differences in changes in outcomes (i.e., SIBS, SEBS, and SSWQ scores) from pretest to posttest. If these omnibus tests yielded significant findings, then post-hoc analyses would be conducted using t tests with Bonferonni adjustments to further locate significant differences between specific conditions. Regardless of findings related to statistical significance, the magnitude of change observed for each condition across each outcome was evaluated by calculating and comparing standardized mean differences (Cohen's d) for pretest–posttest scores. The relative magnitudes of these effect sizes were

interpreted using Cohen's (1988) longstanding decision rules: .20–.49 = small, .50–.79 = medium, .80+ = large. All preliminary and primary data analyses were conducted using SPSS version 23.

Results

Preliminary Analyses

Descriptive statistics for the SIBS, SEBS, and SSWQ scores for the total sample at both time points are presented in Table 2. Scores for all outcomes at both time points, as well as the correlation between T1–T2 scores, are presented by condition in Table 3. In general, descriptive results indicated decreases in SIBS and SEBS scores across all conditions, characterized by large positive T1–T2 correlations. Findings for SSWQ scores did not indicate a clear descriptive trend, as the EAU group showed an increase and the GBG and MST groups showed decreases, with the T1–T2 correlations ranging from negligible (MST) to moderate (EAU) to large (GBG). Furthermore, findings from the series of ANOVA investigating potential between-condition differences for each of the pretest outcome measures indicated non-significant condition effects for the SIBS ($F(2, 70) = 0.01$, $p = .99$, $\eta^2 < .001$), SEBS ($F(2, 70) = 2.73$, $p = .07$, $\eta^2 = .07$), and SSWQ ($F(2, 70) = 1.29$, $p = .28$, $\eta^2 = .04$).

Primary Analyses

Results from the repeated-measures ANOVA for SIBS scores indicated a significant Time main effect ($F = 6.33$ (1, 70), $p = .005$, $\eta^2_p = .11$, power = .82), yet a non-significant Time × Condition interaction effect ($F = 6.33$ (2, 70), $p = .94$, $\eta^2_p < .01$, power = .06). Calculation of effect sizes indicated small-positive therapeutic changes in SIBS scores across all conditions (EAU $d = .31$, GBG $d = .37$, MST $d = .38$). Findings from the repeated-measures ANOVA for SEBS scores also indicated a significant Time main effect ($F = 5.28$ (1, 70), $p = .003$, $\eta^2 = .07$, power = .62), but a non-significant Time × Condition interaction effect ($F = 2.49$ (2, 70), $p = .09$, $\eta^2_p = .07$, power = .49). However, calculation of effect sizes indicated differentiated changes for SEBS scores by condition, with the EAU and GBG groups showing negligible effects ($d = .12$ and .05, respectively) and the MST group showing a moderate positive therapeutic effect ($d = .55$). Finally, results from the repeated-measures ANOVA for SSWQ scores were non-significant for both the Time main effect ($F = 1.48$ (1, 70), $p = .23$, $\eta^2_p = .02$, power = .22) and the Time × Condition interaction effect ($F = 1.88$ (2, 70), $p = .16$, $\eta^2_p = .05$, power = .38). That said, calculation of effect sizes again indicated differentiated changes for SSWQ scores by condition, with the EAU and GBG groups showing negligible

Table 2 Descriptive statistics for outcome measures with the total sample

Outcome/time	M	SD	r_{T1-T2}
Internalizing			
Time 1	9.56	4.15	
Time 2	8.67	3.52	.57**
Externalizing			
Time 1	13.45	5.99	
Time 2	11.70	6.35	.67**
Wellbeing			
Time 1	51.67	6.90	
Time 2	50.49	8.94	.31**

Skewness and kurtosis for all means scores at both time points was < |1|

**$p < .01$

effects ($d = .07$ and $-.04$, respectively) and the MST group showing a small iatrogenic effect ($d = -.45$). Given that none of the omnibus ANOVA yielded significant interaction effects, further post-hoc tests were deemed unnecessary.

Discussion

The critical need to manage classroom behavior is especially salient in high-poverty alternative schools given the elevated rates and severity of problem behavior (Simonsen et al. 2010; Thomas et al. 2008), and their major goal to facilitate students return to traditional community schools. Unfortunately, students from these schools are often underrepresented in the intervention literature. The present study extends the literature on classroom behavior management by examining the effectiveness of two different theoretical approaches to classroom management, one based on behavioral principles (i.e., the GBG) and the other on mindfulness-based practices (i.e., MST), with a sample of fifth-grade, predominantly African American students from an urban, high-poverty alternative school. To the authors' knowledge, there are less than a handful of studies of the GBG and MBIs for youth that have been conducted in alternative schools. Additionally, there are no studies of the GBG and only one investigation of a MBI in a high-poverty alternative school. Thus, this study may shed light on the generalizability of these interventions beyond regularly studied student populations and traditional school contexts. Given the emerging state of MBIs for youth, this study also attempted to further contribute to our understanding of these interventions by singling out a common core component for testing in isolation (i.e., the mindful STOP procedure an informal mindfulness training practice).

Table 3 Descriptive statistics for outcomes measures by condition

Outcome/Time	EAU group			GBG group			MST group		
	M	SD	r_{T1-T2}	M	SD	r_{T1-T2}	M	SD	r_{T1-T2}
Internalizing									
Time 1	9.02	3.39		8.65	3.64		11.23	5.09	
Time 2	8.60	3.55	.51**	8.50	3.63	.67**	8.95	3.50	.58**
Externalizing									
Time 1	13.35	6.57		13.49	6.18		13.55	5.28	
Time 2	11.31	6.31	.51**	11.92	6.94	.80**	11.92	5.95	.71**
Wellbeing									
Time 1	50.31	6.65		51.49	7.27		53.49	6.67	
Time 2	50.98	9.53	.30^{NS}	51.18	9.94	.54**	49.13	7.09	.03^{NS}

Repeated-measures univariate ANOVA results:

Internalizing problems

Time main effect: $F = 6.33$ (1, 70), $p = .005$, $\eta^2_p = .11$, power $= .82$

Time × condition effect: $F = 6.33$ (2, 70), $p = .94$, $\eta^2_p < .01$, power $= .06$

EAU $d = .31$

GBG $d = .37$

MST $d = .38$

*Externalizing problems

Time main effect: $F = 5.28$ (1, 70), $p = .003$, $\eta^2_p = .07$, power $= .62$

Time × condition effect: $F = 2.49$ (2, 70), $p = .09$, $\eta^2_p = .07$, power $= .49$

EAU $d = .12$

GBG $d = .05$

MST $d = .55$

*Student wellbeing

Time main effect: $F = 1.48$ (1, 70), $p = .23$, $\eta^2_p = .02$, power $= .22$

Time × condition effect: $F = 1.88$ (2, 70), $p = .16$, $\eta^2_p = .05$, power $= .38$

EAU $d = .07$

GBG $d = -.04$

MST $d = -.45$

Interpretation of Results

We hypothesized that, compared to students in the EAU condition, those in both the GBG and MST conditions would evidence positive therapeutic changes in behavior across all outcomes of interest: internalizing problems, externalizing problems, and wellbeing. Considering that this study was the first, at least to our knowledge, to directly investigate the effects of behavioral-based and mindfulness-based classroom interventions, we had no expectation regarding potential differential effects across outcomes—only that the intervention conditions would produce more desirable behavior change than education as usual. That said, results from the present study generally failed to support our hypotheses. Specifically, although Time main effects were observed for both SIBS and SEBS scores, the Time × Condition interaction effects were deemed non-significant for all outcomes, suggesting that any differences in changes from pretest to posttest were unlikely attributable to the classroom interventions.

It is noteworthy, however, that all of these findings, which were based on inferential statistics, are characterized by extremely low statistical power, suggesting low likelihoods of identifying meaningful effects, if actually present. Given the context of the present study, we suspect that such low power was strongly influenced by the small sample size. We therefore calculated and considered effect sizes to represent the magnitude of observed change from pretest to posttest for each condition across each outcome. These findings provide a richer description of the results and suggest a more nuanced view of potential intervention effects. Specifically, for internalizing problems, effect size estimates indicated small-positive changes from pretest to posttest for all conditions, suggesting no unique intervention effects for either the GBG or MST. Yet for externalizing problems, MST was the only condition to yield a meaningful pre–post change, suggesting a moderate therapeutic effect in favor of the mindfulness-based classroom intervention. Similarly, for the student wellbeing

outcome, MST was again the only condition to yield a meaningful pre–post change, yet this effect was in the undesirable direction, suggesting the mindfulness-based intervention had an iatrogenic effect. Taken together, these findings provide preliminary evidence that MST might have differential effects on student mental health outcomes, compared with education as usual and more traditional behavioral interventions. However, this evidence is not universally encouraging (given the observed iatrogenic effect on wellbeing), and it should not be interpreted or accepted with high confidence (given the non-significant inferential statistics observed for the interaction effects). Ultimately, then, we conclude that data from the present study suggest that neither classroom management strategy was clearly more effective than education as usual, and that further research is warranted—with much larger sample sizes—to explore the potential differential effects of the GBG and MST on student mental health outcomes.

Implications for Practice

Why did these two research-based interventions fail to produce the anticipated effects? One hypothesis is that the interventions were ineffective for this unique population of students and school intervention context due to limitations in generalizability. Importantly, limitations in generalizability can be revealed by unexpected intervention and implementation outcomes as issues with fit can compromise the effectiveness of treatments as well as the ease in which they are implemented. This may have been the case in the present study and highlights the importance of deliberately considering the generalizability and fit of evidence-based interventions a priori.

Frequently within schools, once interventions have been labeled as evidence-based school practitioners readily apply them with little forethought about their transportability to the new intervention context. They naturally assume that demonstrated treatment effects will transfer. Yet, schools often are implementing evidence-based interventions in contexts and with student populations that differ from those under which they were designed and tested. Therefore, when selecting interventions for use in schools, it seems critical for practitioners to consider both the body of evidence on an intervention (i.e., the quantity and quality), as well as the relevance of this evidence for one's specific intervention context. As an example, Spilt et al. (2013) made an astute observation that while many studies of selected and indicated interventions have examined potential moderators of treatment effects, this research is less commonly done with universal preventive interventions. To address this gap in the literature, they conducted an examination of subgroup differences in the effectiveness of the GBG. This research revealed that although the GBG was successful in reducing externalizing and internalizing behaviors among low-risk children, children with emotional problems, and children victimized by their peers, no treatment effects were found for children from troubled homes (i.e., children whose mothers were distressed, had low parent involvement, and were often low SES) or with a combination of individual risk factors (i.e., children who exhibited multiple elevations across problematic sociobehavioral domains of functioning). Characteristics of these latter subgroups closely match those of the present study sample and expose potential non-responders to the GBG as traditionally designed. In relation, a review of the research literature on the GBG and MBIs for youth revealed that despite the large body of evidence supporting these interventions less than a handful of studies had sample and setting characteristics comparable to this one.

Given that no intervention can be tested with all student populations within all school contexts, practitioners need to become skilled in their ability to transport evidence-based interventions to new student populations and intervention contexts. To do so, school support specialists need to understand the theory behind selected interventions, including the specific mechanisms that drive changes in target outcomes. With this understanding, modifications can be made to intervention content or delivery to optimize fit, thus enhancing the chances of attaining desired outcomes. For example, the effectiveness of the GBG relies heavily on behavioral principles of reinforcement. Accordingly, students' responses will compete with other behavior contingencies in the classroom environment. Therefore, any intervention must take these other contingencies into consideration in its design, as students will allocate their responding to denser schedules of reinforcement and more potent reinforcers. In the present study, students were provided with a tangible reinforcer at the end of the period only if the target behaviors across the entire period fell below a predetermined goal for their team. However, it is highly probable that students in this study were more reinforced by teacher and/or peer attention, as these reinforcers would be easier to access and more frequent than the tangible one. Thus, the efficiency of the reinforcers used for the GBG in this study may have been diminished. This hypothesis is supported by meta-analytic research that suggests the correct application of reward procedures during implementation of the GBG is associated with differential treatment effects (Flower et al. 2014a). In relation, students placed in alternative schools often have a history of negative school experiences related to exclusionary and punitive discipline practices (Jolivette et al. 2014). Thus, it may have been advantageous to apply the alternate version of the GBG which awards points for engagement in appropriate behaviors, rather than demarcations for rule violations. In so doing, the game reward procedures may have helped to counter this negative history and been more efficient.

Potential explanations for the mixture of non-significant treatment effects for the MST intervention may also be related to suboptimal fit. For instance, an examination of the literature on MBIs suggests that Ms. Cooper's competency in

mindfulness procedures and the intensity of the MST intervention may have been insufficient. Throughout the literature on MBIs for youth, intervention agents had extensive personal experiences in mindfulness practices as a prerequisite to carrying out a specific intervention (e.g., Burke 2010; Harnett and Dawe 2012; Meiklejohn et al. 2012). Some recognized good practice guidelines for mindfulness intervention agents include "an established and ongoing personal mindfulness meditation practice, professional training, regular supervision, attendance at teacher-led silent meditation retreats, and ongoing professional development" (Burke 2010, p. 134). Therefore, a customary level of background knowledge and skills in mindfulness may have been required to be a fully effective intervention agent. The approximately 1-hour training afforded Ms. Cooper on MST may have been adequate for promoting adherence to the intervention, but inadequate to result in highly competent implementation, as skilled implementation appears to necessitate expertise in mindfulness practices more broadly.

This study attempted to contribute to the understanding of MBIs for youth by singling out a common core component (i.e., MST) for testing in isolation. Although this was advantageous for gaining a clear picture of the possible effectiveness and importance of this component, it may have resulted in a mismatch between intervention intensity and student need. Most MBIs studied have been supplied in larger doses. In fact, many of the interventions previously studied were comprised of several weeks of experiential learning opportunities including some combination of guided mindfulness practices (e.g., body scan, meditation), interventionist-led enquiry, discussion, and psychoeducation. Additionally, those previously studied often included regular outside-of-session practice (Burke 2010; de Vibe et al. 2012). While there were MBIs found whose design and dosage paralleled that of the MST intervention in this study, these tended to be implemented with non-clinical populations who despite sharing some demographic characteristics with the present sample had less significant needs (e.g., Black and Fernando 2014).

Limitations and Directions for Future Research

As with all research studies, this study has some limitations that readers should be aware of. First, one of the primary limitations of the present study was the small sample size. The small nature of the sample was largely due to the unique student population and setting of focus, as well as the practice-based, collaborative action research design. Like other research designs, collaborative action research designs are subject to the same threats to internal validity. However, due to their collaborative nature, they can be more susceptible to data collector bias, because data collectors are aware of the study purpose and sometimes operate as intervention agents, as in the case of the present study where Ms. Cooper delivered both

interventions under investigation. To help mitigate this potential confound, great care was taken to (a) not provide Ms. Cooper with any information to suggest the merits of one intervention over the other, and (b) impress upon her the importance of not overlooking implementation issues or student responses she did not want to see. Thus, the weekly check-in meetings were essential to supporting the rigor of this study. As stated earlier, the small sample size was also related to extremely low power estimates. Therefore, the moderate positive and negative treatment effects observed for the MST intervention could not be ruled out as being due to chance. Furthermore, the size of these effects may, in part, be impacted by the study being underpowered. Small sample sizes are common in intervention research carried out in alternative school settings. Despite this and other challenges of conducting research in these schools, continued intervention research with specialized student populations in these settings is greatly needed. This research not only helps educators to meet the needs of more students, but also reveals important limitations about the generalizability of evidence-based interventions.

A second limitation of this study was the potential mismatch in intensity between the selected classroom management interventions and the students' demonstrated needs. It is possible that the short duration of the interventions (i.e., 4 weeks), coupled with the biweekly, implementation undermined the ability of these interventions to produce meaningful and significant changes in outcomes. Although both classroom management approaches were grounded in sound theory and had evidence to support their implementation as designed, when conducting effectiveness research, it is critical to consider intervention fit and possible moderators to treatment effectiveness. Only by doing so, can appropriate intervention adjustments be made to maximize the chances for intervention success.

A third study limitation was the inability to gather direct observation data on intervention fidelity and the primary student behaviors targeted. This restriction was placed on the study by the IRB due to the unique vulnerability of this student population and, relatedly, inability to secure active parental consent for all students. This barrier to gathering direct observation data likely hindered our ability to detect treatment effects, because direct observations of student behavior would have been more targeted and sensitive to change than data gathered via the teacher-report measures (the SIBS and SEBS) used in this study. Though, one advantage of this study having breadth in measurement of treatment outcomes was the ability to detect the full scope of possible treatment effects. As a result, this study detected a potential unintended negative outcome, namely iatrogenic effects on student wellbeing as a result of implementing the MST intervention with this student population.

Ms. Cooper rated her own fidelity via an adherence check-list, and while this is better than no fidelity monitoring and common practice in schools, it does not provide indisputable evidence of adequate adherence to the interventions. Intervention fidelity is typically a prerequisite for producing desired changes in target outcomes. Some research suggests that self-report methods of intervention fidelity are problematic because of the increased likelihood for upwardly biased reporting and the low agreement they often have with direct observation methods (see Noell 2008). However, in consultation, direct observation methods can be influenced by observer reactivity (Sheridan et al. 2009) and it is not always practical for researchers to observe a representative proportion of intervention sessions. Additionally, new research has shown the potential reliability and accuracy of self-report methods of intervention fidelity under specialized circumstances. For instance, Gresham et al. (2017) applied generalizability theory to assess the dependability of teacher self-report of intervention fidelity to the GBG and found teacher self-report data produced a reliable estimate of intervention fidelity after roughly seven measurements. The present study had an analogous self-report intervention fidelity measure and captured greater than seven data points for each intervention. In relation, Sanetti and Kratochwill (2011) have suggested that teachers can be accurate and reliable reporters in cases when they have been trained on proper procedures for assessing their intervention fidelity. More specifically, research on self-report methods indicates that "individuals are more accurate reporters as the time frame for self-reporting decreases and the specificity of the required responses increases" (p. 300). This research was explicitly drawn upon when designing intervention fidelity procedures for the present study.

Taken together, a case can be made for the validity of the intervention fidelity data in this study. Anecdotal data gathered from Ms. Cooper provides further support for this validity and guidance for intervention fidelity measurement in future research. Information gleaned from Ms. Cooper during weekly check-ins suggested she was applying intervention fidelity assessment procedures per her training. Although, conversations with her surrounding implementation issues suggested adherence to the interventions was less problematic than her quality or skill of delivery. More specifically, Ms. Cooper described unanticipated interactions between the interventions and her students, which she did not always feel confident she effectively addressed. With respect to the MST, Ms. Cooper reported application of the mindful STOP procedure sometimes negatively impacted the flow of the class lesson. She expressed a lack of familiarity with concepts and exercises used to build mindfulness skills, difficulty gaining buy-in from students, and problems with soliciting active engagement with the intervention. Ms. Cooper stated that some students felt "uncomfortable and unnatural" engaging in the mindfulness training, especially around peers who exhibited

bullying behavior. In contrast, she indicated she was acquainted and comfortable with the strategies comprising the GBG. This may have resulted in the more optimal levels of fidelity reported for this intervention.

Despite her greater comfort with the GBG, she again described some problematic interactions between the intervention and her students. Specifically, students engaged in countercontrol behaviors that resulted in fluctuations of the effectiveness of the game from day to day. Countercontrol relates to intervention recipients' reaction to what they perceive as attempts by the intervention agent to control their behavior. Individuals who engage in countercontrol "may show an emotional reaction of anger or frustration including operant behavior which injures or is otherwise aversive to the controller" (Skinner 1953, p. 321). As an example, students in the present study expressed betrayal or a sense of being "tricked" when a team they were members of did not earn a desirable reinforcer, despite their efforts to follow the class rules. In response, students often refused to follow the class rules the next school day.

The above implementation challenges reveal the importance for intervention researchers to capture more than one dimension of intervention fidelity. Power et al. (2005) recommended that two categories of intervention fidelity be considered when planning assessment of the construct: content (i.e., How much of the intervention is delivered?) and process (i.e., How well is the intervention delivered?). Content dimensions include customary measurement of adherence to intervention components and tracking of intervention dosage. Process dimensions are less frequently studied and include quality/skill of delivery and participant responsiveness to the intervention. Data of intervention process dimensions are essential for gleaning information related to intervention fit and knowledge for guiding recommendations to enhance intervention success.

Conclusion

The present study found some evidence to suggest that mindfulness-based classroom management interventions may have differential effects compared to both education as usual and a traditional behaviorally-oriented intervention. Yet this evidence was not universally promising, and the obtained inferential statistics do not allow for strong confidence in the results. We therefore conclude that neither intervention approach was found to be ideal for use with this unique population of students, and we encourage follow-up research with more robust sampling methods to further investigate these phenomena. That said, the contribution of these study findings to the broader literature is diminished by the inability to rule them out as being due to chance. Despite this limitation, we suggest that results do illustrate the benefit of measuring a diverse set of outcomes when examining the effectiveness of school-based

interventions, as doing so may supply a more complete and accurate picture of the benefits and possible detriments of various classroom-based intervention approaches.

In conclusion, further research is needed to determine how to effectively transport classroom-management interventions to high risk and diverse student populations, such as that reflected by the current sample. If school psychologists are to be effective as consultants, they need to know not only how to select evidence-based interventions but also how to integrate the best research evidence and theory with knowledge about the unique needs and characteristics of the target population and context. It is only by doing so that school psychologists can optimize the chance of attaining desired intervention outcomes for all students.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Research Involving Human Participants and/or Animals All procedures performed in this study involving human participants were in accordance with the ethical standards of the authors' institutional and national research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent For this type of study, no formal consent was required as subjects could not be identified in the research data indirectly, directly or statistically (IRB# E9115).

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Akin-Little, A., Little, S. G., Bray, M. A., & Kehle, T. J. (Eds.). (2009). *Behavioral interventions in schools: evidence-based positive strategies*. Washington: American Psychological Association.

Arslan, G., & Renshaw, T. L. (2017). Student subjective wellbeing as a predictor of adolescents' problem behaviors: A comparison of first-order and second-order factor effects. Child Indicators Research. Advance online publication. doi:https://doi.org/10.1007/s12187-017-9444-0.

Aud, S., Wilkinson-Flicker, S., Kristapovich, P., Rathbun, A., Wang, X., & Zhang, J. (2013). *The condition of education 2013 (NCES 2013-037)*. U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch.

Baer, R. A. (2003). Mindfulness training as a clinical intervention: a conceptual and empirical review. *Clinical Psychology: Science and Practice, 10*, 125–143. https://doi.org/10.1093/clipsy/bpg015.

Barrish, H. H., Saunders, M. S., & 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.

Beebe-Frankenberger, M., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Sorting second-grade students: differentiating those retained from those promoted. *Journal of Educational Psychology, 96*(2), 204–215. https://doi.org/10.1037/0022-0663.96.2.204.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: a proposed operational definition. *Clinical Psychology: Science and Practice, 11*, 230–241. https://doi.org/10.1093/clipsy.bph077.

Black, D. S., & Fernando, R. (2014). Mindfulness training and classroom behavior among lower-income and ethnic minority elementary school children. *Journal of Child and Family Studies, 23*, 1242–1246. https://doi.org/10.1007/s10826-013-9784-4.

Bluth, K., Campo, R. A., Pruteanu-Malinici, S., Reams, A., Mullarkey, M., & Broderick, P. C. (2016). A school-based mindfulness pilot study for ethnically diverse at-risk adolescents. *Mindfulness, 7*, 90–104. https://doi.org/10.1007/s12671-014-0376-1.

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting positive behavior using the good behavior game: A meta-analysis of single case research. *Journal of Positive Behavior Interventions, 18*, 180–190. https://doi.org/10.1177/1098300715592355.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., … Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: a six-site, cross-national study. Developmental Psychology, 39, 222–245. doi:https://doi.org/10.1037/0012-1649.39.2.222.

Burke, C. A. (2010). Mindfulness-based approaches with children and adolescents: a preliminary review of current research in an emergent field. *Journal of Child and Family Studies, 19*, 133–144. https://doi.org/10.1007/s10826-009-9282-x.

Carver, P. R., & Lewis, L. (2010). *Alternative schools and programs for public school students at risk of educational failure: 2007–08 (NCES 2010–026). U.S. Department of Education, National Center for Education Statistics*. Washington: Government Printing Office.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98*, 388–400.

Chaffee, R. K., Briesch, A. M., Johnson, A. H., & Volpe, R. J. (2017). A meta-analysis of class-wide interventions for supporting student behavior. *School Psychology Review, 46*, 149–164. https://doi.org/10.17105/SPR-2017-0015.V46-2.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum.

Cook, C. R., Rasetshwane, K. B., Truelson, E., Grant, S., Dart, E. H., Collins, T. A., et al. (2011). Development and validation of the student internalizing behavior screener: examination of reliability, validity, and classification accuracy. *Assessment for Effective Intervention, 36*, 71–79. https://doi.org/10.1177/1534508410390486.

de Vibe, M., Bjørndal, A., Tipton, E., Hammerstrøm, K. T., & Kowalski, K. (2012). Mindfulness based stress reduction (MBSR) for improving health, quality of life and social functioning in adults. *Campbell Systematic Reviews, 2012*, 3. https://doi.org/10.4073/csr.2012.3.

Embry, D. D. (2002). The good behavior game: a best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*, 273–297. https://doi.org/10.1023/A:1020977107086.

Emmer, E. T., & Sabornie, E. J. (2015). Introduction to the second edition. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of classroom management* (2nd ed., pp. 3–12). New York: Routledge.

Felver, J. C., Celis-de Hoyos, C. E., Tezanos, K., & Singh, N. N. (2016). A systematic review of mindfulness-based interventions for youth in school settings. *Mindfulness, 7*, 34–45. https://doi.org/10.1007/s12671-015-0389-4.

Fishbein, D., Miller, S., Herman-Stahl, M., Williams, J., Lavery, B., Markovitz, L., et al. (2016). Behavioral and psychophysiological effects of a yoga intervention on high-risk adolescents: A randomized control trial. *Journal of Child and Family Studies, 25*, 518–529. https://doi.org/10.1007/s10826-015-0231-6.

Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega, R. J. (2014a). Effects of the good behavior game on challenging

behaviors in school settings. *Review of Educational Research, 84*, 546–571. https://doi.org/10.3102/0034654314536781.

Flower, A., McKenna, J., Muething, C. S., Bryant, D. P., & Bryant, B. R. (2014b). Effects of the good behavior game on classwide off-task behavior in a high school basic algebra resource classroom. *Behavior Modification, 38*, 45–68. https://doi.org/10.1177/0145445513507574.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). New York: McGraw-Hill.

Good, T. L., & Brophy, J. E. (2008). *Looking in classrooms* (10th ed.). Boston: Pearson Education, Inc..

Gresham, F. M., Dart, E. H., & Collins, T. A. (2017). Generalizability of multiple measures of treatment integrity: comparisons among direct observation, permanent products, and self-report. *School Psychology Review, 46*, 108–121. https://doi.org/10.17105/SPR46-1.108-121.

Harnett, P. H., & Dawe, S. (2012). The contribution of mindfulness-based therapies for children and families and proposed conceptual integration. *Child and Adolescent Mental Health, 17*, 195–208. https://doi.org/10.1111/j.1475-3588.2011.00643.x.

Hastings, R. P., & Bham, M. S. (2003). The relationship between student behavior patterns and teacher burnout. *School Psychology International, 24*, 115–127.

Jolivette, K., Patterson, D. P., Swoszowski, N. C., McDaniel, S. C., Kennedy, C., & Ennis, R. P. (2014). School-wide positive behavioral interventions and supports in a residential school for students with emotional and behavioral disorders: first years of implementation and maintenance follow-up focus groups. *Residential Treatment for Children & Youth, 31*, 63–79. https://doi.org/10.1080/0886571X.2014.878584.

Joslyn, P. R., Vollmer, T. R., & Hernández, V. (2014). Implementation of the good behavior game in classrooms for children with delinquent behavior. *Acta de Investigacio'n Psicolo'gica, 4*, 1673–1681. https://doi.org/10.1016/S2007-4719(14)70973-1.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future. *Clinical Psychology: Science and Practice, 10*, 144–156. https://doi.org/10.1093/clipsy/bpg016.

Kallapiran, K., Koo, S., Kirubakaran, R., & Hancock, K. (2015). Effectiveness of mindfulness in improving mental health symptoms of children and adolescents: a meta-analysis. *Child and Adolescent Mental Health, 20*, 182–194. https://doi.org/10.1111/camh.12113.

Kellam, S. G., Brown, C. H., Poduska, J. M., Lalongo, N. S., Wang, W., Toyinbo, P., … 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, 95S*, S5–S28. doi:https://doi.org/10.1016/j.drugalcdep.2008.01.004.

Kleiner, B., Porch, R., & Farris, E. (2002). *Public alternative schools and programs for students at risk of educational failure: 2000–01 (NCES 2002–004)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Klingbeil, D. A., Fischer, A. J., Renshaw, T. L., Bloomfield, B. S., Polakoff, B., Willenbrink, J. B., & ... Chan, K. T. (2017a). Effects of mindfulness-based interventions on disruptive behavior: a meta-analysis of single-case research. *Psychology in the Schools, 54*, 70–87. doi:https://doi.org/10.1002/pits.21982.

Klingbeil, D. A., Renshaw, T. L., Willenbrink, J. B., Copek, R. A., Chan, K. T., Haddock, A., … Yassine, J. (2017b). Mindfulness-based interventions with youth: A comprehensive meta-analysis of group-design studies. *Journal of School Psychology, 63*, 77–103. doi:https://doi.org/10.1016/j.jsp.2017.03.006.

Lehr, C. A., Tan, C. S., & Ysseldyke, J. (2009). Alternative schools: a synthesis of state-level policy and research. *Remedial and Special Education, 30*, 19–32.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-

analysis. *American Psychologist, 48*, 1181–1209. https://doi.org/10.1037/0003-066X.48.12.1181.

McFarland, J., Hussar, B., de Brey, C., Snyder, T., Wang, X., Wilkinson-Flicker, S., … Hinz, S. (2017). *The condition of education 2017 (NCES 2017–144)*. U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017144.

Medland, M. B., & Stachnik, T. J. (1972). Good behavior game: a replication and systematic analysis. *Journal of Applied Behavior Analysis, 5*, 45–51.

Meiklejohn, J., Phillips, C., Freedman, M. L., Griffin, M. L., Biegel, G., Roach, A., … & Isberg, R. (2012). Integrating mindfulness training into K-12 education: fostering the resilience of teachers and students. *Mindfulness, 3*, 291–307. doi:https://doi.org/10.1007/s12671-012-0094-5.

Mitchell, M. M., & Brashaw, C. P. (2013). Examining classroom influences on student perceptions of school climate: the role of classroom management and exclusionary discipline strategies. *Journal of School Psychology, 51*, 599–610. https://doi.org/10.1016/j.jsp.2013.05.005.

Mouzakitis, A., Codding, R. S., & Tryon, G. (2015). The effects of self-monitoring and performance feedback on the treatment integrity of behavior intervention plan implementation and generalization. *Journal of Positive Behavior Interventions, 17*, 223–234. https://doi.org/10.1177/1098300715573629.

National Center for Education Statistics. (2006). *Urban education in America: School locale definitions*. Retrieved from https://nces.ed.gov/surveys/urbaned/definitions.asp.

Noell, G. H. (2008). Research examining the relationships among consultation process, treatment integrity, and outcomes. In W. P. Erchul & S. M. Sheridan (Eds.), *Handbook of research in school consultation: empirical foundations for the field* (pp. 315–334). Mahwah: Erlbaum.

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews, 2011*, 4. https://doi.org/10.4073/csr.2011.4.

Power, T. J., Blom-Hoffman, J., Clarke, A. T., Riley-Tillman, T. C., Kelleher, C., & Manz, P. H. (2005). Reconceptualizing intervention integrity: a partnership-based framework for linking research with practice. *Psychology in the Schools, 42*, 495–507. https://doi.org/10.1002/pits.20087.

Reinke, W. M., Herman, K. C., & Sprick, R. (2011). *Motivational interviewing for effective classroom management: the classroom check-up*. New York: Guilford Press.

Renshaw, T. L. (2012). Mindfulness-based practices for crisis prevention and intervention. In S. E. Brock & S. R. Jimerson (Eds.), *Handbook of school crisis prevention and intervention* (2nd ed., pp. 401–422). Bethesda: National Association of School Psychologists.

Renshaw, T. L. (2015). A replication of the technical adequacy of the student subjective wellbeing questionnaire. Journal of Psychoeducational Assessment. Advanced online publication. doi:https://doi.org/10.1177/0734282915580885.

Renshaw, T. L., & Arslan, G. (2016). Psychometric properties of the student subjective wellbeing questionnaire with Turkish adolescents: a generalizability study. *Canadian Journal of School Psychology, 31*, 139–151. https://doi.org/10.1177/0829573516634644.

Renshaw, T. L., & Chenier, J. S. (2016). Further validation of the student subjective wellbeing questionnaire: Comparing first-order and second-order factor effects on actual school outcomes. Journal of Psychoeducational Assessment. Advance online publication. doi:https://doi.org/10.1177/0734282916678494.

Renshaw, T. L., & O'Malley, M. D. (2014). Cultivating mindfulness in students. In M. J. Furlong, R. Gilman, & E. S. Huebner (Eds.), *Handbook of positive psychology in the schools* (2nd ed., pp. 245–259). New York: Routledge.

Renshaw, T. L., Bolognino, S. J., Fletcher, S. P., & Long, A. C. J. (2015a). Using mindfulness to improve wellbeing in schools. *NASP Communiqué, 43*(6), 4–8.

Renshaw, T. L., Long, A. C. J., & Cook, C. R. (2015b). Assessing adolescents' positive psychological functioning at school: development and validation of the student subjective wellbeing questionnaire. *School Psychology Quarterly, 30*, 534–552. https://doi.org/10.1037/spq0000088.

Richards, L., Heathfield, L., & Jenson, W. (2010). A classwide peer-modeling intervention package to increase on-task behavior. *Psychology in the Schools, 47*, 551–566. https://doi.org/10.1002/pits.20490.

Salend, S. J., Reynolds, C. J., & Coyle, E. M. (1989). Individualizing the good behavior game across type and frequency of behavior with emotionally disturbed adolescents. *Behavior Modification, 13*, 108–126. https://doi.org/10.1177/01454455890131007.

Sanetti, L. M. H., & Kratochwill, T. R. (2011). An evaluation of the treatment integrity planning protocol and two schedules of treatment integrity self-report: impact on implementation and report accuracy. *Journal of Educational & Psychological Consultation, 21*, 284–308. https://doi.org/10.1080/10474412.2011.620927.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57*(1), 1–10. https://doi.org/10.1037/a0018082.

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. *Journal of Clinical Psychology, 62*, 373–386. https://doi.org/10.1002/jclp.20237.

Sheridan, S. M., Swanger-Gagne, M., Welch, G. W., Kwon, K., & Garbacz, S. A. (2009). Fidelity measurement in consultation: Psychometric issues and preliminary examination. *School Psychology Review, 38*, 476–495.

Simonsen, B., Fairbank, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: considerations for research to practice. *Education and Treatment of Children, 31*, 351–380.

Simonsen, B., Myers, D., & DeLuca, C. (2010). Teaching teachers to use prompts, opportunities to respond, and specific praise. *Teacher Education and Special Education: The Journal of Teacher Education Division of the Council for Exceptional Children, 33*, 300–318. https://doi.org/10.1177/0888406409359905.

Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions, 15*, 5–15. https://doi.org/10.1177/1098300712440453.

Simonsen, B., Freeman, J., Goodman S., Mitchell, B., Swain-Bradway, J., Flannery, B., … Putman, B. (2015). Supporting and responding to behavior: evidence-based classroom strategies for teachers. OSEP Technical Assistance Brief. Retrieved from https://www.osepideasthatwork.org/evidencebasedclassroomstrategies.

Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.

Skinner, C. H., Skinner, A. L., & Burton, B. (2009). Applying group-oriented contingencies in the classroom. In A. Akin-Little, S. G. Little, M. A. Bray, & T. J. Kehle (Eds.), Behavioral interventions in schools: evidence-based positive strategies (pp. 157–170). Washington, DC: American Psychological Association.

Snider, L. A., Seligman, L. D., Ketchen, B. R., Levitt, S. J., Bates, L. R., Garvey, M. A., et al. (2002). Tics and problem behaviors in schoolchildren: prevalence, characterization, and associations. *Pediatrics, 110*, 331–336. https://doi.org/10.1542/peds.110.2.331.

Spilt, J. L., Koot, J. M., & van Lier, P. C. (2013). For whom does it work? Subgroup differences in the effects of a school-based universal prevention program. *Prevention Science, 14*, 479–488. https://doi.org/10.1007/s11121-012-0329-7.

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

Sy, J. R., Gratz, O., & Donaldson, J. M. (2016). The good behavior game with students in alternative educational environments: interactions between reinforcement criteria and scoring accuracy. *Journal of Behavioral Education, 25*, 455–477. https://doi.org/10.1007/s10864-016-9257-0.

Thomas, D. E., Bierman, K. L., Thompson, C., & Powers, C. J. (2008). Double jeopardy: child and school characteristics that predict aggressive-disruptive behavior in first grade. *School Psychology Review, 37*, 516–532.

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The good behavior game: 1969-2002. *Behavior Modification, 30*, 225–253. https://doi.org/10.1177/0145445503261165.

U.S. Department of Education, National Center for Education Statistics. (2005). *The condition of education 2005* (NCES 2005–094). Washington, DC: U.S. Government Printing Office.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. (1987). Effectiveness of psychotherapy with children and adolescents: a meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55*, 542–549.

What Works Clearinghouse. (2017). *What works clearinghouse: Standards handbook, version 4.0*. Washington, DC: Institute of Educational Sciences Retrieved from https://ies.ed.gov/ncee/wwc/Handbooks.

Wisner, B. L., & Norton, C. L. (2013). Capitalizing on behavioral and emotional strengths of alternative high school students through group counseling to promote mindfulness skills. *Journal for Specialists in Group Work, 38*, 207–224. https://doi.org/10.1080/01933922.2013.803504.

Ysseldyke, J., Betts, J., Thill, T., & Hannigan, E. (2004). Use of an instructional management system to improve mathematics skills for students in title I programs. *Preventing School Failure, 48*, 10–14. https://doi.org/10.3200/PSFL.48.4.10-14.

Zenner, C., Herrnleben-Kurz, S., & Walach, H. (2014). Mindfulness-based interventions in schools: a systematic review and meta-analysis. *Frontiers in Psychology, 5*, 603. https://doi.org/10.3389/fpsyg.2014.00603.

Zoogman, S., Goldberg, S. B., Hoyt, W. T., & Miller, L. (2015). Mindfulness interventions with youth: a meta-analysis. *Mindfulness, 6*, 290–302. https://doi.org/10.1007/s12671-013-0260-4.

Anna C. J. Long, PhD, is an Assistant Professor of School Psychology and Chair of the Committee on Diversity and Outreach in Psychology within the Department of Psychology at Louisiana State University. Her research interests center around the promotion of treatment integrity to evidence-based interventions in schools and community-based settings, and culturally-responsive, evidence-based practice.

Tyler L. Renshaw, PhD, is an Assistant Professor in the School Psychology program within the Department of Psychology at Utah State University. His research interests are broadly focused on developing and validating assessment and intervention methods for promoting well being and mental health in schools.

Devon L. Camarota is now a middle school Special Education teacher at Democracy Prep Baton Rouge, and will graduate with a Master's in Public Administration from Louisiana State University in May 2018. Devon has been teaching for five years in North Baton Rouge and has a passion for promoting educational equity.