

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Special Issue: *Mentoring: Theoretical Background, Empirical Findings, and Practical Applications*

ORIGINAL ARTICLE

Pilot randomized trial of brief school-based mentoring for middle school students with elevated disruptive behavior

Samuel D. McQuillin¹ and Heather L. McDaniel²¹Department of Psychology, University of South Carolina, Columbia, South Carolina. ²The Curry School of Education, University of Virginia, Charlottesville, Virginia

Address for correspondence: Samuel D. McQuillin, Department of Psychology, University of South Carolina, 1512 Pendleton St., Columbia, SC 29208. MCQUILLIS@mailbox.sc.edu

We report a pilot study of a brief (i.e., 10 sessions) goal-focused mentoring program for middle school students with elevated disruptive behavior. Students with high levels of school misconduct ($n = 67$) were randomly assigned to the program or school as usual. We collected multi-informant emotional, behavioral, and academic functioning assessments pre- and postintervention. Results indicate that the program produced significant positive changes in school behavioral infractions, math grades, students' report of emotional symptoms, and school problems. Mentors found the curriculum acceptable, understood the material from the manual, found implementing the curriculum feasible, and felt the program was congruent with the school mission. Mentors of less impaired students indicated that they desired additional support implementing the manual, perhaps indicating the program is overengineered for students with higher levels of impairment. The manuscript concludes with a discussion of considerations for future research and implementation, including the importance of integrating mentoring programs into existing support systems (e.g., multitiered systems of support) within the school context.

Keywords: mentoring; school-based mentoring; at-risk youth; disruptive behavior

Introduction

Currently, in the United States, between 2 and 2.5 million youth are served in mentoring programs each year,¹ which is an increase of around 2 million since 1990. With this expansion, the characteristics of children being served by these programs have shifted, with programs serving children who have increasingly greater needs.^{2,3} It is estimated that in formal mentoring programs, approximately 55% of youth served are academically at-risk, 20% have mental health needs, and 14% are adjudicated or juvenile justice-involved.² This change in the risk status of young people in mentoring programs also corresponds with changes in the priorities of programs. Garringer and colleagues² note that 20 years ago, 100% of programs rated a top priority as providing a caring adult relationship, while in 2016, only 44% rated this as within their top four priorities. Current priorities include foci such as mental

health promotion, violence prevention, and life and social skills development.

The shift in mentoring program focus and the characteristics of children being served also corresponds with a shift in setting, with more mentoring organizations inserting programs in schools.² School-based mentoring is reported to be the fastest growing form of mentoring. This increasing preference for school-based programs may be guided by research indicating that school-based programs are cheaper⁴ and equally effective as community-based programs.⁵ Schools also provide safe centralized locations for mentoring activities, increased access to youth from under-resourced backgrounds, reduced program liability for transporting youth, and a greater presence of other qualified helping adults within the same vicinity as the mentoring activities.

On average, the effect sizes of school-based mentoring are modest. In meta-analyses, researchers

find the results of school-based mentoring to range between $g = 0.09^{6,7}$ and $g = 0.27^{5,7}$. However, there is considerable heterogeneity in the effect sizes on different outcomes within studies and on the same outcomes between studies.^{5,6} For example, within mentoring studies, it has been observed that survey-based assessments are associated with larger effect sizes than other types of measures, such as archival school record data.⁵ Between studies, larger effects have been observed when a higher proportion of male mentees and male mentors participate as well as when mentors have a background in a helping profession.⁵ Moreover, within existing evaluations of school-based mentoring, very little is known about what mentors actually do when they are with their mentees, which makes it difficult to understand why, how, and for whom mentoring may be the most helpful. In a review of how researchers specified treatments in school-based mentoring, McQuillin *et al.*⁸ found that only 23% of evaluations prescribed practices and less than 10% measured what mentors did when they were with mentees. This is particularly problematic, as recent research suggests that student outcomes are maximized when mentors develop a close relationship with their mentees (i.e., developmental mentoring), as well as engage in skill development and goal-directed activities (i.e., instrumental mentoring).⁹ Thus, the existing cumulative literature on school-based mentoring is practically and scientifically limited in informing how mentors might best help young people. Additional research is needed to operationalize effective instrumental strategies for use in effective school-based mentoring. Similarly, most randomized evaluations of school-based mentoring with volunteer mentors do not focus on young people with high levels of impairment. We posit that the expansion of school-based mentoring programs for children with greater needs has outpaced the evidence base, and more research is necessary to understand the safety, effectiveness, and best practices for extending mentoring services to children with high levels of school impairment. To bridge this gap, the purpose of our current study was to pilot a randomized trial of a youth school-based mentoring program designed to support students demonstrating elevated disruptive behavior in middle school.

Limitations of existing research on mentoring youth with disruptive behavior

Research focused on mentoring youth with elevated disruptive behavior includes mixed and sometimes contradictory results. Some researchers have found that children with elevated environmental stressors (e.g., socioeconomic stress or absent parent) and behavioral risk (e.g., poor academic performance, misconduct, or substance use) show fewer benefits from mentoring, have shorter mentoring relationships, and are less satisfied with their mentoring relationships than those at lower risk.^{10,11} Furthermore, a meta-analytic review of mentoring programs suggests that there was appreciable variability in outcomes for youth presenting with behavioral or other personal difficulties.¹¹ However, a more recent review suggests that mentoring programs are associated with more favorable outcomes when youth demonstrate elevated levels of disruptive behavior.¹² Of 27 studies identified that included youth involved in problem behavior, the estimated effect size of mentoring was 0.28, while the effect of mentoring in the 55 studies, not including youth with disruptive behavior, reported a smaller effect size of 0.19.¹² The authors of this later review suggest that the observed improvement in outcomes for these youth in more recent years may be related to the implementation of more goal-directed mentoring programs in schools.¹² Another recent systematic review focused solely on mentoring youth at risk for or currently demonstrating delinquent behaviors echoes these positive findings, indicating that mentoring programs resulted in significant, small-to-moderate reductions in delinquency and increases in academic functioning, although programs notably lacked specificity around prescribed practices and theoretical mechanisms.¹³ While these more recent reviews of the impact of mentoring for youth demonstrating disruptive behaviors indicate some promise, evidence from experiments that use volunteer mentors to help young people with disruptive behavior in the school setting is inconclusive (e.g., Refs. 14 and 15), and there is no known extant literature on brief, instrumental, school-based mentoring programs for students exhibiting disruptive behavior. More research is needed focusing on if

and what type of school-based mentoring is an appropriate service for students who demonstrate elevated disruptive behavior difficulties.

Purpose of the current study

In the current study, we piloted a randomized controlled trial of brief, instrumental mentoring for middle school-aged adolescents who displayed the highest levels of school-reported disciplinary infractions. The purpose of this study was to estimate treatment effects and variance components to inform power analyses for the future evaluations of similar interventions and provide preliminary support for the safety of the intervention in a supervised and controlled setting. Additionally, the pilot served three functions: (1) testing assessment demands on teachers, students, and parents; (2) understanding mentors' perceptions of acceptability, feasibility, and usability of the curriculum; and (3) assessing practical barriers associated with implementation procedures and dosage.

First, we evaluated completion rates of a multi-informant assessment schedule for school functioning, as well as emotional and behavioral functioning. Previous evaluations of similar mentoring interventions have relied mostly on student self-reported data or school records, and have not incorporated formal, psychometrically valid measures of school behavioral and mental health, or behavioral and mental health reports from parents. To this end, we collected the Behavior Assessment System for Children (BASCTM) 3rd edition¹⁶ from parents, teachers, and students, and the Classroom Performance Survey¹⁷ (CPS) measures from teachers. *A priori* goals of the study were to collect 90% of the BASCs from each informant pre- and postexperiment. Second, we assessed mentors' perceptions of the program, the usability of the manual, and their reaction to mentoring this population of children. We were also interested in mentors' opinions of the training and treatment manual. In addition to testing the feasibility of the study procedures, we report our experience coordinating this program and providing onsite technical assistance to mentors with an emphasis on identifying barriers, facilitators, and demands of conducting this randomized trial. We provide suggestions for improving the coordination of the program in future evaluations of similar interventions.

Method

Overview

We recruited volunteer undergraduate students from a large public university to serve as mentors. Mentees in the program were middle school students with elevated school disciplinary referrals from a public middle school in a southeastern state. To pilot the randomized trial and estimate treatment effects, students were randomly assigned to treatment or "school as usual." Pre- and postoutcomes included school records (i.e., grades and disciplinary infractions), teacher reports of classroom performance, and teacher-, parent-, and self-reports of children's emotional and behavioral functioning. Students in the intervention group received approximately 10 45-min mentoring sessions during the school day over the course of the 18-week semester.

Participants

School. The participating middle school was a public middle school in the southeastern United States. In the 2016–2017 school year, the school enrolled approximately 900 students with approximately 55% of the student population being Black, 34% White, and 11% of other racial or ethnic backgrounds. The school had a 96% attendance rate, 23% suspension rate, and 42% of students received a free or reduced-price lunch.

Students. Middle school students and their caregivers were recruited in the fall of the 2016 school year. Potential mentees received consent forms to take home to their caregivers for completion, and parents received phone calls from the community coordinator describing the study. Students were considered eligible for participation if they presented within the top 10% of behavior infractions in the school (i.e., defined by the number of office discipline referrals). We invited approximately 90 participants using this procedure, and 57 consented to participate in the study. However, males were overrepresented in this sample. The Community Coordinator, a school support personnel, then assisted with the identification of females at risk to present within the top 10% of behavior problems in the school. This resulted in the enrollment of an additional 11 female middle school students into the project. Participating students were in the 6th (28%), 7th (49%), and 8th (23%) grades. The participating students primarily identified as male

Table 1. Demographic characteristics of treatment and control

	Control	Treatment
Total sample	33	34
6th Grade	10	9
7th Grade	16	17
8th Grade	7	8
Male	20	24
Black	28	29
White	5	4
Latin/Hispanic	0	1

(69%) and Black/African American (88%). Table 1 displays demographic characteristics of students by the assigned condition.

Mentors. Mentors were recruited in the fall semester of the 2016–2017 academic year. The targeted mentor population was all college undergraduates in a University honors program and/or students in the College of Arts and Sciences. The authors attended undergraduate courses for the honors program and in the College of Arts and Sciences to share information about the volunteer mentoring opportunity. Prospective mentors completed an online application and in-person interview with the program director. While recruitment efforts were intended to recruit both male and female mentors, all 34 participating mentors were female undergraduate students and 67.6% of the mentors were psychology majors. Furthermore, 82% of mentors were White, 6% Black/African American, 3% Latin American/Hispanic, 3% Asian, and 6% did not report on race/ethnicity. All mentors also completed a background check and program training before meeting with their mentees. Mentors were recruited as volunteers and did not receive any compensation for their participation.

Measures

BASC. The BASC is a comprehensive, validated, multi-informant system for assessing the behavior and emotions of youth.¹⁶ The student’s perspective of their own behaviors and emotions was measured by the Self-Report of Personality (SRP). The teacher and parent perspectives of student behavior were measured by the Teacher Rating Scales (TRS) and the Parent Rating Scales, respectively. The assessments were collected from each reporter pre- and postintervention. The teacher rater for each student

was randomly selected from each of the student’s core classes (i.e., Math, English Language Arts, or Science).

Teacher CPS. This 18-item rating scale assessed a student’s behavior in the classroom in the past month, as reported by the teacher.¹⁷ The teacher that was randomly selected to complete the BASC for a given student also completed the CPS.¹⁷ Fifteen Likert-style items, such as “Brings necessary items to class” and “Turns in completed work,” were rated on a 5-point scale ranging from “always” to “never” and yielded two sum scores representing academic competence and interpersonal competence. Lower scores on these subscales are indicative of greater competence. Previous research suggests that the CPS demonstrates adequate psychometric properties.

School records. Student school record data were collected to serve as an ecologically meaningful measure of school performance.¹⁸ Student grades in core subjects and a count of behavior referrals were collected from school administrative records. Second-quarter school records were used as the preassessment time point as these were finalized prior to the intervention. Fourth-quarter records were used as the postassessment time point, as these reports were released when approximately 95% of mentoring dyads completed the full cycle of the intervention.

Grades in core subjects. Grades in the core subjects of Math, Science, and English Language Arts were recorded as percentages (i.e., with a valid range from 0% to 100%) for the second (i.e., preintervention assessment) and fourth (i.e., postintervention assessment) academic quarters.

Behavior referrals. A count of behavior referrals was used as an indicator of disruptive behavior exhibited in the school setting.¹⁹ These data were released from school records and initially recorded by the school administration. Behavioral referrals in this study represent conduct infractions that require school disciplinary action. These offenses include refusal to obey, excessive tardiness or class cutting, disruption of class, use of profanity, unlawful activities, threats of violence, theft, disruption of a school bus, sexual harassment, fighting, property destruction, possession of gang-related materials, gang-related activity, and violation of the district’s technology agreement. Depending on the offense,

the frequency of the offense, and the severity of the offense, the school used these indicators to implement the following discipline procedures: in-school conference with parent, lunch and after-school detentions, in-school suspension, and out-of-school suspension between 1 and 10 days. The records released for this study did not include information on the specific infraction or the resulting disciplinary action, and, as such, can only be understood as a frequency indicator of discipline problems (i.e., it does not include information on the severity).

Student's life satisfaction scale. This 7-item measure assessed students' self-reported levels of life satisfaction over the past several weeks.²⁰ Students rated items, such as "My life is going well" and "I have what I want in life" on a 6-point, Likert-style scale ranging from "strongly disagree" to "strongly agree." We used a mean score of this scale in our analyses, with higher scores indicative of higher life satisfaction. In the previous research, this scale has demonstrated adequate psychometric properties.^{20,21}

Usage rating profile-intervention revised (URP-IR). A revised version of this form was used to evaluate acceptability, feasibility, and usability of the mentoring intervention as reported by the mentors.²² The items were used to reflect wording more aligned with the current project (e.g., "program" replaced "intervention"). Mentors rated 29 items, such as "This program is an effective way to help students succeed in school" and "I understand how to use the program's manual," on a 6-point scale from "strongly disagree" to "strongly agree." Items were summed to create six subscale scores representing acceptability, understanding, home-school collaboration, feasibility, system climate, and system support. The URP-IR is an established measure of treatment acceptability and related factors.²³

Analytic plan

Descriptive statistics. To provide information on the feasibility of the assessment schedule and the usability of the mentoring curriculum, we provide basic descriptive statistics on pre- and postassessment and record completion rates. Additionally, we provide raw percentage data on the items that make up the six constructs of a revised version of the URP-IR.

Inferential models. All statistical models were conducted in the statistical analysis package R.²⁴ To evaluate the effect of the mentoring program, we included an intent-to-treat (ITT) dummy code in each model. All models focused on student-level outcomes at postassessment, given the student-randomized research design. In addition to the ITT dummy code, covariates for all models included student gender, student grade, and baseline assessment on the outcome being modeled. The four student-reported outcomes were modeled in a general linear model with ordinary least squares estimation (i.e., life satisfaction, SRP BASC school problems, SRP BASC emotional symptoms index, and the SRP BASC inattention/hyperactivity scale). The four teacher-reported outcomes, as well as the four grade outcomes, were modeled as linear mixed models fit by restricted maximum likelihood (i.e., TRS BASC internalizing and externalizing problems, as well as academic competence and interpersonal competence). A random error term was included on the intercept of teacher-reported outcomes to account for the clustering of students within teachers. In addition to the ITT dummy code and covariates described above that were included in all models (i.e., student's gender and grade, and baseline assessment on the outcome being modeled), teacher report of baseline academic competence and interpersonal competence were included as covariates in all teacher-reported outcome models. For the student- and teacher-reported outcomes, as well as student grades at postassessment, a standardized estimate of treatment effect size, Cohen's *d* was calculated by dividing the treatment parameter by the pooled standard deviation of the outcome from a model without covariates.⁷ Finally, we used a zero-inflated Poisson regression model to analyze behavioral count data. This model included all covariates discussed above, as well as the teacher report of academic and interpersonal competence.

Procedures

Mentoring program. The mentoring program is a brief (i.e., 10 sessions), goal-focused, one-on-one school-based mentoring program that occurs during nonacademic electives.²⁵ Mentors in the program are trained in principles and practices of motivational interviewing (MI),²⁶ a person-centered approach to having conversations about behavior change. Specifically, the program was adapted

from a single-session MI intervention, the Student Checkup.²⁷ Mentors use the process of MI to develop a positive working alliance with mentees and guide mentees toward school-related goals. Although the process of MI is used throughout the program, mentors use a manual that guides program activities for each session. The first three sessions of the manual are structured to facilitate rapport-building activities, helping the mentor identify and understand student strengths and values, and a structured activity adapted from Homework Organization and Planning Skills.²⁸ Beginning in the third session, mentors help guide the mentee to identify a school-related goal to work toward using the aforementioned Student-Checkup protocol.²⁵ Following the third session, mentors and mentees review progress toward weekly goals, adjust plans of action, and learn new skills related to identified goals. To assist in skill development, the manual includes modular, single-session interventions adapted from cognitive behavioral therapy (e.g., understanding how activating events, beliefs, and consequences are connected; setting goals for positive activities; and progressive muscle relaxation), academic intervention (e.g., expository reading strategies), and modules used to help facilitate MI conversations.

The theoretical frame of the mentoring program is consistent with Cavell and colleagues' commentary on the role of socialization in the development and mitigation of antisocial and disruptive behavior.²⁹ Specifically, the mentoring program is designed to engage young people at-risk for delinquency or antisocial behavior by providing a prosocial context that might be limited or inaccessible otherwise. For example, youth with chronic disruptive behavior in school have frequent negative social interactions with adult helpers at school, which builds mutual distrust between the students and teachers, and lowers the students' outcome expectations of engaging in prosocial contexts. As Cavell and colleagues note, "Social disadvantage and prejudice conspire with mutual distrust and fear to keep many youth away from settings and opportunities that could enrich their developmental experiences" (pp. 55). In the mentoring program, mentors use MI to provide an intentional approach to collaborating around goals that the child sets while also offering, through a menu of options, the opportunity to develop skills derived from cognitive and

behavioral sciences. Consistent with Cavell and colleagues' commentary, this type of mentoring "could directly enhance the degree to which targeted children achieve transactional success over the course of the intervention period and, in turn, develop more positive outcome expectancies about their ability to function in prosocial contexts" (pp. 55), like school.

Mentor training. All mentors completed a series of trainings, including two 1.5-h in-person trainings in the fall semester and one 1-h on-site training in the spring semester. The first 90-min in-person training was a didactic overview of MI and other key program components. The second 90-min training was a role-play and rehearsal session in which prospective mentors practiced key MI skills introduced at the first training (i.e., open-ended questions, affirmations, reflections, and summaries). The prospective mentors practiced the use of these skills in rotating triads; one mentor practiced the skills, one mentor role played a protégé, and the third mentor recorded the successful use of the skills. Records of key skills included asking open-ended questions instead of closed-ended questions, making approximately two reflections for each question, making authentic affirmations, and summarizing the conversation.

Throughout the course of the mentoring relationship, mentors were also required to complete five "just-in-time" online trainings prior to specific sessions of the program (i.e., approximately weekly for sessions one through five). The just-in-time trainings involved dialogue examples that juxtapose verbal behavior consistent and inconsistent with MI, procedural steps for the upcoming session, and how to handle some common issues within the upcoming session (e.g., "If your mentee doesn't want to set a goal"). Each just-in-time training was less than 5 minutes. Examples of content include (1) asking open-ended questions that evoke change talk; (2) affirming effort toward goals and inherent worth; (3) reflections that emphasize change talk and soften sustain talk; and (4) avoiding the righting reflex (e.g., arguing, telling war stories, warning, etc.). Supervisors checked for mentor completion and understanding of online trainings.

Assignment to conditions. Students were randomly assigned using the R statistical computing program following the informed consent and

pretest measures. Students assigned to the mentoring program were matched with mentors based on similar interests collected from both mentors and mentees by the school social worker.

Implementation structure. Mentors were supervised by the authors on site at the school. Fidelity was monitored by the supervisors and with the assistance of a school-based social worker, and by two advanced undergraduate students in psychology (i.e., a total estimated 1.5 full-time equivalent (FTE)). To encourage implementation of the manual, we used a proactive support approach.³⁰ At check-in, the mentor briefly met with a supervisor to check the knowledge of session content and allow time for the mentor to ask any clarifying questions. At check-out, the mentor verbally reported to the supervisor about coverage of all session content in the manual and goals that the student had set with their mentor that week. Fidelity was measured using mentor self-reported completion of the session.

Results

Assessment completion rates

Importantly, although the original design of the study included parent-reported outcomes, we encountered significant barriers to retrieving parent reports, with between 40% and 53% of parent measures being incomplete at the pre- and posttest, respectively. Because of this level of missingness, we did not analyze parent-reported data. The maximum number of missing school records for any outcome at the pretest was three records (i.e., 4%). These records were absent because the students were reenrolled with a different teacher between the first and second quarters ($n = 2$) or the student was a transfer student from another school ($n = 1$). The maximum number of missing teacher-reported outcomes (i.e., CPS and BASC-TRP) was five (i.e., 7%). Similarly, the maximum number of missing self-reported outcomes was five (i.e., 7%). Excepting parent measures, data collection rates exceeded our *a priori* goal (i.e., 90%).

Program fidelity data

Mentors self-reported fidelity data during check-out supervision. Across all mentors, the average number of self-reported completed sessions was 8.82, with a minimum of 3 ($n = 1$) and a maximum of 10 ($n = 15$). This number is a downwardly biased, as during a 2-week supervisor transition, the super-

visor did not track fidelity data for nine matches. Ignoring these nine matches, the average number of completed sessions was 9.2. In the circumstance that a match did not meet at the expected time, mentors were asked to complete the session missed prior to advancing in the manual. In the case that mentors would not likely be able to finish all 10 sessions, duplicate goal-setting sessions (i.e., identical procedures to set a new goal) were dropped.

Outcome statistics

Raw unadjusted means, standard deviations, and pre- to postassessment effect sizes for primary outcomes are displayed in Table 2 and the results of ITT analyses are presented in Tables 3 and 4. For student-reported outcomes, there was a statistically significant, medium-sized effect of the program on school problems ($b_{ITT} = -6.21(2.19)$, $t = -2.80$, $P < 0.01$, 95% CI: -10.61 to -1.81 , $d = -0.58$). Youth randomly assigned to the mentoring program reported significantly lower school problems at postassessment as compared with youth in the control group. Results also demonstrated a statistically significant, small-to-medium effect of the program on the Emotional Symptoms Index ($b_{ITT} = -4.11(2.03)$, $t = -2.02$, $P = 0.049$, 95% CI: -8.20 to -0.02 , $d = -0.34$). Random assignment to the treatment group, as compared with control, was associated with a 4.11-unit lower emotional symptoms index t -score at postassessment. While there were no statistically significant effects of the intervention on student-reported life satisfaction or inattention and hyperactivity, trends were in the hypothesized direction. As denoted in Table 2, youth in the mentoring program showed a larger mean increase in life satisfaction than youth in the control group from pre- to postassessment (i.e., $d_{\Delta\text{treatment}} = 0.26$ versus $d_{\Delta\text{control}} = 0.05$). Similarly, youth participating in the mentoring program showed a decline in their mean level of inattention and hyperactivity, while youth in the control group demonstrated a negligible increase from pre- to postassessment (i.e., $d_{\Delta\text{treatment}} = -0.19$ versus $d_{\Delta\text{control}} = 0.03$).

When examining the school record outcomes, we found that students in the treatment group had significantly fewer behavior-related office referrals than those in the control group at postassessment ($b_{ITT} = -0.74(0.34)$, $z = -2.35$, $P = 0.02$, 95% CI: -1.36 to -0.12). Random assignment

Table 2. Raw unadjusted means, standard deviations, change scores, and effect sizes of primary outcomes

Source	Tx M	Tx M	Tx post	Control	Control	Control	Tx Δd	Control
Primary outcomes	post	pre	SD	M post	M pre	post SD		Δd
School record								
Math	74.8	75.3	13.7	70.7	77.7	13.3	−0.04	−0.52
ELA	68.9	70.4	24.4	71.3	74.9	16.3	−0.06	−0.22
Science	72.2	73.5	22.1	64.3	75.8	26.6	−0.06	−0.43
History	79	70.3	12.5	80.8	72.8	14	0.69	0.57
Office referrals	0.82	10.8	1.11	0.97	9.61	1.67	−8.93	−5.16
Teacher report								
CPA	3.09	3.22	0.95	2.91	3.01	0.97	−0.14	−0.1
CPI	2.62	2.77	1.07	2.62	2.5	0.78	−0.14	0.15
BASC external	72.1	68.6	19.2	69.2	66.8	19.4	0.18	0.12
BASC internal	53.8	52.8	12	56.2	55.6	14.2	0.086	0.044
Self-report								
Life satisfaction	4.48	4.25	0.89	4.26	4.22	0.84	0.26	0.05
BASC school problems	54.4	57.6	0.98	58.4	55.1	11.3	−0.33	0.29
BASC-ESI	52.3	57.8	12.8	54.3	54.8	11.2	−0.42	−0.045
BASC-IH	55.5	57.7	10.9	54.9	54.6	10.4	−0.19	0.03

ELA, English and Language Arts grades; CPA, classroom performance survey academic functioning; CPI, classroom performance survey interpersonal functioning; BASC, behavioral assessment system for children; BASC external, BASC externalizing problems index; BASC internal, BASC internalizing problems index; BASC-ESI, BASC emotional symptoms index; BASC-IH, BASC inattention hyperactivity; Tx, treatment; M, mean; Tx Δd, Tx post – Tx pre; Control Δd, control post – control pre.

to the mentoring program was also associated with statistically significant, small-to-moderate gains in math grades ($b_{ITT} = 5.77(2.68)$, $t = 2.15$, $P = 0.04$, 95% CI: 0.26–11.31, $d = 0.42$). On average, students randomly assigned to the mentoring program, as compared with youth in the control group, demonstrated math grades 5.77 percentage points higher at postassessment. The intervention was not associated with a statistically significant impact on the other school grade outcomes. However, trends appear to be more favorable for youth in the treatment group than the control group. Science and ELA grade means for youth in both the treatment and control groups declined from pre- to postassessment, and the mean decline in grades for the youth in the mentoring group was smaller (i.e., raw science grades: $d_{\Delta treatment} = -0.06$ versus $d_{\Delta control} = -0.43$; raw ELA grades: $d_{\Delta treatment} = -0.06$ versus $d_{\Delta control} = -0.22$).

We did not observe any statistically significant impact of the mentoring program on teacher-reported outcomes. Though, the impact of the program on externalizing behavior approached statistical significance ($b_{ITT} = 4.30(2.69)$, $t = 1.60$, $P = 0.12$, 95% CI: −0.96 to 9.39, $d = 0.22$), suggesting that teachers were reporting higher levels of externalizing problems at postassessment for stu-

dents in the mentoring program as compared with students in the control group. Similarly, the examination of raw means for the treatment and control group students from pre- to postassessment suggests that students in both the mentoring and control groups had increases in internalizing difficulties from pre- to postassessment, but this was incrementally greater in magnitude for students in the treatment group (i.e., $d_{\Delta treatment} = 0.09$ versus $d_{\Delta control} = 0.04$). Raw means also suggest that, on average, teachers reported gains in interpersonal competence for students in the mentoring program and losses for students in the control group (i.e., $d_{\Delta treatment} = -0.14$ versus $d_{\Delta control} = 0.15$). And, on average, students in both the treatment and control groups showed similar gains in teacher-reported academic competence (i.e., $d_{\Delta treatment} = -0.14$ versus $d_{\Delta control} = -0.10$).

Usability ratings from mentors

Table 5 includes item-level percentages of mentor endorsement for each item of the URP-I. In general, the program was found to be acceptable, feasible, understood by mentors, and congruent with the system climate. Some items suggested greater variability in the mentor’s appraisal of feasibility. While there was not an explicit parent component of the program, mentors disagreed as to the importance

Table 3. Multilevel intent to treat outcomes

Outcome	Parameter (intercept, ITT)	Coefficient	SE	σ	τ_{00}	Cohen's <i>d</i>	95% CI	<i>t</i> ratio	<i>P</i>
Math									
Intercept		74.79	15.80	9.65	8.19	–	–	–	–
ITT		5.77	2.68	–	–	0.42	(0.26 to 11.31)	2.15	0.04*
ELA									
Intercept		59.77	24.47	18.82	4.01	–	–	–	–
ITT		0.75	5.19	–	–	0.04	(–9.47 to 11.18)	0.14	0.88
Science									
Intercept		69.30	31.70	21.04	5.86	–	–	–	–
ITT		7.97	5.75	–	–	0.32	(–3.11 to 18.70)	1.39	0.17
History									
Intercept		68.39	19.89	11.09	3.31	–	–	–	–
ITT		3.02	3.09	–	–	0.23	(–3.20 to 9.86)	0.97	0.33
CPA									
Intercept		0.54	0.49	0.67	0.37	–	–	–	–
ITT		0.003	0.19	–	–	0.003	(–0.36 to 0.46)	0.02	0.99
CPI									
Intercept		0.72	0.48	0.70	0.20	–	–	–	–
ITT		–0.12	0.19	–	–	–0.13	(–0.50 to 0.23)	–0.64	0.52
BASC external teacher									
Intercept		67.91	7.65	8.91	7.42	–	–	–	–
ITT		4.31	2.69	–	–	0.22	(–0.96 to 9.39)	1.60	0.12
BASC internal teacher									
Intercept		52.38	4.26	4.53	6.96	–	–	–	–
ITT		0.78	1.51	–	–	0.06	(–2.48 to 3.66)	0.52	0.61

σ , standard deviation of the residual; τ_{00} , standard deviation of the random effect of the intercept; **P* < 0.05.

of home–school collaboration in the mentoring program, with some endorsing home–school collaboration as essential to the success of the program and others suggesting it as not important. We suspect this variability was likely due to variability in access to parents, which was inconsistent, hence the low return of parent-reported outcomes. Importantly, we also found greater variability in mentors’ report of system support necessary to successfully implement the program, with some mentors feeling more training and support was necessary. Specifically, 29% of mentors endorsed some level of agreement with the statement, “I was in need of additional training in order to implement this program’s manual.” Similarly, 22% of mentors agreed to some level with the statement, “I needed additional resources to carry out this program’s manual.” Based on supervision records and experience, we hypothesized that mentors who were matched with mentees with greater impairment may have desired more training than those with less impairment. We tested this hypothesis in post

hoc analyses, whereby we predicted system support from students’ pretest teacher-reported BASC Behavioral Symptoms Index (BASC-BSI; i.e., global assessment of impairment) and students’ BASC self-report of Emotional Symptoms Index (BASC-ESI). Contrary to our hypothesis, we found that mentors who were matched with more impaired young people felt more comfortable with system support for the manual. This trend was statistically significant for teacher reports of impairment on the BASC-BSI, *B* = –0.02(0.01), *t* = –2.53, *P* = 0.017. Although not statistically significant, the trend was shared with student reports on the BASC-ESI, *B* = –0.02(0.01), *t* = 1.87, *P* = 0.07.

Discussion

Results suggest that students randomly assigned to participate in the mentoring program demonstrated significantly improved math grades and decreases in referrals to the office for behavioral challenges, as well as decreases in self-reported school problems and emotional symptoms, as compared

Table 4. Self-reported outcomes

Outcome						
Parameter (intercept, ITT)	Coefficient	SE	Cohen's <i>d</i>	95% CI	<i>t</i>	<i>P</i>
Life satisfaction						
Intercept	2.13	0.46	–	–	–	–
ITT	0.21	0.18	0.24	(–0.16 to 0.58)	1.15	0.26
BASC school problems						
Intercept	58.07	3.01	–	–	–	–
ITT	–6.21	2.19	–0.58	(–10.61 to –1.81)	–2.83	0.007*
BASC inattention/hyperactivity						
Intercept	55.08	2.57	–	–	–	–
ITT	–2.26	1.89	–0.21	(–6.05 to 1.54)	–1.19	0.24
BASC emotional symptoms index						
Intercept	56.03	2.77	–	–	–	–
ITT	–4.11	2.03	–0.34	(–8.20 to –0.02)	–2.02	0.049*

**P* < 0.05.

with students in the control group at postassessment. These results provide some preliminary evidence of promise that brief, instrumental, school-based mentoring may be an effective way to help young people with elevated disruptive behavior perform better in school (i.e., increased grades and decreased behavior referrals) and feel better (i.e., decreased self-report of school problems and emotional symptoms) after participation in this mentoring program. The results of this study align with previous research on this program, which suggests that participation in brief, instrumental, school-based mentoring is associated with increases in academic performance and a reduction in behavioral referrals.^{25,31} Notably, the effect size of the intervention on math grades in this study ($d = 0.42$) is commensurate with or exceeds effect sizes observed in previous trials of this program.^{25,31} This effect exceeds the estimated average effect of mentoring on academic achievement for youth at risk for delinquency (i.e., $d = 0.11$).¹³ However, previous trials have also recorded significant impacts of the intervention on other core subject grades as well as student self-report of life satisfaction that were not observed in the current study.

We believe our positive result on student misconduct has school discipline policy implications, particularly considering that this result corroborates the results of a previous evaluation of the same intervention with a less impaired population.³¹ Importantly, infractions recorded in this study were used by the school to make decisions for disciplinary

action, including exclusionary discipline. Use of exclusionary discipline practices in U.S. schools has progressively risen, in part due to “zero tolerance” discipline policies.³² In the 2015–2016 school year, approximately 2.7 million students received one or more out-of-school suspensions³³ and a corresponding loss of over 11 million days of instruction.³⁴ School suspensions, both in and out of school, are associated with poorer academic performance and increased risk for school dropout across empirical studies,³⁵ as well as increased involvement with the juvenile justice system (see Ref. 36). Although our data preclude this analysis, we hypothesize that by reducing disciplinary infractions, youth in the treatment group may have been exposed to more instruction, which may cascade to other long-term positive outcomes (e.g., grade promotion). While it is critical for schools to consider practices outside of exclusionary discipline, this research suggests that it is also important to consider the implementation of preventive interventions that may scaffold student learning of critical social, emotional, and academic skills that may protect them from exclusionary discipline. This work contributes to a growing body of literature indicating that school-based mentoring may be helpful in promoting more positive outcomes for youth with elevated disruptive behavior.^{9,13,37}

Additionally, some mentors indicated that they needed more support to implement the program. Contrary to our original hypothesis, we found an *inverse* relationship between mentee’s global level of

Table 5. Usage rating profile—school-based mentoring mentor report

Subscale/item	Percentage of mentors endorsing a response (N = 31)					
	Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
<i>Acceptability</i>						
This program is an effective way to help students succeed in school.	0	0	0	0	58	42
The program is an appropriate way to teach students skills for school success.	0	0	3	6	51	39
I am not interested in participating in this program.	74	26	0	0	0	0
I have a positive attitude about participating in this program.	0	0	0	0	71	29
This program is a good way to help students succeed.	0	0	0	0	45	55
I implemented this program's manual with a good deal of enthusiasm.	0	0	0	3	55	42
This program is not disruptive to other students.	0	0	3	10	42	45
I was committed to carrying out this program's manual.	0	0	3	10	35	52
This program's manual easily fits in with my current practices as a mentor.	0	0	3	13	42	42
<i>Understanding</i>						
I understood how to use the program's manual.	0	0	0	10	38	52
I am knowledgeable about the program's manual.	0	0	0	13	42	45
I understood the procedures laid out in this program's manual.	0	0	3	6	42	48
<i>Home-school collaboration</i>						
A positive home-school relationship is needed to implement this program.	3	0	6	23	39	29
Parental collaboration is required in order to implement this program.	3	19	13	32	13	19
Regular home-school communication is needed to implement this program.	3	3	26	19	36	13
<i>Feasibility</i>						
I was able to allocate my time to participate in this program.	0	0	0	3	23	74
The total time required to implement the program's manual is manageable.	0	0	6	10	29	55
Preparation of materials needed for this program is minimal.	3	3	3	19	52	19
Material resources needed for this program are reasonable.	0	0	0	9	39	52

Continued

Table 5. Continued

Subscale/item	Percentage of mentors endorsing a response (N = 31)					
	Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
This program's manual is too complex to carry out accurately.	39	39	9	13	0	0
The amount of time required for completing "check ins" and other forms is reasonable.	0	0	0	3	42	55
<i>System climate</i>						
My supervisor was supportive of my use of this program.	0	0	0	0	23	77
Use of this program is consistent with the mission of the school.	0	0	0	6	45	48
Implementation of this program's manual is well matched to what is expected in my role.	0	0	6	10	42	42
The program's procedures are consistent with the way things are done in this school system.	0	3	3	26	42	26
My daily life and responsibilities are conducive to participating in a program like this one.	0	0	3	6	48	42
<i>System support</i>						
I needed additional resources to carry out this program's manual.	10	48	19	16	6	0
I needed additional supervision to implement this program's manual.	16	48	19	16	0	0
I was in need of additional training in order to implement this program's manual.	10	52	10	13	13	3

behavioral difficulties (i.e., teacher-reported BASC-BSI) and mentor’s appraisal of system support, indicating that mentors matched with mentees who had fewer difficulties desired more support. Perhaps, this indicates that the program curriculum is overengineered toward students with higher levels of impairment and less relevant to students with lower levels of impairment. Or perhaps students with less impairment were more easily able to access the curriculum and needed less instruction and behavioral support, increasing the efficiency by which the mentor and mentee were able to complete activities and leaving the mentor with additional time for activities with their mentee. This might indicate a need to have a more adaptive curriculum that permits mentors to individualize the

program experience for youth with a wider range of impairment.

Considerations for future research and implementation

Based on our experience coordinating this study and the results of our analyses, we suggest three modifications to future trials of this or similar interventions. First, we propose greater integration between the mentoring program and existing tiered systems of support. These systems, such as positive behavior interventions and supports (PBIS³⁸) or expanded school mental health (SMH³⁹) services, are implemented within schools in increasing frequency and involve the coordination and support of various professionals who are native to the school. As coordinators, we

often found it challenging to help address the needs of this population without the support and involvement of other school professionals, including school counselors and psychologists, and social workers. This type of formal integration was not included in the current procedural manual or the coordinator procedures; however, it was at times necessary and integrated in an ad hoc manner when student needs were particularly acute. Although mentors received training to report acute symptoms to supervisors, future efforts should include training and procedures for engaging and interacting with SMH or PBIS teams to help coordinate care. Because of the needs of this population, it is unlikely that mentoring in isolation would be sufficient to robustly reduce impairment. We suspect that even our meager efforts to involve other school staff in helping these children improved the quality of the program.

Second, though mentors positively appraised the training and support procedures, we suspect that further refinement of the “just-in-time” online training modules that involve explicit instruction in managing disruptive behavior during sessions would improve the usability of the curriculum. Students in this iteration of the program demonstrated off-task and disruptive behaviors at a greater frequency than previous trials. Mentors indicated a need for training in more targeted skills to ameliorate these behaviors during their mentoring sessions.

Third, coordinating this study was a demanding task. Although our team consisted of approximately 1.5 FTE staff for support, we suggest that two FTE staff with backgrounds in SMH, PBIS, or other tiered frameworks for student social, emotional, behavioral, and academic supports would be most appropriate for coordinating the evaluation of this program with a similar sample, especially if a similarly intensive assessment strategy is planned.

Limitations

There are a variety of limitations in the current research that warrant consideration. First, this is a small, underpowered pilot study. In addition to being underpowered to detect a true effect, small samples are also associated with overestimation of effect sizes.⁴⁰ In addition, students for the current project were enrolled from one middle school in a southeastern state. It is possible that students

enrolled in the current project are not representative of students with increased behavioral difficulty in other school settings, limiting the generalizability of these effects. Thus, while some of the effects found in this study are promising, they are intended to inform power estimates for future, large-scale research that would be able to provide more precise estimates of population-level effects of the mentoring program. Similarly, our mentors were predominately White, females, and not representative of all possible undergraduate mentors. Moreover, there was a general mismatch between the mentors who were mentoring predominately Black male youth. This reflects the national snapshot of mentoring demographics with only about 24% of mentees identifying as White but more than half of mentors identifying as White, as well as approximately 60% identifying as female.² However, race and gender match warrant careful consideration. While this is not an uncommon scenario, additional research is needed to understand if and how this gender mismatch impacts student experiences of the mentoring relationship as well as outcomes.

Furthermore, coverage of session content was assessed by mentor self-report. Although self-report is one of the most common and feasible ways to assess implementation, the implementation science literature suggests that self-report measures may not correspond well with other measures of program implementation, such as more objective observation measures, and may not be as clearly linked to program outcomes.^{41,42} Future research should incorporate more objective measures of program implementation, such as observational ratings by a trained observer.

In addition, the aim of this study was to understand the feasibility of collecting parent-reported surveys of student functioning, as well as generating estimates of preliminary effects of the program on parent-reported outcomes. Unfortunately, parent survey completion rates were poor (i.e., 40% completion at baseline and 53% completion at postassessment), indicating that alternative methods to collect parent-report data are warranted in future trials. In addition, given the poor completion rates, we were unable to estimate effects of the intervention on parent-reported outcomes. In future research, it will be important to understand how the intervention impacts parent-reported

student outcomes to discern whether treatment gains are generalizing beyond the school setting.

Conclusions

Altogether, the results of this study suggest that this mentoring program may be a feasible strategy to implement with students in middle school with elevated disruptive behavior. Mentors generally reported the program was acceptable, feasible, understood, and congruent with the system climate. However, mentors who are working with youth with lower levels of disruptive behavior indicated that they could benefit from additional support. Furthermore, while additional, larger scale research is needed to garner more precise estimates of the intervention effects, the results of this pilot study suggest that, as compared with students in the control group, students receiving the school-based mentoring program had fewer behavioral referrals, school problems, emotional symptoms, and higher math grades at postassessment. Future research will aim to embed the intervention in existing multitiered support to ensure that these students are receiving the full continuum of tiered supports.

Acknowledgments

This work was funded by the Society for the Study of School Psychology through the Early Career Research Award mechanism.

Author contributions

S.D.M. wrote the grant for this project, pursued institutional review for the protection of human participants, ensured the integrity of data collection, conducted statistical analyses, and contributed to writing the manuscript. H.L.M. coordinated the study, wrote substantial portions of the manuscript, interpreted the statistical results, and provided methodological and statistical guidance throughout the course of the project. Both authors agree that our efforts on this manuscript are equally weighted and we share credit for the first authorship. We also thank Regina Schreiber who double-entered all data and prepared analytic datasets. Correspondence concerning the program's training, support, or manual should be addressed to S.D.M.

Competing interests

The authors declare no competing interests.

References

1. Raposa, E.B., N. Dietz & J.E. Rhodes. 2017. Trends in volunteer mentoring in the United States: analysis of a decade of census survey data. *Am. J. Community Psychol.* **59**: 3–14.
2. Garringer, M., S. McQuillin & H. McDaniel. Examining youth mentoring services across America: findings from the 2016 National Mentoring Program Survey. MENTOR: The National Mentoring Partnership.
3. Wheeler, M., T. Keller & D. DuBois. 2010. Review of three recent randomized trials of school-based mentoring: making sense of mixed findings. Society for Research in Child Development.
4. Alfonso, Y.N., S. Lindstrom Johnson, T. Cheng, *et al.* 2019. A marginal cost analysis of a Big Brothers Big Sisters of America youth mentoring program: new evidence using statistical analysis. *Child. Youth Serv. Rev.* **101**: 23–32.
5. Raposa, E.B., J. Rhodes, G.J.J.M. Stams, *et al.* 2019. The effects of youth mentoring programs: a meta-analysis of outcome studies. *J. Youth Adolesc.* **48**: 423–443.
6. Wood, S. & E. Mayo-Wilson. 2012. School-based mentoring for adolescents: a systematic review and meta-analysis. *Res. Soc. Work Pract.* **22**: 257–269.
7. Hedges, L.V. & I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. Elsevier.
8. McQuillin, S.D., M.D. Lyons, R.J. Clayton, *et al.* 2018. Assessing the impact of school-based mentoring: common problems and solutions associated with evaluating non-prescriptive youth development programs. *Appl. Dev. Sci.* <https://doi.org/10.1080/10888691.2018.1454837>.
9. Lyons, M.D., S.D. McQuillin & L.J. Henderson. 2019. Finding the sweet spot: investigating the effects of relationship closeness and instrumental activities in school-based mentoring. *Am. J. Community Psychol.* **63**: 88–98.
10. Raposa, E.B., J.E. Rhodes & C. Herrera. 2016. The impact of youth risk on mentoring relationship quality: do mentor characteristics matter? *Am. J. Community Psychol.* **57**: 320–329.
11. DuBois, D.L., B.E. Holloway, J.C. Valentine, *et al.* 2002. Effectiveness of mentoring programs for youth: a meta-analytic review. *Am. J. Community Psychol.* **30**: 157–197.
12. DuBois, D.L., N. Portillo, J.E. Rhodes, *et al.* 2011. How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychol. Sci. Public Interest* **12**: 57–91.
13. Tolan, P., D. Henry, M. Schoeny, *et al.* 2013. Mentoring interventions to affect juvenile delinquency and associated problems: a systematic review. *Campbell Syst. Rev.* **9**: 1–158.
14. Cavell, T.A. & J.N. Hughes. 2000. Secondary prevention as context for assessing change processes in aggressive children. *J. Sch. Psychol.* **38**: 199–235.
15. Hughes, J.N., T.A. Cavell, B.T. Meehan, *et al.* 2005. Adverse school context moderates the outcomes of selective interventions for aggressive children. *J. Consult. Clin. Psychol.* **73**: 731–736.
16. Reynolds, C.R. & R.W. Kamphaus. 2015. *Behavior Assessment System for Children*. 3rd ed. [Assessment instrument]. Bloomington, MN: Pearson.
17. Brady, C.E., S.W. Evans, K.S. Berlin, *et al.* 2012. Evaluating school impairment with adolescents using the classroom performance survey. *Sch. Psychol. Rev.* **41**: 429–446.

18. Sechrest, L., P. McKnight & K. McKnight. 1996. Calibration of measures for psychotherapy outcome studies. *Am. Psychol.* **51**: 1065–1071.
19. Pas, E.T., C.P. Bradshaw & M.M. Mitchell. 2011. Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychol. Sch.* **48**: 541–555.
20. Huebner, E.S. 1991. Initial development of the student's life satisfaction scale. *Sch. Psychol. Int.* **12**: 231–240.
21. Gilman, R. & S. Huebner. 2003. A review of life satisfaction research with children and adolescents. *Sch. Psychol. Q.* **18**: 192–205.
22. Chafouleas, S.M., A.M. Briesch, S. Neugebauer & T.C. Riley-Tillman. 2011. *Usage Rating Profile—Intervention, Revised [Assessment Instrument]*. Storrs, CT: University of Connecticut.
23. Briesch, A.M., S.M. Chafouleas, S.R. Neugebauer, *et al.* 2013. Assessing influences on intervention implementation: revision of the Usage Rating Profile-Intervention. *J. Sch. Psychol.* **51**: 81–96.
24. R Core Team. 2019. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
25. McQuillin, S.D. & M.D. Lyons. 2016. Brief instrumental school-based mentoring for middle school students: theory and impact. *Adv. Sch. Ment. Health Promot.* **9**: 73–89.
26. Miller, W.R. & G.S. Rose. 2009. Toward a theory of motivational interviewing. *Am. Psychol.* **64**: 527.
27. Strait, G., S. McQuillin & J. Anderson, *et al.* 2018. Practical applications of student focused MI. In *Motivational Interviewing: Children and Young People III "Education and Community Settings"*. E. McNamara, Ed.: 22–34. Merseyside, UK: Ainsdale.
28. Langberg, J.M., M.R. Dvorsky, S.J. Molitor, *et al.* 2018. Overcoming the research-to-practice gap: a randomized trial with two brief homework and organization interventions for students with ADHD as implemented by school mental health providers. *J. Consult. Clin. Psychol.* **86**: 39–55.
29. Cavell, T., S. Hymel, K. Malcom & A. Seay. 2007. Socialization and interventions for antisocial youth. In *Handbook of Socialization: Theory and Research*. J. Grusec & P. Hastings, Eds.: 42–67. New York: Guilford.
30. McQuillin, S.D., G.G. Straight & E. Saeki. 2015. Program support and value of training in mentors' satisfaction and anticipated continuation of school-based mentoring relationships. *Mentor. Tutoring Partnersh. Learn.* **23**: 133–148.
31. McQuillin, S., G. Strait, B. Smith, *et al.* 2015. Brief instrumental school-based mentoring for first- and second-year middle school students: a randomized evaluation. *J. Community Psychol.* **43**: 885–899.
32. Losen, D. & R. Skiba. 2010. Suspended education: urban middle schools in crisis. The Civil Rights Project at UCLA and the Southern Poverty Law Center.
33. United States Department of Education Office for Civil Rights. 2018. 2015–2016 Civil rights data collection: school climate and safety.
34. Losen, D.J. & A. Whitaker. 2018. 11 Million days lost: race, discipline and safety at U.S. public schools. The Center for Civil Rights Remedies of UCLA's Civil Rights Project and the American Civil Liberties Union of Southern California.
35. Noltmeyer, A.L., R.M. Ward & C. Mcloughlin. 2015. Relationship between school suspension and student outcomes: a meta-analysis. *Sch. Psychol. Rev.* **44**: 224–240.
36. Skiba, R.J., M.I. Arredondo & N.T. Williams. 2014. More than a metaphor: the contribution of exclusionary discipline to a school-to-prison pipeline. *Equity Excell. Educ.* **47**: 546–564.
37. Lyons, M.D. & S.D. McQuillin. 2019. Risks and rewards of school-based mentoring relationships: a reanalysis of the student mentoring program evaluation. *Sch. Psychol.* **34**: 76–85.
38. Sugai, G. & R.R. Horner. 2006. A promising approach for expanding and sustaining school-wide positive behavior support. *Sch. Psychol. Rev.* **35**: 15.
39. Weist, M.D. 1997. Expanded school mental health services. In *Advances in Clinical Child Psychology*. T.H. Ollendick & R.J. Prinz, Eds.: 319–352. Boston, MA: Springer US.
40. Ferguson, C.J. & M. Heene. 2012. A vast graveyard of undead theories: publication bias and psychological science's aversion to the null. *Perspect. Psychol. Sci.* **7**: 555–561.
41. Durlak, J. & E. DuPre. 2008. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* **41**: 327–350.
42. Schoenwald, S.K., A.F. Garland, J.E. Chapman, *et al.* 2011. Toward the effective and efficient measurement of implementation fidelity. *Adm. Policy Ment. Health Ment. Health Serv. Res.* **38**: 32–43.