School-based Supports and Interventions to Improve Social and Behavioral Outcomes with

Racially and Ethnically Minoritized Youth: A Review of Recent Quantitative Research

Lindsay M. Fallon, Ph.D., BCBA-D[1], Emily R. DeFouw, Ph.D., BCBA[2],

Sadie C. Cathcart, M.A.[1], Talia S. Berkman, B.A.[1], Patrick Robinson-Link, M.A.[1],

Breda V. O'Keeffe, Ph.D.[3], and George Sugai, Ph.D.[4]

[1]University of Massachusetts Boston, [2]Ohio University,

[3]University of Utah, [4]University of Connecticut

Citation:

Author Note

Abstract

School discipline disproportionality has long been documented in educational research, primarily impacting Black/African American and non-White Hispanic/Latinx students. In response, federal policymakers have encouraged educators to change their disciplinary practice, emphasizing that more proactive support is critical to promoting students' social and behavioral outcomes in school. Results from a literature review conducted nearly a decade ago indicated that there was, at that point, a paucity of empirical research related to considering students' culture (e.g., race, ethnicity) and supporting school behavior. The purpose of this study is to replicate and expand the previous review to summarize the characteristics of the most recent school-based quantitative research addressing interventions to promote social and behavioral outcomes for racially and ethnically minoritized youth. We screened 1,687 articles for inclusion in the review. Upon coding 32 eligible research studies, we found that intervention and implementer characteristics within these studies varied, but noted strong intervention effects in studies that included established evidence-based practices, adapted interventions, as well as new practices piloted with student participants. Results inform recommendations to continue to study interventions that promote positive social and behavioral outcomes for racially and ethnically minoritized students to disrupt a long history of subjection to exclusionary discipline disproportionately.

Keywords: behavior, diversity, culture, systematic review

School-based Supports and Interventions to Improve Social and Behavioral Outcomes with

Racially and Ethnically Minoritized Youth: A Review of Recent Quantitative Research

In a unanimous decision, the United States Supreme Court ruled in *Brown v. Board of Education of Topeka* that school segregation violates the Equal Protection Clause of the Fourteenth Amendment. This ruling prompted school integration, propelled the Civil Rights Movement, and led to policy changes including Title VI of the Civil Rights Act of 1964. Title VI prohibits discrimination on the basis of race and includes school discipline policies, procedures, and practices (U.S. Department of Education, 2015). Yet, despite these protections, well-documented racial and ethnic disparities in both school discipline and achievement persist decades after the 1960s (Welsh & Little, 2018). Specifically, there continues to be vast evidence of disproportionality by race and ethnicity related to (a) office discipline referrals for major (Bradshaw et al., 2010; Skiba et al., 2011), minor (Marchbanks et al., 2018) and subjective student behavior (Girvan et al., 2017), (b) out-of-school suspension (Brenda, 2000; Davis & Jordan, 1994; Skiba et al., 2002; Skiba et al., 2014), (c) expulsion (Brenda, 2000; Skiba et al., 2014), (d) school-based arrests (Mowen & Brent, 2016), as well as academic achievement (Davis & Jordan, 1994; Morris & Perry, 2016; Rausch & Skiba, 2006). Black/African American and non-White Hispanic/Latinx students are most often impacted (e.g., Gage, Katsiyannis, et al., 2020; Gage, Whitford et al., 2019).

Federal response efforts have taken many forms; the Office of Civil Rights has monitored school disproportionality rates since the late 1960s, and the No Child Left Behind Act of 2001 (NCLB) was passed to hold schools accountable by requiring student scores in reading and mathematics be in the proficient range across important subgroups including race/ethnicity in order to access federal funding (Dawoody, 2008). Recently, the Every Student Succeeds Act

(ESSA) of 2015 was passed which empowers state-level education officials to develop goals, generate systems, and enact practices to promote achievement for all students. Similarly, the Individuals with Disabilities Education Act (IDEA, 2004) requires state and local education agencies receiving federal funds to collect and analyze data to determine if students from certain ethnic groups are being disproportionately disciplined and excluded from the learning environment (e.g., suspended, expelled). IDEA (2004) also specifies that if significant disproportionality is determined, school practices and procedures must be reviewed and revised to focus on prevention and reverse noted trends. Yet, the law provides little guidance about how to do this.

Existing Research

Disparities in disciplinary practices are concerning due to the relationship between exclusionary school discipline and an increased risk of academic failure, criminalization, and incarceration in adulthood (Wolf & Kupchik, 2017). Gregory and colleagues (2017) proposed a framework for eliminating these disparities that integrates both prevention and intervention efforts. The theory emphasizes the importance of student and family voice in prevention (e.g., creation of policies, procedures, practices) and intervention (e.g., coordinating re-entry) after a disciplinary incident occurs. Other components of the framework include (a) building supportive relationships, (b) promoting respectful learning environments, (c) engaging in culturally responsive teaching and behavioral instruction, (d) reviewing data for evidence of discrepant treatment of certain groups, (e) using a problem-solving approach to discipline, and (f) implementing a multi-tiered system of support framework for the provision of behavior support. Although many studies have included analysis of discipline disproportionality in schools and theorized effective solutions, few published, quantitative intervention studies have

evaluated suggestions made by Gregory and colleagues (2017). Specifically, it is unclear what intervention and practices have been studies that promote racially and ethnically minoritized students' social and behavioral outcomes proactively.

Authors (2012a) published results of a literature review related to culturally and contextually relevant behavior management practices published between 1991 and 2010 in peer-reviewed journals. They defined student culture as "the extent to which a group of individuals engage in overt and verbal behavior reflecting shared behavioral learning histories, serving to differentiate the group from other groups, and predicting how individuals within the group act in specific setting conditions" (Authors, 2012b; p. 9). Behavior support was characterized as action taken to decrease the likelihood of problem behavior and promote positive educational outcomes (Authors, 2012a). Authors (2012a) noted that the vast majority of research found and reviewed was qualitative, and summarized suggestions based on practices described in these studies. Just seven quantitative studies were found. Of these, four were record reviews, two evaluated the impact of Positive Behavioral Interventions and Supports (PBIS; Sugai & Horner, 2002) broadly on disciplinary outcomes in diverse schools, and one was an experimental single case design study with four African American males receiving individualized behavior plans.

The authors noted the paucity of research and called for additional studies utilizing rigorous, quantitative methodology to support recommendations for best practice in schools. A similar suggestion was made more recently in a review of single-case design studies conducted in majority (> 50%) "ethnic–racial minority classrooms" (Long et al., 2019). Twenty-two studies were identified, authored from 1973 to 2014, and common behavioral interventions (e.g., group contingencies) were found to be generally effective, yet the review identified single case design studies only. The authors also focused the analysis on intervention characteristics. The current

study seeks to expand this approach by including (a) both group and single case design studies, as well as (b) details relevant to implementation (treatment integrity, treatment intensity).

Variables Associated with Treatment Integrity and Intensity

Data relevant to treatment integrity and treatment intensity are useful to understand the relationship between intervention implementation and treatment effects. Treatment integrity can be measured in several ways (e.g., direct observation, self-report), and can include estimates of various dimensions such as adherence and quality (Sanetti & Kratochwill, 2009). Treatment intensity is defined as the length, frequency and duration of an intervention (DeFouw et al., 2019). In evaluating the available quantitative research targeting intervention or practices to promote social and behavioral outcomes with racially and ethnically minoritized learners, it is imperative to analyze not only student outcomes (the dependent variables), but implementation (the independent variable) to determine what has been effective to date.

Purpose of Study

The purpose of this study was to produce a systematic replication of Authors (2012a) with quantitative studies from the past 10 years to identify articles published since the prior review. The authors searched for literature that described interventions or supports delivered in schools with samples of racially and ethnically minoritized students. Specifically, the research questions were as follows:

1. What quantitative research on interventions and practices to support social and behavioral outcomes with racially and ethnically minoritized youth in school settings has been published from 2010-2020?

2. Of the studies that met inclusion criteria, what are the study and student participant characteristics?

3. Of the studies that met inclusion criteria, what were the characteristics of the

interventions or practices provided?

4. Of the studies that met inclusion criteria, what was measured to determine

effectiveness? How was significance or magnitude of the intervention effect determined?

5. Of the studies that met inclusion criteria, what treatment integrity and treatment

intensity data were reported?

Method

Study Identification Procedures

Two advanced doctoral student coders located peer-reviewed journal articles and

dissertations through a two-step process. First, coders searched abstracts in PsycINFO,

PsycARTICLES, Educational Resources Information Center (ERIC), and ProQuest Dissertations

and Theses for articles published between 2010 – 2020 using the following search terms used in

Authors (2012b): classroom management OR disciplin* OR behavior* AND cultur* OR divers*

OR race, OR linguistic* OR ethnic*. The use of truncated terms (indicated with an asterisk) was

meant to capture relevant articles that may use variations of the root word (e.g., discipline,

disciplinary). In addition, the search was expanded to include additional, more specific terms

related to race and ethnicity: African American OR Black OR Hispanic OR Latino/a/x OR

Chicano/a OR Asian OR Native American OR Indigenous OR American Indian OR Native

Hawaiian OR Pacific Islander OR multiracial OR multiethnic OR multiple races. The search

process produced 1,687 studies.

Similar to the approach used in Authors (2012b) review, coders eliminated a study if it

did not (a) focus on an intervention or practice to support social and behavioral outcomes of

racially and ethnically minoritized youth in schools, (b) was published in a language other than

English, or (c) was conducted outside of the United States. Also, due to the research question

focused on implementation (i.e., treatment integrity and intensity), coders only gathered

quantitative (i.e., quasi-experimental or experimental group and single case design) studies. As a

result, 56 articles and dissertations were identified for screening. Coders also conducted an

ancestral search of these 56 reference lists to produce an additional 17 studies to screen to

produce a total of 73 studies to screen. (See Figure 1 for search and screening procedures.)

Screening Procedures

The primary coder (second author) screened the 73 studies for five inclusion criteria.

Specifically, studies had to include a (a) practice or intervention to impact a (b) social or

behavioral target as the dependent variable (e.g., disruptive behavior, problem behavior, on-task

behavior, academic engagement, bullying behavior, relational aggression, self-regulation,

conflict resolution), with a (c) racially and ethnically minoritized student sample (i.e., ≥ 50%

identifying as a race or ethnicity other than White) (d) in Grades K – 12, (e) at school (not home,

hospital, community, etc.). Upon screening, 32 studies met all inclusion criteria.

Study Coding Procedures

The first author created a coding manual to collect information to answer each research

question. For instance, to answer the first research question about general study characteristics,

items were generated for coders to indicate the type of research design, primary implementer(s)

of study interventions, and school setting. The manual was similar to that which was used in

previous systematic reviews (Authors, 2018a; Maggin et al., 2012; available upon request). The

first author provided a systematic training to coders by first describing each item within the

manual with coders. Then, coders reviewed a sample study that was coded using the manual. The

first author then answered all coder's questions prior to coder's independent practice. Coders

then rated a study independently with the goal of achieving greater than 90% agreement with the sample study.

Agreement was determined by dividing the number of agreements by the number of agreements and disagreements. This value was then converted to a percentage. Raters' agreement was 99.6% which exceeded this criterion. Discrepancies in coding were discussed until consensus was reached between the coders and first author. Coders then proceed to rate all studies in the review independently. Below is a description of how studies were coded.

Study Characteristics

Coders documented the study design (experimental group design, quasi-experimental group design or single case design) and the implementer's professional role (e.g., general education teacher, special education teacher, paraeducator, researcher). Coders noted the type of school setting (i.e., public, private, charter school) and the urbanicity of the school (i.e., urban, suburban, rural).

Student Participant Characteristics

Coders also reviewed student characteristics including (a) student grade level (i.e., elementary, middle, high), (b) the number of student participants, (b) the number of student English Learners (EL) in the sample, (c) the number of participants receiving special education, and (d) race/ethnicity (i.e., Hispanic/Latinx, African American/Black, White, Asian, Native Hawaiian/Pacific Islander, Multiracial, multiple races, other).

Intervention Characteristics

Coders documented the name of the intervention, reported intervention description, intervention setting (i.e., general education classroom, general education classroom with resource support, special education self-contained, separate room, other), and the method(s) of

intervention delivery applied (i.e., individual, group, classwide, schoolwide). Coders then noted

the primary dependent variable(s), reported methods of determining significance and/or the

magnitude of intervention effect (e.g., effect size, significance testing), and reported the

estimates across participants in the study.

Treatment Integrity and Treatment Intensity Characteristics

Coders documented (a) if treatment integrity was assessed, (b) the measurement

method(s) used to assess treatment integrity (e.g., direct observation, self-report, permanent

product), (c) the percentage of sessions for which treatment integrity was assessed and

percentage of treatment integrity reported overall, (d) the treatment integrity dimensions reported

(e.g., adherence, quality), and (e) the percentage of sessions for which interobserver agreement

(IOA) data was collected for treatment integrity. Coders also recorded the treatment intensity of

intervention implementation in each study, specifically the intervention (a) session length (i.e.,

the number of minutes of each session), (b) session frequency (i.e., number of sessions delivered

per day/week/month), and the (c) total treatment duration (i.e., in weeks).

Interrater Agreement

Fifteen randomly selected studies (20.5%) were double screened and coded by a second

member of the research team to calculate interrater agreement. Interrater agreement was

calculated as the total number of coding agreements (i.e., same choice selected) divided by the

total opportunities for agreement. Overall, the interrater agreement for the screening and coding

procedures were 95.8% and 93.6% respectively. Any discrepancies between coders were

reviewed as a team and discussed until consensus was met.

Data Analysis

Descriptive statistics (i.e., frequency counts, percentages) were used to summarize coding variables across the 32 studies that met inclusionary criteria.

<div align="center">

Results

</div>

Results are organized by research question. First, to address what quantitative research on interventions and practices to support social and behavioral outcomes with racially and ethnically minoritized youth in school settings has been published from 2010-2020, we report general study characteristics (Question 1). Then, we describe student participant characteristics (Question 2). Next, we summarize characteristics of the interventions or practices provided in the studies (Question 3). We then describe what was measured to determine effectiveness, and how was significance or magnitude of intervention effects were determined (Question 4). Finally, we describe treatment integrity and treatment intensity data reported in studies included in the review (Question 5).

Study and Implementer Characteristics

Overall, as depicted in Table 1, many of the 32 studies that met the inclusion criteria used a quasi-experimental ($n = 10$; 31.25%) or experimental ($n = 4$; 12.50%) group research design. However, the majority of studies coded were from single case research papers ($n = 18$; 56.25%), many utilizing a withdrawal/reversal design ($n = 7$; 21.86%). Often, in the studies coded, the implementers were teachers ($n = 19$; 59.36%) and/or researchers ($n = 12$; 37.50%). The majority of the studies occurred in a public school ($n = 23$; 71.86%) followed by a charter ($n = 5$; 15.63%) or alternative school ($n = 2$; 6.25%). One study occurred on an American Indian reservation ($n = 1$; 3.13%). Many studies took place in an urban setting ($n = 21$; 65.63%), although some occurred in a suburban school ($n = 3$; 9.36%). No study reportedly took place in a rural environment.

Student Participant Characteristics

The majority of studies included student participants in elementary grades ($n = 18$; 56.25%), although many studies occurred in secondary settings including middle ($n = 7$; 21.86%) and high ($n = 7$; 21.86%) schools. Students who participated in studies coded ranged in number from one to 1,791. In some studies, participants were English learners ($n = 10$; 31.25%) and/or were receiving special education services ($n = 14$; 43.75%). In addition, the race/ethnicity of student participants varied. In many studies, the participant sample was identified as mostly African American/Black ($n = 15$; 46.86%), non-White Latinx/Hispanic ($n = 12$; 36.50%) or Native American ($n = 2$; 6.25%).

Intervention Characteristics

As evidenced in Table 2, studies included a variety of evidence-based behavioral interventions such as the Good Behavior Game ($n = 5$; 15.63%) and Check-in, Check-out ($n = 3$; 9.36%), as well as mystery motivator, increasing opportunities to respond, implementing home-school notes, and offering choice in academic assignments (all $n = 1$; 3.16%). These interventions were typically not adapted in any way, but rather implemented with a sample of racially and ethnically minoritized students.

Many other studies ($n = 7$; 21.86%) utilized an adapted a social-emotional learning curriculum (e.g., *Strong Start*; Graves et al., 2017) to increase cultural relevance. In other studies, interventions were designed to be culturally relevant from inception. This includes *Sisters of Nia* ($n = 2$; 6.25%) and *Brothers of Ujima* ($n = 1$; 3.13%), two strengths-based small group interventions to teach about Afrocentric values and identity. Finally, non-behavioral interventions were coded including mindfulness-based interventions ($n = 3$; 9.36%) and restorative practices ($n = 2$; 6.25%). The implementer delivered the intervention most often

classwide ($n = 14$; 43.75%) or in a small group ($n = 9$; 28.13%) format within the general education classroom ($n = 16$; 50.00%) or a separate room ($n = 8$; 25.00%).

In studies reviewed, the dependent variables measured varied. Some researchers measured on-task/academically engaged ($n = 5$; 15.63%) and/or disruptive behavior ($n = 8$; 25.00%). Others reported teachers' perceptions of students' skills or students' perceptions of identity and/or emotions ($n = 14$; 43.75%). There was also variation in how it was determined if effects were significant and the magnitude of observed effects. Depending on the design, some researchers reported significance based on a t test ($n = 2$; 6.25%) or analysis of variance ($n = 3$; 9.36%). Others calculated effect size estimates to determine the magnitude of an observed effect ($n = 16$; 50.00%). Still others utilized visual analysis and a descriptive approach to analyzing treatment effects ($n = 4$; 12.50%).

Treatment Integrity and Treatment Intensity Characteristics

As depicted in Table 3, most studies ($n = 24$; 75.00%) reported if treatment integrity was assessed throughout the intervention. Of these studies, the most common measurement method included direct observation with a checklist ($n = 19$; 59.38%) and self-report checklist ($n = 7$; 21.88%). Only a few studies assessed a dimension other than adherence ($n = 22$; 68.75%); two studies each assessed treatment quality (e.g., interventionist enthusiasm, fluency; 6.25%) and treatment exposure (e.g., number of sessions, duration; 6.25%). Additionally, of the 24 studies that monitored treatment integrity, nine (28.13%) also reported IOA of treatment integrity data.

The majority of studies reported aspects of treatment intensity related to the session length, or the number of minutes per session ($n = 22$; 68.75%). Sessions ranged from five to 90 min. The number of sessions per week was reported often ($n = 26$, 81.25%). Intervention or practice sessions often occurred once or twice a week ($n = 13$; 40.61%), but often more

frequently (e.g., 5x/week; $n = 11$; 34.38%). Finally, the total treatment duration, or the total number of weeks, was reported in most studies ($n = 26$, 81.25%). Intervention or practices implemented in studies reviewed ranged in duration from three weeks (Featherston, 2014) to three years (Usera, 2017).

Discussion

Students receiving public education in the United States are entitled to receive equal protection under the law and not be subject to discriminatory school discipline policies and practices. Yet, despite legal protections, evidence continues to indicate that racially and ethnically minoritized students are disproportionality disciplined (Gage et al., 2020, Gage et al., 2020), and that achievement is also impacted (Welsh & Little, 2018). To reduce the use of exclusion from the classroom environment, researchers have conducted systematic research reviews to provide suggestions to educators about how to design more inclusive, supportive learning environments (e.g., Long et al., 2019). One such review by Authors (2012a) provided many recommendations for culturally relevant classroom behavior support practices, but studies reviewed only included research published prior to 2010. In addition, these recommendations were drawn from primarily qualitative studies and expert recommendations. The current study is a systematic replication of Authors (2012a) with a specific emphasis on quantitative research and implementer action specifically.

Results indicate that coders located 32 quantitative research studies that met inclusion criteria. Inclusion criteria required that studies include a behavior management practice or intervention, a behavioral target as the dependent variable, and include a school-based sample with most participants being racially and ethnically minoritized youth ($\geq 50\%$ identifying as a race or ethnicity other than White). After screening, 32 studies were coded. Roughly half utilized

a single case research design whereas the other half employed quasi-experimental or experimental group designs. Although in the current review search terms were expanded to ensure all relevant articles were located, results stand in contrast to what was located in the previous review (2012a), indicating that there may be much more quantitative research on this topic published in the past 10 years than the 20 years prior.

Results also indicated that teachers were most often implementers of study interventions or practices. This is relevant as generalizing the effectiveness of interventions studied in applied settings hinges on the capacity of school-based personnel to implement the interventions with the resources required (i.e., without support from researchers who may not always be involved or available). In the future, research might seek to gauge implementers' perceptions of feasibility due to its relationship with higher levels of implementation integrity and ultimately better outcomes (Gadke et al., 2020).

Studies primarily took place in urban, public elementary schools, although nearly 44% of studies occurred in secondary settings. This was a surprising finding given the paucity of school-based research in middle and high schools (e.g., Sansosti et al., 2010). It appears based on significance testing and effect size estimates reported that interventions were effective across these settings.

Many students were Black/African American or Hispanic/Latinx and ranged in number from one to over 1,500. It was uncommon for researchers to report English Learner and disability status. This was unexpected given the relevance of the demographic information to the study topics. However, these topics are not listed explicitly as important characteristics of participants by journal reporting standards in psychology (Appelbaum et al., 2018). Future

researchers might report these details, though, to continue to expand the literature base for effective behavioral support practices for English Learners and students with disabilities.

The type of intervention implemented varied, but as described in the results, several studies provided students with established evidence-based behavioral interventions. This included the Good Behavior Game, which has a wealth of research support (Tingstrom et al., 2006), including positive outcomes in classrooms with minoritized students (Nolan et al., 2014). No adaptions were made to increase its cultural relevance. Based on analyses of the five studies in which it was implemented, two reported strong treatment effects. One study found large effect size estimates for two students who decreased interrupting behaviors while playing the game in class (Ortiz et al., 2017). The other study found strong effect sizes for students' externalizing behaviors, internalizing behaviors, and well-being upon receiving the Good Behavior Game with an embedded mindfulness intervention (Long et al., 2018).

Some evidence-based interventions were adapted to be more culturally relevant. For instance, Behavior Bingo (Collins et al., 2018) was an interdependent group contingency intervention (an evidence-based practice; Maggin et al., 2017) adapted to incorporate a student preference assessment to identify rewards for displays of behavioral expectations (rather than relying on teacher assumptions). It also included culturally relevant peer models (i.e., students of similar age, ethnicity, gender) to promote positive behavior. Although effect size estimates were not reported, descriptive data and visual analysis demonstrated the intervention was effective in two classrooms upon decreasing average displays of on-task and disruptive behavior and increasing average displays of on-task behavior.

As described in the results, some implementers provided participants with an adapted social-emotional learning curriculum. For instance, in Graves et al. (2017), *Strong Start*

(Whitcomb & Merrell, 2012) was adapted by altering some of the instructional materials (e.g.,

books read during lessons) to increase the lesson's relevance to the Black/African American

participant sample. Results were mixed as researchers noted large effect size for variables such

as social regulation but not for externalizing behaviors. In other studies, interventions were

designed to be culturally relevant including *Sisters of Nia* and *Brothers of Ujima* which teach

about Afrocentric values and identity. For *Sisters of Nia*, authors found strong effect sizes for

participants' perceptions of social strengths in one study (Aston & Graves, 2016) and decreases

in verbal aggression in another (Aston et al., 2017). For *Brothers of Ujima,* strong effect sizes

were noted for social resiliency (Graves & Aston, 2018).

Most often, studies took place in the general education environment, although some

supports were delivered classwide, while others were targeted to small groups in a separate

setting. In many studies reviewed, intervention implementation aimed to reduce displays of less

desirable behaviors such as disruptions (e.g., Bunch-Crump, 2016) and aggression (e.g., Farrell

et al., 2018), rather than increase adaptive or more desirable behaviors.

The current study sought to code variables related to implementer action specifically.

This included implementers' treatment integrity and intensity. Many studies provided

information about implementers' treatment integrity despite research to indicate that treatment

integrity data have not consistently been reported in the past (Sanetti et al., 2012). Most

commonly, in the studies reviewed, implementers were observed directly. Direct observation of

implementation is considered to be the most reliable method of treatment integrity assessment

(Sanetti & Kratochwill, 2009). In 12 of the 32 studies, treatment integrity was assessed during

each intervention session and adherence was found to be high (88-100% of intervention steps

were implemented, on average). This indicates treatments were delivered as planned

consistently. Less often, interobserver agreement of treatment integrity data collection was reported. When it was, researchers reported IOA of at least 20% of sessions (e.g., Graves et al., 2017). In addition, when reported, the intensity of interventions provided ranged from a few minutes to an hour a day minutes, from once a week or every day and over six weeks to up to one year. Strong treatment effects were noted for studies with a range of treatment intensity characteristics.

Limitations

There are limitations to consider when interpreting results of the current review. First, we conducted database and hand searches systematically, yet it is possible that studies that meet criteria were missed. For instance, if the race and ethnicity of student participants were either not reported or not reported clearly, coders may not have located the study with the search procedures applied. However, we made every effort to locate relevant research. In addition, we were interested in a review of quantitative studies, as there has been a dearth of such research reported in the past (Authors, 2012a). However, this meant that qualitative studies were excluded from the current review. Qualitative studies excluded were largely descriptive and investigated topics such as student-teacher interactions (Scott, 2010) and student preferences for classroom management practices (Hubbard, 2015). Future research might include a synthesis of qualitative studies to supplement the findings of the current review.

Furthermore, our review is a systematic replication of Authors (2012a) and is therefore descriptive. Thus, we report study estimates from significance testing and effect size calculations rather than computing these estimates independently. This was due to the wide breadth of independent variables (i.e., interventions or practices) and dependent variables included in the

studies reviewed. This approach was suitable for addressing the study research questions which, broadly, inquired about characteristics of research published in the last 10 years.

Implications

Implications for research, practice and policy are presented below.

Research

As described above, future research might extend the results of the current study to include qualitative and mixed methods studies. We chose to synthesize quantitative studies to gauge if more quantitative research had been conducted since Authors (2012a) was published (as the need for additional quantitative research was a primary recommendation in that paper). We also focused on quantitative works due to our interest in implementer action (i.e., treatment integrity, treatment intensity). However, qualitative data may capture implementer and student perceptions (e.g., of acceptability, alignment with culture). Such perceptions could be very useful to guiding recommendations for school-based practice. Additionally, future reviews might target the utility of certain interventions implemented in the studies more comprehensively. For instance, several studies reviewed included the implementation of the Good Behavior Game. Although the Good Behavior Game is well-researched and has been called a "behavioral vaccine" (Embry, 2002, p. 273), there is no known comprehensive, systematic review of its impact on the behavior of students from racially and ethnically minoritized backgrounds. Such a review could also guide recommendations pertaining to how it might be modified to be more culturally and contextually relevant.

In addition, replications of studies found to be effective in the current review would expand the literature base. Specifically, studies that target increasing proactive classroom behaviors (versus reducing negative behaviors; e.g., aggression) would broaden the impact of

research efforts to promote positive, inclusive learning environments. Researchers might also

develop interventions that incorporate elements of Gregory and colleagues (2017) framework

which included incorporating student and family voice and building supportive relationships in

the classroom. Quantitative evaluation of such approaches would add to the scope of the

literature base.

Practice

Based on the results of the current review, there is preliminary evidence that when

empirically-supported behavioral interventions are implemented with high levels of treatment

integrity, students from racially and ethnically minoritized backgrounds benefit. It is important to

note that having the time required to deliver practices and interventions will likely impact

intervention effectiveness. In the current review, studies that reported data related to treatment

intensity indicated session length was on average 40 min at least once per week. School leaders

must provide the time and training required to implement supports with appropriate levels of

treatment integrity and intensity to promote the best outcomes for youth.

Results also demonstrated examples of cultural adaptions to established interventions that

successfully promote students' social and behavioral outcomes. This is aligned with Gregory and

colleagues' (2107) recommendation to provide culturally responsive teaching and behavioral

instruction. When classroom instruction and behavior supports reflect students' culture, a sense

of belonging and academic engagement can result. Studies reviewed provided a variety of

examples of culturally adapted and culturally responsive practices and interventions. These

examples may help educators create an environment in which behavior is supported proactively

and exclusionary discipline is not used, a step critical to addressing discipline disproportionality

and preventing the long-term determinantal effects of time out of class, suspension, etc. (Wolf &

Kupchik, 2017).

Policy

Results of the current review indicate that educators might support discipline policies to

shift to reflect more positive, proactive behavioral support to engage learners of racially-

ethnically minoritized background in instructional environments. These supports might be

included as part of a school-, district- and/or state-wide initiative to implement a multi-tiered

system of support framework (e.g., PBIS) in accordance with federal policy such as the

Individuals with Disabilities Education Act (IDEA) of 2004 and Every Student Succeeds Act

(ESSA) of 2015. These federal policies encourage school staff to implement a multi-tiered

systems of support framework to improve equitable access to instruction in schools.

Implementation can be considerate of student culture by including student, family and

community voice in determining schoolwide expectations, acknowledgment procedures, as well

as how to respond to behavior that interferes with learning (e.g., reteaching expectations)

(Martinez et al., 2019). With implementing MTSS comes the need for resources to provide staff

with adequate training, time and support to engage in evidence-based practices effectively.

According to the results of the current study, this includes appropriate training to provide

supports with high levels of treatment integrity and intensity, as well as personnel to coordinate

the collection and monitoring of implementation and student outcome data.

Conclusion

Authors (2012a) conducted a literature review of studies that, up until 2010, addressed

culturally relevant classroom behavior support practices in schools. The authors acknowledged

that schools are becoming increasingly diverse and discipline disproportionality persists, yet

high-quality, rigorous research-based recommendations for support practices to promote social and behavioral outcomes in schools was limited. Results of this study indicate there has been an increase in quantitative research studies published pertaining to this topic in the past decade. Furthermore, of the studies that met inclusion criteria, many implemented well-researched behavior interventions (e.g., Good Behavior Game, mystery motivator, offering choice), but some with cultural adaptions. In other studies, new interventions were introduced, but also were reported to be effective. Teachers and researchers primarily provided the studies in a range of public school environments.

Researchers demonstrated statistically significant and large treatment effects for decreases in verbal aggression, externalizing behaviors, and increases in academic engagement, on-task behavior and self-regulation. However, several studies also reported non-significant findings or small effect size estimates. Without treatment integrity and intensity data for all studies, it is difficult to determine if non-effects were result of an ineffective intervention or poor implementation. Overall, it is important to continue this line of inquiry to improve recommendations for both policy and practice regarding ways to foster educational environments that are supportive to a variety of individuals in a range of contexts.

References

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018).

    Journal article reporting standards for quantitative research in psychology: The APA

    publications and communications board task force report. *American Psychologist, 73*(1),

    3-25. https://doi.org/10.1037/amp0000191

*Aston, C., & Graves Jr., S. L. (2016). Challenges and barriers to implementing a school-based

    Afrocentric intervention in urban schools: A pilot study of the Sisters of Nia Cultural

    Program. *School Psychology Forum, 10*(2), 165-176.

*Aston, C., Graves Jr., S. L., McGoey, K., Lovelace, T., & Townsend, T. (2017). Promoting

    sisterhood: The impact of a culturally focused program to address verbally aggressive

    behaviors in Blacks girls. *Psychology in the Schools, 55*(1), 50-62.

    https://10.1002/pits.22089

Authors (2012a). Blinded for review.

Authors (2012b). Blinded for review.

Authors (2018a). Blinded for review.

Authors (2018b). Blinded for review.

*Barrasso, M. M. (2020). *Supporting teachers to use restorative practice circles to improve

    student outcomes in elementary school classrooms*. (Doctoral dissertation.) University of

    Massachusetts Boston.

*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*(7), 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*(1), 90-104. https://doi.org/10.1007/s12671-014-0376-1

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration
of factors contributing to the overrepresentation of black students in office disciplinary
referrals. *Journal of Educational Psychology*, *102*(2), 508-
520. https://doi.org/10.1037/a0018450

Brenda, L. (2000). The disproportionate discipline of African American learners: Reducing
school suspensions and expulsions. *Exceptional Children, 66*(3), 381-391.
https://doi.org/10.1177/001440290006600308

*Bunch-Crump, R. K. (2015). *The effects of Tier II Check-In Check-Out intervention and Tier III
function-based self-management on the disruptive behavior and academic engagement of
selected African American male students*. [Doctoral dissertation, The University of North
Carolina at Charlotte].

*Castro-Olivo, S. M., & Merrell, K. W. (2012). Validating cultural adaptations of a school-based
social-emotional learning programme for use with Latino immigrant adolescents.
Advances in School Mental Health Promotion, 5(2), 78-92.
https://doi.org/10.1080/1754730X.2012.689193

*Castro-Olivo, S., Preciado, J., Le, L., Marciante, M., & Garcia, M. (2018). The effects of
culturally adapted version of First Steps to Success for Latino English language learners:
Preliminary pilot study. *Psychology in the Schools*, *55*(1), 36-49.
https://doi.org/10.1002/pits.22092

*Collins, T. A., Hawkins, R. O., Flowers, E. M., Kalra, H. D., Richard, J., & Haas, L. E. (2018).
Behavior Bingo: The effects of a culturally relevant group contingency intervention for

students with EBD. *Psychology in the Schools*, *55*(1), 63-75.

https://doi.org/10.1002/pits.22091

*Cramer, K. M., & Castro-Olivo, S. (2016). Effects of a culturally adapted social-emotional

learning intervention program on students' mental health. *Contemporary School

Psychology, 20*, 118-129. https://doi.org/10.1007/s40688-015-0057-7

Davis, J. E., & Jordan, W. J. (1994). The effects of school context, structure, and experiences on

African American males in middle and high school. *The Journal of Negro

Education*, *63*(4), 570-587. https://doi.org/10.2307/2967296

DeFouw, E. R., Codding, R. S., Collier-Meek, M. A., & Gould, K. M. (2019). Examining

dimensions of treatment intensity and treatment fidelity in mathematics intervention

research for students at risk. *Remedial and Special Education*, *40*(5), 298-312.

https://doi.org/0741932518774801

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal

behavioral vaccine. *Clinical Child and Family Psychology Review*, *5*(4), 273-297.

https://doi.org/10.1023/A:1020977107086

*Farrell, A. D., Sullivan, T. N., Sutherland, K. S., Corona, R., & Masho, S. (2018). Evaluation of

the Olweus Bully Prevention Program in an urban school system in the USA. *Prevention

Science, 19*(6) 833-847. https://doi.org/10.1007/s11121-018-0923-4

*Featherston, T. R. (2014). *An experimental study on the effectiveness of a restorative justice

intervention on the social aggression, social problem solving skills, and prosocial

behaviors of African American adolescent girls*. [Doctoral dissertation, Capella

University].

Gadke, D. L., Kratochwill, T. R., & Gettinger, M. (2020). Incorporating feasibility protocols in

    intervention research. *Journal of School Psychology*, *84*, 1-18.

    https://doi.org/10.1016/j.jsp.2020.11.004

Gage, N. A., Katsiyannis, A., Carrero, K. M., Miller, R., & Pico, D. (2020). Exploring

    disproportionate discipline for Latinx students with and without disabilities: A national

    analysis. *Behavioral Disorders*, https://doi.org/10.1177/0198742920961356

Gage, N. A., Whitford, D. K., Katsiyannis, A., Adams, S., & Jasper, A. (2019). National

    Analysis of the Disciplinary Exclusion of Black Students with and without

    Disabilities. *Journal of Child and Family Studies*, *28*(7), 1754-1764.

    https://doi.org/10.1007/s10826-019-01407-7

*Garcia-Dubon, E. R. (2018). *Exploring the impact of increased opportunities to respond on the

    behavior of middle school students with emotional and behavior disorders*. [Doctoral

    dissertation, University of California Los Angeles].

Girvan, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2017). The relative contribution of

    subjective office referrals to racial disproportionality in school discipline. *School

    Psychology Quarterly*, *32*(3), 392-404. https://doi.org/10.1037/spq0000178

*Graves S. L, Aston C. (2018). A mixed-methods study of a social emotional curriculum for

    black male success: A school-based pilot study of the Brothers of Ujima. *Psychology in

    the Schools, 55*(1), 76-84. https://doi.org/10.1002/pits.22088

*Graves Jr., S. L., Herndon-Sobalvarro, A., Nichols, K., Aston, C., Ryan, A., Blefari, A., ... &

    Prier, D. (2017). Examining the effectiveness of a culturally adapted social-emotional

    intervention for African American males in an urban setting. *School Psychology

    Quarterly*, *32*(1), 62-74. http://dx.doi.org/10.1037/spq0000145

Gregory, A., Skiba, R. J., & Mediratta, K. (2017). Eliminating disparities in school discipline: A

framework for intervention. *Review of Research in Education*, *41*(1), 253-278.

https://doi.org/10.3102/0091732X17690499

Hubbard, J. H. (2015). *Classroom management practices for male African American students

with behavioral disorders.* [Doctoral dissertation, Walden University].

 *Kleinman, K. E. & Saigh, P. A. (2011). The effects of the Good Behavior Game on the conduct

of regular education New York City high school students. *Behavior Modification, 35*(1),

95-105. https://doi.org/10.1177/0145445510392213

*Kowalewicz, E. A. & Coffee, G. (2014). Mystery motivator: A Tier 1 classroom behavioral

intervention. *School Psychology Quarterly, 29*(2), 138-156.

https://doi.org/10.1037/spq0000030

*Lambert, A. M., Tingstrom, D. H., Sterling, H. E., Dufrene, B. A., & Lynne, S. (2015). Effects

of tootling on classwide disruptive and appropriate behavior of upper-elementary

students. *Behavior Modification, 39*(3), 413-430.

https://doi.org/10.1177/0145445514566506

*Long, A. C. J., Miller, F. G., & Upright, J. J. (2019). Classroom management for ethnic–racial

minority students: A meta-analysis of single-case design studies. *School Psychology,

34*(1), 1-13. http://dx.doi.org/10.1037/spq0000305

*Lopach, L. C. (2016). The Electronic Daily School Note: *A study examining an evidence-based

school intervention package for improving on-task behavior, academics, and home-

school collaboration.* [Doctoral dissertation, University of Utah].

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A

systematic evidence review of school-based group contingency interventions for students

with challenging behavior. *Journal of School Psychology*, *50*(5), 625-654.

https://doi.org/10.1016/j.jsp.2012.06.001

Maggin, D. M., Pustejovsky, J. E., & Johnson, A. H. (2017). A meta-analysis of school-based

group contingency interventions for students with challenging behavior: An

update. *Remedial and Special Education*, *38*(6), 353-370.

https://doi.org/10.1177/0741932517716900

Marchbanks III, M. P. T., Peguero, A. A., Varela, K. S., Blake, J. J., & Eason, J. M. (2018).

School strictness and disproportionate minority contact: Investigating racial and ethnic

disparities with the "School-to-Prison Pipeline." *Youth Violence and Juvenile*

Justice, *16*(2), 241-259. https://doi.org/10.1177/1541204016680403

Martinez, S., Kern, L., Hershfeldt, P., George, H. P., White, A., Flannery, B., Freeman, J. (2019).

High school PBIS implementation: Student voice. Office of Special Education Technical

Assistance Center on Positive Behavior Interventions and Support, University of Oregon.

Retrieved from www.pbis.org.

*Mitchell, R. R. (2014). *Evaluating the effectiveness of the Good Behavior Game with general*

education high school students utilizing a changing criterion component. [Doctoral

dissertation, University of Southern Mississippi].

*Mohn, T. (2019). *Effects of Check-In Check-Out intervention for an African American child*

with Attention Deficit Hyperactivity Disorder. [Doctoral dissertation, The Chicago

School of Professional Psychology].

Morris, E. W., & Perry, B. L. (2016). The punishment gap: School suspension and racial

disparities in achievement. *Social Problems*, *63*(1), 68-86.

https://doi.org/10.1093/socpro/spv026

Mowen, T., & Brent, J. (2016). School discipline as a turning point: The cumulative effect of

suspension on arrest. *Journal of Research in Crime and Delinquency*, *53*(5), 628-653.

https://doi.org/10.1177/0022427816643135

Nolan, J. D., Houlihan, D., Wanzek, M., & Jenson, W. R. (2014). The Good Behavior Game: A

classroom-behavior intervention effective across cultures. *School Psychology

International*, *35*(2), 191-205. https://doi.org/10.1177/0143034312471473

*Ortiz, J., Bray, M. A., Bilias-Lolis, E., & Kehle, T. J. (2017). The Good Behavior Game for

Latino English Language Learners in a small-group setting. *International Journal of

School & Educational Psychology*, *5*(1), 26-38.

https://doi.org/10.1080/21683603.2016.1169961

Rausch, M. K., & Skiba, R. J. (2006). Exclusion is not the only alternative: The children left

behind project. In A. Reyes (Ed.), *Discipline, achievement, and race: Is zero tolerance

the answer?* (pp. 105-126). Rowman & Littlefield Publishing Group.

*Robinson-Ervin, P. D. (2012). *The effects of culturally responsive computer-based social skills

instruction on the social skill acquisition and generalization of urban 6th-grade students

with emotional and behavioral disorders.* [Doctoral dissertation, The Ohio State

University].

*Ryan, A., Graves Jr, S., Sobalvarro, A., Nichols, K., Schutte, K., Aston, C., & Griffin, A.

(2016, June). An evaluation of Strong Kids in an urban African American female sample:

The need for gender-specific and culturally focused interventions. *School Psychology

Forum*, 10(2), 157-164.

Sanetti, L. M. H., Dobey, L. M., & Gritter, K. L. (2012). Treatment integrity of interventions

with children in the *Journal of Positive Behavior Interventions* from 1999 to

2009. *Journal of Positive Behavior Interventions*, *14*(1), 29-46.

https://doi.org/10.1177/1098300711405853

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment

integrity: Introduction to the special series. *School Psychology Review*, *38*(4), 445-459.

https://doi.org/10.1037/14275-001

Sansosti, F. J., Noltemeyer, A., & Goss, S. (2010). Principals' perceptions of the importance and

availability of response to intervention practices within high school settings. *School

Psychology Review*, *39*(2), 286-295. https://doi.org/10.1080/02796015.2010.12087779

*Simonsen, B., Myers, D., & Briere III, D. E. (2011). Comparing a behavioral check-in/check-

out (CICO) intervention to standard practice in an urban middle school setting using an

experimental group design. *Journal of Positive Behavior Interventions*, *13*(1), 31-48.

https://doi.org/10.1177/1098300709359026

*Skerbetz, M. D. & Kostewicz, D. E. (2013). Academic choice for included students with

emotional and behavioral disorders. *Preventing School Failure, 57*(4), 212-222.

https://doi.org/10.1080/1045988X.2012.701252

Scott, M. D. R. (2010). *Teacher practices while interacting with preschoolers in inclusive

settings.* [Doctoral dissertation, University of North Carolina, Greensboro].

Skiba, R. J., Chung, C. G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014).

Parsing disciplinary disproportionality: Contributions of infraction, student, and school

characteristics to out-of-school suspension and expulsion. *American Educational

Research Journal*, *51*(4), 640-670. https://doi.org/10.3102/0002831214541670

Skiba, R. J., Horner, R. H., Chung, C., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is

not neutral: A national investigation of African American and Latino disproportionality

in school discipline. *School Psychology Review, 40*(1), 85-

107. https://doi.org/10.1080/02796015.2011.12087730

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline:

Sources of racial and gender disproportionality in school punishment. *The Urban

Review*, *34*(3), 317-342. https://doi.org/10.1023/A:1021320817372

*Speight, D. R. (2018). *Impact of the implementation of Class-Wide Function-related

Intervention Teams on adolescent student behavior.* [Doctoral dissertation, University of

Arkansas, Fayetteville].

*Stoll-Juredine, N. (2015). *Goodwill Girls: Examining the effectiveness of a relational

aggression intervention with predominantly African American females*. (Doctoral

dissertation.) The Ohio State University.

*Tanol, G., Johnson, L., McComas, J., & Cote, E. (2010). Responding to rule violations or rule

following: A comparison of two versions of the Good Behavior Game with kindergarten

students. *Journal of School Psychology, 48*(5), 337-355.

https://doi.org/10.1016/j.jsp.2010.06.001

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The Good Behavior

Game: 1969-2002. *Behavior Modification*, *30*(2), 225-253.

https://doi.org/10.1177/0145445503261165

*Usera, J. J. (2017). The efficacy of an American Indian culturally-based risk prevention

program for upper elementary school youth residing on the northern plains

reservations. *The Journal of Primary Prevention*, *38*(1-2), 175-194.

https://doi.org/10.1007/s10935-016-0462-3

Welsh, R. O., & Little, S. (2018). The school discipline dilemma: A comprehensive review of

   disparities and alternative approaches. *Review of Educational Research*, *88*(5), 752-794.

   https://doi.org/10.3102/0034654318791582

Whitcomb, S. A., & Merrell, K. W. (2012). Understanding implementation and effectiveness of

   Strong Start K-2 on social-emotional behavior. *Early Childhood Education*

   Journal, *40*(1), 63-71.

   https://doi.org/10.1007/s10643-011-0490-9

*Willenbrink, J.B. (2018). Effects of a mindfulness-based program on children's social skills,

   problem behavior, and emotion regulation. [Doctoral dissertation, University of

   Wisconsin-Milwaukee].

Wolf, K. C., & Kupchik, A. (2017). School suspensions and adverse experiences in

   adulthood. *Justice Quarterly*, *34*(3), 407-430.

   https://doi.org/10.1080/07418825.2016.1168475

Yeaton, W. H., & Sechrest, L. (1981). Meaningful measures of effect. *Journal of Consulting and*

   Clinical Psychology, 49(5), 766–767. https://doi.org/10.1037/0022-006X.49.5.766

Table 1
Study Characteristics

Authors	Study Characteristics				Student Characteristics				
	Study Design[1]	Implementer Type[2] and Number (n)	School Setting	Reported Urbanicity	Grade	Number of Participants	Number of English Learner Participants	Number of Participants Receiving SPED	% Participant Race/Ethnicity[3]
Aston & Graves (2016)	QED	R (2)	Public	Urban	Elem	5	-	-	100% AA
Aston et al. (2017)	SCD - MB	R (1)	Charter	Urban	Middle	12	-	-	100% AA
Barrasso (2020)	SCD - MB	T (4)	Public	Urban	Elem	101	-	11	48% H 29% W 11% AA 7% O, 6% M
Black & Fernando (2014)	QED	MI (2), T (17)	Public	Urban	Elem	409	-	-	52% H 28 % B 15 % A 4 % W <1 % O
Bluth et al. (2016)	QED	R (1), MI (1)	Alternative	-	High	27	0	-	56% H 19% AA 19% W 7% O
Bunch-Crump (2016)	SCD – W/R, MB	T (1), SPED (1), AP (1)	Public	Urban	Elem	3	-	1	100% AA
Castro-Olivo & Merrell (2012)	QED	T (2)	Public	-	High	40	8	-	100% H
Castro-Olivo et al. (2018)	SCD - MB	R (1), T (3), P (3)	-	Suburban	Elem	3	3	-	100% H
Collins et al. (2018)	SCD W/R	T (1)	Alternative	Urban	High	15	-	15	60% AA 27% W 13% H

Cramer & Castro-Olivo (2016)	QED	R (3)	Public	Suburban	High	34	-	-	75% H 15% AA 3% W
Farrell et al. (2018)	SCD - MB	-	Public	Urban	Middle	1791	-	-	15% H 81% AA
Featherston (2014)	ED	R (1)	Charter	-	High	48	-	-	100% AA
Garcia-Dubon (2018)	SCD - W/R	T (1)	Public	Urban	Middle	5	0	5	80% H 20% AA
Graves & Aston (2018)	QED	R (1), T (4)	Charter	Urban	Middle	14	-	-	100% AA
Graves et al. (2017)	ED	R (4)	Public	Urban	Elem	61	-	-	100% AA
Kleinman & Saigh (2011)	SCD - W/R	T (1)	Public	Urban	High	26	-	0	73% H 23% AA 4% NR
Kowalewicz & Coffee (2014)	SCD – WR, CC	T (10)	Public	Urban Suburban	Elem	188	13	15	45% H 33% W 9% AA 8% M 4% A 1% O
Lambert (2015)	SCD - W/R	T (2)	Public	-	Elem	36	-	-	11% H 30% AA 3% A 50% W
Long et al. (2018)	QED	T (1)	Public	Urban	Elem	73	0	-	99% AA 1% W
Lopach (2017)	SCD - MP	T (2)	Public	Urban	Elem	4	0	-	25% H 25% W 25% NH/PI 25% M
Mitchell (2014)	SCD – W/R, CC	T (3)	Public	-	High	63	-	9	2% H 86% AA 11% M

									2% W
Mohn (2019)	SCD –W/R	T (5), PE (1), SSP (1), O (1)	Public	Suburban	Elem	1	1	0	100% AA
Ortiz et al. (2017)	SCD - MB	R (1)	Public	Urban	Elem	2	2	2	100% H
Robinson-Ervin (2012)	SCD - MP	Computer-facilitated	Public	Urban	Middle	6	-	6	100% AA
Ryan et al. (2016)	QED	R (2)	Public	Urban	Elem	39	-	-	100% AA
Simonsen et al. (2011)	ED	SSP (7), R (3)	Public	Urban	Middle	42	-	7	79% H 12% W 7% AA 2% NA
Skerbetz & Kostewicz (2013)	SCD - W/R	T (1) R (1)	Charter	Urban	Elem	5	-	5	100% AA
Speight (2018)	SCD – W/R, CC	T (4)	Public	-	Middle	8	8	4	75% H 25% NH/PI
Stoll-Juredine (2017)	ED	T (3)	Public	Urban	Elem	17	-	-	76% AA 74% W
Tanol et al. (2010)	SCD – W/R	T (2)	Public	-	Elem	6	-	1	100% NA
Usera (2017)	QED	-	American Indian Reservation	-	Elem	1531 (pre) / 1145 (post)	-	-	87% NA
Willenbrink (2019)	QED	MI (1)	Charter	Urban	Elem	188	-	-	54% H 39% AA 7% O

Note. [1]Designs included ED = experimental group design, QED = quasi-experimental group design, SCD = single-case design (MB = multiple baseline, W/R = withdrawal or reversal, CC = changing criterion, MP = multiple probe); [2]Numbers of implementers by type included R = researcher, T = teacher, SPED = special education teacher, AP = assistant principal, MI = mindfulness instructor, PE = paraeduator, PA = parent, SSP = school support personnel (in Simonsen et al. (2011), that included three school counselors, a social worker and three graduate student interns); [3]% Student race/ethnicity included AA = African American/Black, H = Hispanic/Latinx, NA = Native American, NH/PI = Native Hawaiian/Pacific Islander, A = Asian, W = white, M = multiple races, O = other.

Table 2
Intervention Characteristics

Authors	Intervention	Reported Intervention Description	Intervention Setting	Delivery Method	Primary Dependent Variables	Effect Size Estimate or Significance Test	Reported Estimates Across Participants
Aston & Graves (2016)	*Sisters of Nia*	Cultural enhancement intervention for African American females in early adolescence to increase cultural values and beliefs	Library during lunch	Small group	Perceptions of social strengths Perceptions of racial identity Perceptions of ethnic identity Perceptions of global worth Perceptions of physical attractiveness Perceptions of scholastic competence	Dependent sample t-tests and effect size estimates	$t = 24.61$ (.010)* $t = 4.96$ (.008)* $t = 2.46$ (.070) $t = 22.77$ (.050)* $t = 27.20$ (.002) $t = 24.00$ (.016)*
Aston et al. (2017)	*Sisters of Nia*	See AG2016	-	Small group	Verbal aggression	PND, TauU	PND $= 63 - 100$ TauU $= .85 - 1.00$**
Barrasso (2020)	Restorative Practices Circles	Community-building intervention; students form a circle and engage in discussions to strengthen relationships	GenED classroom	Classwide	Academic engagement after circle Disruptive behavior after circle	IRD, TauU	IRD $= .21 - .53$ TauU $= .14 - .68$ IRD $= .10 - .55$ TauU $= .02 - .31$
Black & Fernando (2014)	*Mindful Schools* (MS) Curriculum	5-week curriculum with lessons on awareness of breathing, visualization, thoughts, gratitude, and emotions	GenED classroom	Classwide	Paying attention Self-control Participation Respect for others	Cohen's d corrected for within-subjects dependence[1]	$d = .46$ $d = .43$ $d = .39$ $d = .35$
Bluth et al. (2016)	*Learning to BREATHE* (L2B) Mindfulness Curriculum	Derived from Mindfulness Based Stress Reduction to teach adolescents skills (e.g., body scan, meditation, mindful movement)	Classroom and gymnasium	Large group	Mindfulness Self-compassion Social connectedness Perceived stress Anxiety Depression	Hedge's g (95% confidence interval)	$g = 0.51$ (-0.37, 1.40) -0.04 (-0.91, 0.83) -0.23 (-1.11, 0.64) 0.46 (-0.43, 1.34) -0.29 (-1.18, 0.59) -1.26 (-2.21, -0.30)
Bunch-Crump (2016)	Check-in Check-out (CICO) and I-Connect (IC)	CICO is a daily behavior report card with check ins; IC is a function-based self-management intervention	GenED classroom / Separate room	Individual	Disruptive behavior Academic engagement	Reported visual analysis, means and standard deviations only	-
Castro-Olivo & Merrell (2012)	Culturally-adapted *Strong Teens* Curriculum	Adapted SEL curriculum with translation to Spanish and addition of concepts (e.g., ethnic pride, acculturative stress)	GenED classroom	Classwide	*Strong Teens* knowledge test Acculturative stress Sense of school belonging Internalizing concerns (mental health)	Cohen's d	$d = .98$** $d = .20$ $d = .37$ $d = .15$

| Castro-Olivo et al. (2018) | Culturally-adapted *First Steps to Success (FSS)* | Adapted SEL curriculum with translation to Spanish and additional concepts (e.g., Latino customs and folk tales) | GenED classroom | Individual | Academic engagement

Problem behavior | Percentage of all nonoverlapping data (PAND); Pearson's phi (ϕ) | PAND = 91.23% ϕ = .82 PAND = 92.68% ϕ = .84 |
|---|---|---|---|---|---|---|---|
| Collins et al. (2018) | Behavior Bingo | Culturally relevant interdependent group contingency intervention with student choice and peer models | SPED self-contained classroom | Classwide | On-task behavior Off-task behavior Disruptive behavior | Reported visual analysis and phase means only | - |
| Cramer & Castro-Olivo (2016) | Culturally-adapted *Strong Teens* Curriculum | Adapted SEL curriculum with choice between delivery in English or Spanish and emphasis on cultural values | GenEd classroom | Small group | Student resiliency Internalizing symptoms | Dependent sample t-tests | t = 2.12 (0.048)* t = 0.82 (0.424) |
| Farrell et al. (2018) | Olweus Bullying Prevention | Curriculum implemented to develop and monitor bullying plans at individual, classwide, and schoolwide levels | GenEd classroom | Combination | Teachers' ratings of student aggression and victimization | Cohen's d | *Numerous significant values reported* |
| Featherston (2014) | Real Talk 4 Girls | Social problem-solving skills and prosocial behavior program that uses restorative circle process to create psychological safety | Separate room | Small group | Social aggression

Social problem-solving skills

Prosocial behaviors | MANCOVA | $F(1, 47) = 23.90, p < .001$ $F(1, 47) = 16.52, p < .001$ $F(1,47) = 4.43, p = .041$ |
Garcia-Dubon (2018)	Increased opportunities to respond (OTR)	OTR with Kahoot.com, a game-based learning platform, to supplement a culturally responsive teaching approach	SPED classroom	Classwide	Academic engaged time Disruptive behaviors	Reported visual analysis and phase means only	-
Graves & Aston (2018)	*Brothers of Ujima*	Strength-based intervention for African American boys based on Afrocentric world view principles			Afrocentric values Racial identity - Centrality Racial identity – Private Regard Social resiliency	Dependent sample t-tests and effect size estimates	t = 7.77 (0.76)** t = 3.40 (0.40)* t = 2.65 (0.16)* t = 1.68 (0.46)*
Graves et al. (2017)	Strong Start	Adapted SEL curriculum lessons to reflect student interests, problems, community, language, and culture	Separate room	Small group	Social-regulation Self-competence Empathy Responsibility Externalizing behaviors	Cohen's d	d = .99 d = 1.38 d = .47 d = .10 d = .36

Kleinman & Saigh (2011)	Good Behavior Game	Team-based group contingency intervention with access to reinforcement informed by preference questionnaire	GenEd classroom	Classwide	Verbal disruption/aggression	Reported visual analysis, means only	-
Kowalewicz & Coffee (2014)	Mystery Motivator	Group contingency using a calendar to randomly signify reinforcement with a "M"	GenEd classroom	Classwide	Disruptive behavior	Reported visual analysis, means only	-
Lambert (2015)	Tootling	Students recorded instances of prosocial peer behaviors	GenEd classroom	Classwide	Disruptive behavior Appropriate behaviors	Nonoverlap of All Pairs (NAP)	NAP = .875 – 1.0 NAP = .8958 – 1.0
Long et al. (2018)	Good Behavior Game (GBG) / Mindfulness Skills Training	Mindfulness practice to develop awareness and responsivity using the STOP strategy and GBG	Separate room	Small group	Internalizing behaviors Externalizing behaviors Student well-being	Pearson's r	$r = .57**$ $r = .67**$ $r = .31**$
Lopach (2017)	Electronic Daily Hone-School Note	Electronic home-school note to communicate student behaviors and reward day	Separate room	Small group	On-task behavior	Tau-U, IRD	Tau-U = .90 IRD = .88
Mitchell (2014)	Good Behavior Game	Team-based group contingency intervention with access to reinforcement.	GenEd classroom	Classwide	Disruptive behaviors	PND	PND = 86.9 - 96.7
Mohn (2019)	Check-In/Check-Out	Daily progress report (with adult check ins and home-school communication	GenEd classroom	Individual	Percentage of total DPR points earned for respectful, responsible, safe, and kind behavior	TauU	1.0 with 95% CI [0.692, >1]
Ortiz et al. (2017)	Good Behavior Game	Team-based group contingency with access to reinforcement	Separate room	Small group	Frequency of out-of-seat behaviors Frequency of interrupting behaviors	NAP, TauU	Student 1 NAP=0.37 Student 1 TauU=0.27 Student 2 NAP=0.57 Student 2 TauU=0.13 Student 1 NAP=0.73 Student 1 TauU=0.47 Student 2 NAP=0.83 Student 2 TauU=0.67
Robinson-Ervin (2012)	Culturally responsive social skills program	Pre-recorded, computer-based, adapted social skills program	SPED self-contained	Individual	Following adult directions, participating, and entering conversations appropriately	-	-
Ryan et al. (2016)	*Strong Kids*	Skill-based social and emotional learning program	Separate room	Small group	Self-regulation Self-competence	Hedges *g*	$g = .26$ $g = .35$

		developed to promote resiliency			Empathy Responsibility		$g = .25$ $g = .32$
Simonsen et al. (2011)	Check-in/Check-Out	Daily progress report (DPR) with adult check ins and home-school communication	Separate room	Individual	Student off-task behavior Problem behaviors Social skills Academic competence	Cohen's d	$d = -0.90$ $d = 0.65$ $d = -0.12$ $d = -0.17$
Skerbetz & Kostewicz (2013)	Choice vs. No choice	Students were given the option to choose assignments	GenEd classroom	Classwide	Academic engagement, assignment accuracy, time to completion for assignments	Reported visual analysis and phase means only	-
Speight (2018)	Class wide Function-related Intervention Teams (CW-FIT)	Group contingency game; student teams were awarded points for displaying target behaviors and rewarded	Inclusion Classroom	Classwide	On-task behavior	TauU	TauU = 1 TauU = 1 TauU = 0.95
Stoll-Juredine (2017)	*Goodwill Girls*	A tertiary prevention program developed for students at-risk for relational aggression and peer victimization	-	Group	Relational Aggression Peers Belief Inventory	-	-
Tanol et al. (2010)	Good Behavior Game	Team-based group contingency intervention with reinforcement	GenEd classroom	Classwide	Student rule following, rule violations; teacher praise, response to rule violations	-	-
Usera (2017)	*Lakota Circles of Hope* (LCH)	Prevention program for youth in Grades 2-5. Targets healthy decisions in the context of Lakota traditions and values.	GenEd classroom	Classwide	Respect Lakota identity Risk behaviors Communication Self-esteem Conflict resolution	MANOVA	$F(1, 2675) = 7.56$, p <0.010 $F(1,2675) = 2.48$, ns $F(1,2675) = 0.98$, ns $F(1,2675) = 5.75$, p <0.020 $F(1,2675) = 7.93$, p <0.001 $F(1,2675) = 0.03$, ns
Willenbrink (2019)	*Growing Minds*	Curriculum in which students and teachers practice mindfulness together	GenEd classroom	Classwide	Teacher report of social skills Child report of problem behaviors Teacher report of emotion regulation Child report of mindfulness	MANCOVA, ANCOVA	$F(4, 164)=4.71$, $p<0.01$ $F(2, 168)=0.85$, ns $F(1, 170)=0.96$, *ns* $F(1, 170)=0.58$, ns

Note. **p < .01, *p < .05; [1]In BF2014, the primary treatment under investigation was the 5-week curriculum plus 7 additional weekly sessions. We report effect sizes based on pre- and post-intervention scores for this group. There was also a treatment group for which only a 5-week intervention was provided and results were similar (see Black & Fernando (2014), indicating the additional sessions may not be necessary.

Table 3

Treatment Integrity and Treatment Intensity Characteristics

Authors	Treatment integrity assessed?	Measurement Method	Percentage (%) of Sessions Assessed	Average Treatment Integrity (%)	Dimension Assessed	IOA (% of Sessions)	Session Length	Session Frequency	Total Treatment Duration
			Treatment Integrity				**Treatment Intensity**		
Aston & Graves (2016)	No	-	0	-	-	-	35 min	2x/week	8 weeks
Aston et al. (2017)	Yes	DO – checklist	100	100	Adherence	-	35 min	1x/week	8 weeks
Barrasso (2020)	Yes	DO – checklist	100	94 – 99%	Adherence	20	15 min	2x/week	3 - 5 weeks
Black & Fernando (2014)	No	-	0	-	-	-	15 min	3x/week (first 5 weeks); 1x week (next 7 weeks)	12 weeks
Bluth et al. (2016)	Yes	Method not reported	100	-	-	-	50 min	1x/week	11 weeks
Bunch-Crump (2016)	Yes	DO – checklist	51% (CICO) 20% (IC)	100% (CICO) 92% (IC)	Adherence	-	5 min (CICO) 30 min (IC)	2x/day (CICO) 1x/day (IC)	10 – 31 sessions (CICO); 5 sessions (IC)
Castro-Olivo & Merrell (2012)	Yes	DO – checklist	25	-	Adherence	-	-	-	10 – 12 weeks
Castro-Olivo et al. (2018)	Yes	DO – checklist	-	-	Adherence	-	-	-	2 – 4 weeks
Collins et al. (2018)	Yes	DO – checklist	100	100	Adherence	-	40 min	-	12 – 18 sessions
Cramer & Castro-Olivo (2016)	No	-	0	-	-	-	60 min	2x/week	6 weeks
Farrell et al. (2018)	Yes	DO – rating scale	20	-	Adherence, Quality	-	-	-	3-4 school years

Study									
Featherston (2014)	No	-	0	-	-	-	60 min	2x/week	3 weeks
Garcia-Dubon (2018)	Yes	DO – checklist	50	100	Adherence	15	10 min	3x/week	7 weeks
Graves & Aston (2018)	No	-	0	-	-	-	-	1x/week	12 weeks
Graves et al. (2017)	Yes	DO – checklist	20	100	Adherence	-	-	1x/week	1 year
Kleinman & Saigh (2011)	No	-	0	-	-	-	30-60 min	5x/week	6 weeks
Kowalewicz & Coffee (2014)	Yes	DO – checklist	100	100	Adherence	28%	40 min	5x/week	8 weeks
Lambert (2015)	Yes	DO –checklist; SR – checklist	100	97	Adherence	37%	2 hours	5x/week	-
Long et al. (2018)	Yes	SR – checklist	-	-	Adherence	-	90 min	2x/week	4.5 weeks
Lopach (2017)	-	-	-	-	-	-	30-40 min	2x/week	5 weeks
Mitchell (2014)	Yes	DO – checklist	100	92, 90, 88	Adherence	25%	-	-	-
Mohn (2019)	Yes	SR – checklist	-	-	Adherence	-	-	10x/day	6 weeks
Ortiz et al. (2017)	Yes	SR – checklist	100	100	Adherence	-	20 min	-	-
Robinson-Ervin (2012)	Yes	DO – checklist	30	-	Adherence	-	30 min	3-4x/week	3-7 weeks
Ryan et al. (2016)	Yes	DO – checklist	10	100	Adherence, Exposure	-	-	1x/week	14 weeks
Simonsen et al. (2011)	Yes	DO – checklist	-	93	Adherence, Exposure	-	-	2x/day	6 weeks
Skerbetz & Kostewicz (2013)	Yes	DO –checklist; SR – checklist	100	100	Adherence	55%	-	4x/week	-

Speight (2018)	Yes	DO –checklist, SR – checklist	100	-	Adherence	30%	45 min / 90 min	5x/week	6 weeks
Stoll-Juredine (2017)	No	-	-	-	-	-	30-40 min	2x/week	5 weeks
Tanol et al. (2010)	Yes	DO – checklist	100	90.6 - 98.0	Adherence, Quality	25%	10 min	5x/week	8 weeks
Usera (2017)	Yes	DO –checklist, SR – checklist, SR – interview	100	-	-	10%	45 min	10x/year	3 years
Willenbrink (2019)	Yes	DO – checklist	25	82.5	Adherence	-	20 min	2x/week	10 weeks

Note. For measurement method, DO = direct observation, SR = self-report. IOA = interobserver agreement. CICO = Check in/Check Out. IC = I-Connect.