School Psychology Review

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uspr20

In-School and Out-of-School Suspension: Behavioral and Psychological Outcomes in a Predominately Black Sample of Middle School Students

Daniel R. Cohen, Crystal Lewis, Colleen L. Eddy, Lauren Henry, Caroline Hodgson, Francis L. Huang, Wendy M. Reinke & Keith C. Herman

To cite this article: Daniel R. Cohen, Crystal Lewis, Colleen L. Eddy, Lauren Henry, Caroline Hodgson, Francis L. Huang, Wendy M. Reinke & Keith C. Herman (2023) In-School and Out-of-School Suspension: Behavioral and Psychological Outcomes in a Predominately Black Sample of Middle School Students, School Psychology Review, 52:1, 1-14, DOI: 10.1080/2372966X.2020.1856628

To link to this article: https://doi.org/10.1080/2372966X.2020.1856628

Published online: 02 Feb 2021.

SCHOOL PSYCHOLOGY REVIEW
2023, VOL. 52, NO. 1, 1–14
https://doi.org/10.1080/2372966X.2020.1856628

In-School and Out-of-School Suspension: Behavioral and Psychological Outcomes in a Predominately Black Sample of Middle School Students

Daniel R. Cohen[a], Crystal Lewis[b], Colleen L. Eddy[b], Lauren Henry[c], Caroline Hodgson[b]*, Francis L. Huang[b], Wendy M. Reinke[b], and Keith C. Herman[b]

[a]University of Alabama; [b]University of Missouri; [c]Children's Hospital Colorado

ABSTRACT

The use of suspension practices is extremely widespread but few studies have examined the behavioral and psychological outcomes associated with their application. Using a predominantly Black sample of 788 middle school students from the Midwestern United States, the current study evaluates the relations between in-school suspensions (ISS) and out-of-school suspensions (OSS) received during the course of the school year and student self-efficacy, engagement, prosocial behavior, emotion regulation, concentration, internalizing problems, and disruptive behavior based on student and teacher ratings collected at the end of the school year. Regression models were used to evaluate associations between the total number of ISS and OSS exposures on end of school year outcome measures controlling for beginning of school year measures and demographic characteristics. Results indicated that ISS and OSS are both associated with less prosocial behavior, lower levels of emotion regulation, and a greater extent of disruptive behavior and concentration problems at the end of the school year, even after controlling for these behaviors at the start of the school year. Implications of the potential impacts and distribution of suspension practices are discussed.

IMPACT STATEMENT

Within a predominantly Black sample of middle school students, in-school and out-of-school suspensions are associated with lower prosocial behavior and emotion regulation, and more concentration problems and disruptive behavior at the end of the school year, even after accounting for ratings on these outcomes from the beginning of the school year.

ARTICLE HISTORY
Received November 19, 2020
Accepted November 23, 2020

KEYWORDS
suspension, behavior, mental health, social justice

ASSOCIATE EDITOR
Erin Dowdy

THE USE OF SUSPENSION IN SCHOOLS

The use of disciplinary practices that remove students from the classroom environment is widespread. According to the U.S. Department of Education (2018), 2.7 million students throughout the United States received one or more out-of-school suspensions during the 2015–2016 school year, which represents 4.7% of all students (Harper et al., 2019). Although suspension rates across the country vary dramatically by district and by subgroups of students, in the 2015–2016 school year, school children in this country lost over 11 million days of instruction (11,360,004) as a result of out-of-school suspension (Losen & Whitaker, 2017).

It is critical to recognize that students of color, particularly Black students are disproportionally exposed to suspension. Based on the 2015–2016 civil rights data collection, although Black males make up roughly 8% of students attending US public schools, they represent 25% of those receiving out of school suspensions (U.S. Department of Education, 2018). The U.S. Government Accountability Office (2016) notes that high poverty schools with high proportions of students of color (75% or greater Black and or Latinx) represent 22% of students with suspensions, even though such schools represent only 16% of all students.

Given the disproportionate application of exclusionary practices with Black students and its apparent association with the racial opportunity gap (Gregory et al., 2010), the in-depth study of suspension and associated outcomes is essential. A large body of existing scholarship indicates that these practices are associated with poor student outcomes such as lower academic achievement and higher probability of dropout (Noltemeyer et al., 2015), that they are not an effective strategy for promoting safety or order (Lamont et al., 2013; Way, 2011) and that they facilitate school pathways to the juvenile justice system (Hughes et al., 2020; Skiba et al., 2014). One study found that exposure to a single suspension in 9th grade was associated with

CONTACT Daniel R. Cohen drcohen1@ua.edu Educational Studies in Psychology, Research Methodology, and Counseling, University of Alabama, 215F Graves Hall, Box 870231, Tuscaloosa, AL 35487, USA.
*Nationwide Children's Hospital
This article has been corrected with minor changes. These changes do not impact the academic content of the article.

 DOI: 10.1080/2372966X.2020.1856628

two times the risk of subsequent dropout (Balfanz et al., 2014). This is noteworthy given that those who drop out of high school are more likely to subsequently experience a variety of adverse outcomes including being arrested, getting fired from a job, using illegal drugs, and having poor health (Lansford et al., 2016). Another study found that each day of suspension was associated with decreasing reading achievement gains over time (Arcia, 2006). Suspension is also associated with incarceration. For example, Cuellar and Markowitz (2015) matched school data describing the date of out-of-school suspensions with juvenile justice system data describing the date arrest for individual students. Their results indicate that when students are outside of the school environment due to suspension, they are more than twice as likely to be suspended. In short, exposure to suspension can be life-altering.

The Link between Suspension and Negative Outcomes

Theory and research on dynamic systems theory and the process of developmental cascades can help to explain how suspensions contribute to the risk of negative outcomes. Outcomes like dropping out of school, having a substance use problem, and being incarcerated are often the result of a string of risk and compounding systemic failures in multiple domains across childhood and adolescence (Dodge et al., 2009; Paterson et al., 1989).

Suspensions fit into this developmental model as a potential risk factor for multiple domains because suspensions exclude students from the typical learning environment, which increases risk and undermines protective factors (Reinke & Walker, 2006). Children spend more time in school than in any other formal institutional structure and as such, schools play a key part in children's development in a variety of domains. Schools are a critical setting for mastering important developmental tasks including social interactions with peers and other adults, academic learning and cognitive progress, emotional and behavioral control, and physical and moral development (Fazel et al., 2014).

Racism plays an essential role in the disproportionate distribution of suspensions and the associated impacts represent a critical civil rights issue (Simson, 2013). Although the factors that result in disproportionate discipline and subsequent harm to Black students are multidimensional and complex, structural (e.g., school policy) and interpersonal (e.g., educator bias) forms of racism serve as underlying and enduring factors (Crutchfield et al., 2020; Pachter et al., 2010; Thomas et al., 2009). Therefore, efforts to respond to the burden of exclusionary discipline on Black youth must address sources of racism.

Empirical work by Okonofua and Eberhardt (2015) suggests that racial bias contributes to disparities in discipline, finding within an experimental context that teachers are more likely to interpret behaviors of Black students as problematic and are more likely view such behaviors among Black students as part of an ongoing pattern. This bolsters conclusions from previous work highlighting that Black students are more likely to be punished for subjective infractions (Skiba et al., 2002) and that problem behavior does not explain racial Black-White disparities in discipline (Huang, 2016a; 2020).

School policy also plays an important role in persisting disproportionality in school discipline. The behavioral expectations that undergird school polices are overwhelmingly based on White middle-class social norms and when students of color behave outside of these standards, even in the absence of objectively dangerous or otherwise harmful student actions, school polices often require that educational professionals respond with suspension (Fenning & Rose, 2007). Another important and related factor is the emphasis on zero tolerance policies. Based on the premise that harsh discipline would prevent violence, congress passed the Gun Free Schools Act of 1993, requiring school to expel students who are found in possession of firearms (Curran, 2016). From this initial law, school policies evolved to apply the zero tolerance philosophy across variety of infractions establishing a number of immediate pathways to school exclusion (Martinez, 2009). Ultimately, these approaches elevate exclusion and harsh punishment rather than prevention and positive youth development strategies as key frameworks for responding to student behavior problems. Coupled with interpersonal racial bias, zero tolerance policies have conferred extensive harm upon students of color.

Negative developmental processes associated with the disproportionate application of suspensions culminate in school pathways to the juvenile justice system (Hughes et al., 2020). Cuellar and Markowitz (2015) found that suspension is associated with greater risk of arrest, especially for Black students compared to White students. Another study found that even after controlling for delinquent behaviors, racial disparities in suspension are associated with racial disparities in arrest (Nicholson-Crotty et al., 2009).

In mapping out the connection between school exclusion and incarceration, Skiba et al. (2014) describe some aspects of a developmental pathway that can be applied to understand this link. Specifically, they propose that exclusionary practices lead to individual level and school level risk factors that negatively impact student outcomes. Within this model, these processes subsequently increase risk for dropout which is followed by a higher probability of incarceration. Even though Skiba et al. (2014) focus on decreased student engagement as a key proximal consequence of exclusion, it is likely that this factor alone does not capture all the critical dispositional consequences that

directly emerge from suspension exposure. Research gaps associated with the direct impacts of suspension are consistent with the limitations of this model, particularly because there are few studies that evaluate social and emotional impacts of exclusionary practice (Welsh & Little, 2018). In addition to engagement, outcomes that are theoretically and in some cases empirically relevant to suspension include externalizing and internalizing problems, and self-efficacy.

Impact on Engagement

Given that suspension by definition involves the removal of a student from the classroom environment, it is logical to hypothesize that reduced engagement is associated with suspension, and several cross-sectional studies support this notion. For example, Davis and Jordan (1994) investigated a subsample of over 1000 Black students from the National Education Longitudinal Study of 1988 at grades 8 and 10 and found that in high school, the number of out-of-school suspensions received was associated with decreased school engagement. Hinojosa (2008) examined the association between individual student-level factors and the probability of receiving in-school and out-of-school suspension in a sample of over 18,000 students in the Midwestern US. Within models controlling for a number of factors including student behavior, home environment, and beliefs about teachers, academic engagement was negatively associated with both types of suspension. Another study utilized data from the National Longitudinal Health and found that students attending schools with discipline policies that allow for the use of out-of-school suspension for less severe infractions reported having lower levels of school connectedness (McNeely et al., 2002). Given the cross-sectional nature of these studies, it is difficult to determine whether suspension precedes low engagement or vice versa; however, as noted by Toldson et al. (2015), it appears that discipline and engagement have a cyclical relationship, where low engagement leads to more discipline, and more discipline subsequently leads to lower engagement.

Impact on Externalizing and Internalizing Problems

In addition to interfering with student engagement, a number of studies have found that exposure to suspension is associated with externalizing and internalizing symptoms. A salient explanation for these findings provided by Hemphill et al. hypothesize that "suspending students from school may disconnect them from a positive social environment and increase exposure to risk factors" (2006, p. 741). One study of roughly 4000 students between the ages of 12-16 years old in Australia and the US found that exposure to suspension was associated with an increased probability of subsequent antisocial behavior including theft, violence, and substance abuse, even after controlling for a variety of risk and protective factors (Hemphill et al., 2006). The presence of these consequences suggests that students exposed to suspension may engage in fewer prosocial behaviors and be less equipped to regulate their emotions compared to other students. This is consistent with the research indicating that the application of punitive disciplinary practices in school and home contexts with youth displaying risk factors are associated with more maladaptive behaviors and poor emotion regulation (Capaldi et al., 1997; Scaramella & Leve, 2004)

With respect to internalizing problems, a number of cross-sectional studies have found associations between suspension and depressive symptomology (e.g. Flament et al., 2001; Rushton et al., 2002). In addition, one longitudinal study of over 5000 British children and adolescents found that students excluded from school experienced higher levels of psychological distress (Ford et al., 2018).

Potential Impact on Concentration Problems at School

Given the well-established associations between suspension and poor academic outcomes (Noltemeyer et al., 2015), it is possible that when children are removed from the classroom, they miss out on opportunities to further develop concentration and executive functioning skills. While there are no studies evaluating these associations for students exposed to suspensions, if present, this mechanism would also be applicable to students with low attendance. This is supported by research indicating that students with high levels of truancy display poor concentration in the classroom (Reid, 2005). Consistent with the theoretical model provided by Skiba et al. (2014), the presence of concentration problems in classroom settings may be another example of how exclusionary discipline leads to the development of individual risk factors, but further study in this area is needed.

Potential Impact on Self-Efficacy

According to Bandura (1997), a student's perceived self-efficacy in academic contexts is largely driven by their previous efforts in related areas, also referred to as mastery experiences. As such, suspension experiences are likely to degrade student self-efficacy because if a student is repeatedly removed from the classroom environment they lose opportunities to obtain mastery experiences. In the case of Black students and their disproportionate exposure to suspension,

 DOI: 10.1080/2372966X.2020.1856628

this may represent an important opportunity gap (Welner & Carter, 2013). Caprara et al. (2008) found that self-efficacy was associated with middle school grades and mediated the relation between middle and high school grades and subsequent dropout. Even though there are no studies to our knowledge that assess the impact of suspension exposure on student reported self-efficacy, given its potential impact as a mastery experience and the importance of self-efficacy with respect to student outcomes (Usher & Pajares, 2006), the assessment of this link is an essential area of study.

Literature Gaps in the Study of Suspension in Black Students

To our knowledge, the vast majority of extant studies on the impact of suspension on Black students focus on school level outcomes [see Cuellar and Markowitz (2015) and Davis and Jordan (1994) discussed above as exceptions]. One study by Lee et al. (2011) in a sample of almost 300 hundred public schools in Virginia found that the number of students receiving one or more suspensions was associated with the school level dropout rate for both White and Black students, but the strength of the association was stronger for White students. Another study utilized longitudinal data from a large urban school district in Kentucky and found that exposure to suspensions by Black students is responsible for roughly one-fifth of the differences in academic achievement between Black and White students (Morris & Perry, 2016). The current study contributes to the extremely limited literature on student-level behavioral and psychological outcomes associated with suspension within a predominantly Black student population.

Purpose of This Study

Although there is an extensive body of research documenting poor academic outcomes for students receiving exclusionary discipline (Noltemeyer et al., 2015), limited research has evaluated associations with other areas of student functioning such as prosocial behaviors, self-efficacy, and emotional regulation (Welsh & Little, 2018). Also, existing studies tend to utilize purely cross sectional study designs with data from a single time point (e.g. Davis & Jordan, 1994; Rushton et al., 2002). This is an important gap because it would be expected that social and emotional problems would be present prior to exposure to exclusionary practices. As such, studies that control for baseline levels of outcomes of interest (i.e., beginning of the year ratings) discrete frequency of suspension exposures, and outcome data from a subsequent time point (i.e. end of year ratings) are needed. In addition, the vast majority of studies evaluate the associations with out-of-school suspension, but very few focus on impacts of in-school suspension. The present study extends this work by using beginning and end-of-year data to evaluate how the frequency of in-school and out-of-school suspension impacts a number of behavioral and psychological outcomes, controlling for baseline ratings of those measures. Dependent variables include self-report and teacher-report measures of behavioral engagement, self-efficacy, concentration problems, disruptive behavior, prosocial behavior, emotion regulation, depression, and internalizing problems,

METHOD

Sample

The current study uses data collected as part of a cluster randomized controlled trial evaluating the impact of a classroom management intervention in the Midwestern United States over four years. CHAMPS, the intervention that was the basis of the larger trial, is a classroom management training program that includes ongoing coaching for teachers focused on positive reinforcement, promoting on-task behavior, and establishing clear expectations for the classroom (Herman et al., 2020). The study received IRB approval. All middle schools (grades 6-8) within two neighboring school districts were included in the study and all math and reading teachers in those schools were invited to participate. In the first year of the study (2013–14), demographics of schools were 77% Black, 66% free or reduced lunch. Teachers that consented were included in the study and then randomly assigned to treatment or control conditions. Each teacher chose their most challenging classroom to participate in the study. Overall student consent rate to participate in the study was 75%. The study design included randomization at the classroom level, and as a result, some students were in the sample multiple times because they had class periods with multiple teachers in the study. Duplicate students were randomly sampled so that only one record per student was retained. The sample includes 788 middle school students in 75 classrooms across eight middle schools, who did not move during the school year[1] (50% male; 75% free or reduced lunch; 74% Black, 22% White, 2% Hispanic/Latino, 1% Asian, 1% other). Students comprised three cohorts given that the intervention was delivered to three separate groups of classrooms for a duration of one year across three total years. Of 75 teachers in the sample, 73% identified as female, 71% were White, 23% Black, 4% other, 1% Asian and 1% Pacific Islander.

Measures

Across all student and teacher-reported measures, data were collected in the early fall (October) and late spring (May). Student self-report measures were collected in the

classroom setting and teacher report measures were collected using Qualtrics. Teachers completed all measures for each student in the sample except for cases with unintentional missing data (i.e., data could not be collected due to unforeseen circumstances). Administrative data were provided at the end of the school year.

Frequency of in-School and Out-of-School Suspensions

Sum of in-school and out-of-school suspension during the course of the entire school year were provided by school districts through the use of administrative data.

Behavioral Engagement (Student Reported)

The Behavioral Engagement subscale, which is part of the Engagement Versus Disaffection With Learning: Student-Report measure is comprised of the mean of 5 items that assess a student's active involvement in classroom learning activities (Skinner et al., 2008). Skinner et al. (2009) found that structural analyses, teacher ratings, and behavioral observations supported construct validity of the subscale. Items included "I try hard to do well in this class" and "I pay attention in class". Responses ranged from 1= *not at all true* to 4 = *very true*. Alphas in the trial study ranged from .77 to .82.

Self-Efficacy in School (Student Reported)

The mean of student reported self-efficacy was assessed using the three item Self Efficacy in School measure, which evaluates student perceptions of their ability to function successfully in the school context (Cook et al., 2000). Cook et al. (1999) used this measure to evaluate a school-wide behavior program and found strong evidence of convergent validity. Items included "How well can you live up to what your teachers expect of you?" and "How well can you get teachers to help you when you get stuck on schoolwork?". Response options ranged from 1 = *not at all well* to 6 = *very well*. Alphas in the trial study ranged from .56 to .74.

Depression (Student Reported)

In order to assess student reported depression, the sum of 8 items of the Patient Health Questionnaire-9 (PHQ9, Kroenke et al., 2001) were administered to students, with the self-harm item omitted due to practical and ethical considerations associated with asking about suicide risk in the context of a large-scale survey. The PHQ9 has high levels of discriminant validity and sensitivity to change. Questions included "In the past two weeks, how often have you been bothered by feeling down, depressed, irritable, or hopeless?". Response options ranged from 0 = *not at all* to 3 = *nearly every day*. Alphas in the trial study ranged from .79 to .88.

TOCA-C Scales (Teacher Reported)

Concentration Problems, Disruptive Behavior, Prosocial Behavior, Emotion Dysregulation, and Internalization subscales were derived from the Teacher Observation of Classroom Adaptation-Checklist (TOCA-C) which is completed by teachers based on their classroom observations (Koth et al., 2009). Koth et al. (2009) found strong evidence of construct validity using exploratory factor analysis. The measure lists a number of characteristics with the following prompt at the beginning, "would you say the following statements were never, rarely, sometimes, often, very often, or almost always true of this child." The 7 item Concentration Problems scale assesses difficulties related to on task-behavior in the classroom environment. Examples of student characteristics listed in this subscale include, "concentrates", "works hard", and "completes assignments." The 9 item Disruptive Behavior scale assesses rule-breaking and aggression. Examples of student characteristics listed in this subscale include "breaks rules", "teases classmates", and "harms property." The 5 item Prosocial Behavior scale assesses having positive interactions with others and compliance. Examples of student characteristics listed in this subscale include "is friendly", "shows empathy and compassion for others feelings", and "is liked by classmates." The 5 item Emotional Dysregulation scale assesses problems with self-control and managing emotional states. Examples of student characteristics listed in this subscale include "impulsive", "easily frustrated", and "changes moods quickly." Finally, the Internalization scale assesses symptoms of anxiety and depression. Examples of student characteristics listed in this subscale include "nervous", "withdrawn", and "sad." Means of scales were used in analysis. Response options ranged from 1 = *never* to 6 = *almost always*. Subscale alphas in the trial study ranged from .85 to .98.

Analytic Strategy

Multiple linear regression was used to evaluate the relationship between the frequency of in-school and out-of-school suspension and behavioral and psychological ratings in the spring. Self-reported dependent variables include behavioral engagement, self-efficacy, and depression. Teacher reported dependent variables include concentration problems, disruptive behavior, prosocial behavior, emotional dysregulation, and internalization. For each outcome, first a model was run controlling for fall ratings of outcomes of interest, and demographic factors including race, gender, and free or reduced lunch status. Then, to determine the change in R^2, a second model was run adding either the ISS or OSS variables as predictors. Since the only variables of interest were at the student level, teacher fixed effects were used to control for the clustering of students within classrooms and intervention status,

 DOI: 10.1080/2372966X.2020.1856628

as this was considered noise that simply needed to be accounted for (Huang, 2016b). Also, analyses included cluster robust standard errors at the teacher level. Assumptions of linear regression were checked and linear regression was an appropriate method to use for these analyses. Given the positive skew of ISS and OSS variables, we compared predicted versus residual graphs and found no evidence of non-linearity. The Benjamini-Hochberg correction was used with the independent variables of interest (ISS Sum and OSS Sum) to reduce the likelihood of type 1 error (Institute of Educational Sciences, 2017). Analyses were run using Stata 14.2.

Our analytics sample included 788 students who were in school for the academic year over three cohort years. Of the 788 students, 4.8% were missing data on at least one variable of interest. Data for these students were missing because of missed questions or survey incompletion. The variables with the largest amount of missing data were student reported spring Depression (2.4%) and student reported spring Behavioral Engagement (2.2%). Missing data were handled using full-information maximum likelihood estimation which, like multiple imputation, is considered a principled way of accounting for missingness (Dong & Peng, 2013; Enders, 2010). Two regression models are listed below which provide a general overview of models with (1) in-school suspension sum as the independent variable, and (2) out-of-school suspension sum as the independent variable. In each model, $(\text{Dependent Variable T2})_{it}$ represents the score for each of the of the seven outcome variables under study for student *i* with teacher *t* collected at the end of the school year. $(\text{Dependent Variable T1})_{it}$ represents the score of the corresponding variable collected at the beginning the of the school year. φ_t refers to teacher fixed effects and ξ is the error term.

Model 1 – In-School Suspension Models

$$(\text{Dependent Variable T2})_{it} = \beta_0 + \beta_1(\text{ISS Sum})_{it} + \beta_2(\text{Dependent Variable T1})_{it} + \beta_3(\text{Female})_{it} + \beta_4(\text{Free or Reduced Lunch Status})_{it} + \beta_5(\text{Black})_{it} + \beta_6(\text{Other Race})_{it} + \beta_7(\text{TeacherID})_{it} + \varphi_t + \xi.$$

Model 2 – Out-of-School Suspension Models

$$(\text{Dependent Variable T2})_{it} = \beta_0 + \beta_1(\text{OSS Sum})_{it} + \beta_2(\text{Dependent Variable T1})_{it} + \beta_3(\text{Female})_{it} + \beta_4(\text{Free or Reduced Lunch Status})_{it} + \beta_5(\text{Black})_{it} + \beta_6(\text{Other Race})_{it} + \beta_7(\text{Teacher-ID})_{it} + \varphi_t + \xi.$$

RESULTS

Descriptive Statistics

Descriptive statistics associated with key variables are provided in Table 1 and 2. A total of 246 (31%) of students in the sample were exposed to a least one ISS with a range from 1 to 27. A total of 205 (26%) of students in the sample were exposed to a least one OSS with a range from 1 to 8. In order to assess the presence of disproportionality in suspensions across gender, race, and free or reduced lunch status, we assessed the relative risk of having one or more in-school and out-of-school suspension. For male compared to female students, the relative risk of receiving in-school suspension was 1.51 (95% CI: 1.22, 1.87) and 1.22 (95% CI: 0.97, 1.55) for out-of-school suspension suggesting disproportionality related to gender for ISS only. For Black compared to non-Black students, the relative risk of receiving in-school suspension was 1.85 (95% CI: 1.37, 2.49) and 2.18 (95% CI: 1.52, 3.12) for out-of-school suspension suggesting disproportionality related to race for both types of suspension. For student receiving free or reduced lunch compared to students without free or reduced lunch, the relative risk of receiving in-school suspension was 1.98 (95% CI: 1.45, 2.72) and 2.24 (95% CI: 1.55, 3.25) for out-of-school suspension suggesting disproportionality related to free or reduced lunch status for both types of suspension.

Associations Between ISS and OSS on Student Outcomes

Detailed results are provided in Tables 3–10. For all outcomes except student reported depression, the frequency of in-school suspension explained changes in the

Table 1. Pearson Correlation Matrix of Key Variables (n = 788)

	1	2	3	4	5	6	7	8	9
1. ISS Count									
2. OSS Count	0.53*								
3. Behavioral Engagement (Spring)	−0.12*	−0.13*							
4. PHQ-8 (Spring)	0.04	0.09*	−0.18*						
5. Self-Efficacy (Spring)	−0.10*	−0.04*	0.49*	−0.27*					
6. Concentration Problems (Spring)	0.39*	0.35*	−0.34*	0.18*	−0.29*				
7. Disruptive Behavior (Spring)	0.50*	0.47*	−0.25*	0.11*	−0.17*	0.66*			
8. Prosocial Behavior (Spring)	−0.37*	−0.38*	0.27*	−0.18*	0.26*	−0.71*	−0.73*		
9. Emotion Dysregulation (Spring)	0.40*	0.38*	−0.21*	0.18*	−0.18*	0.66*	0.85*	−0.72*	
10. Internalization (Spring)	0.10*	0.11*	−0.14*	0.17*	−0.13*	0.39*	0.40*	−0.52*	0.53*

Note. *p<.05.

Table 2. Descriptive Statistics for Key Study Variables (n = 788)

Variable	Mean	SD
ISS Count	1.06	2.47
OSS Count	0.49	1.10
Behavioral Engagement (Spring)	3.35	0.48
PHQ-8 (Spring)	5.74	5.34
Self-Efficacy (Spring)	4.07	0.72
Concentration Problems (Spring)	3.00	1.30
Disruptive Behavior (Spring)	2.04	0.86
Prosocial Behavior (Spring)	4.46	1.03
Emotion Dysregulation (Spring)	2.56	1.14
Internalization (Spring)	2.02	0.79

outcomes above demographics and the baseline of the outcomes (ΔR^2 ranging from .002 to .025). Similarly, for all outcomes except student reported self-efficacy, the frequency of out-of-school suspensions predicted changes in the outcomes above demographics and the baseline of the outcomes (ΔR^2 ranging from .001 to .020). Multiple regression analyses controlling for baseline scores of study outcomes indicate that the frequency of in-school suspension significantly predicted lower ratings of prosocial behavior ($b = -0.065$, $p \leq .001$, $f^2 = 0.049$), and higher ratings of concentration problems ($b = 0.078$, $p \leq .001$, $f^2 = 0.052$), disruptive behavior ($b = 0.071$, $p \leq .001$, $f^2 = 0.086$), and emotion dysregulation ($b = 0.068$, $p \leq .001$, $f^2 = 0.048$). The frequency of out-of-school suspension significantly predicted lower ratings of prosocial behavior ($b = -0.156$, $p \leq .001$, $f^2 = 0.058$) and higher ratings of concentration problems ($b = 0.118$, $p \leq .001$, $f^2 = 0.027$), disruptive behavior ($b = 0.119$, $p \leq .001$, $f^2 = 0.050$) and emotion dysregulation ($b = 0.114$, $p \leq .01$, $f^2 = 0.023$). For all statistically significant associations between suspension and study outcomes, Cohen's f^2 values were indicative of small effect sizes (Cohen, 1988). Neither form of exclusionary discipline was significantly associated with measures of self-efficacy, behavioral engagement, depression, or internalizing problems.

The Relation Between Race and Student Outcomes

Multiple regression analyses controlling for suspension exposure and other demographic characteristics indicated that compared to White students, Black students were rated as having more problem behaviors by teachers. Specifically, in ISS and OSS models teachers rated Black students as having greater concentration problems ($b = 0.291$, $p \leq .01$; $b = 0.310$, $p \leq .001$), being more disruptive ($b = 0.169$, $p \leq .01$; $b = 0.180$, $p \leq .01$) and emotionally dysregulated ($b = 0.186$, $p \leq .01$; $b = 0.201$, $p \leq .01$), and engaging in less prosocial behavior ($b = -0.176$, $p \leq .05$; $b = -0.178$, $p \leq .05$).

DISCUSSION

This study provides further evidence that exclusionary discipline is associated with important negative outcomes related to youth behavior and psychological well-being.

Table 3. In-School Suspension Models Part 1 (n = 788)

	Self-Report: Behavioral Engagement (Spring)				Self-Report: Self-Efficacy (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
ISS Sum	−0.007	0.007	−0.037	−0.11, 0.036	−0.015	0.012	−0.053	−0.135, 0.029
Female	0.079*	0.032	0.163	0.036, 0.291	−0.025	0.033	−0.035	−0.126, 0.057
Black	−0.037	0.048	−0.076	−0.271, 0.118	0.028	0.080	0.039	−0.181, 0.258
Other Race	−0.049	0.086	−0.101	−0.448, 0.246	−0.139	0.099	−0.194	−0.465, 0.077
Free or Reduced Lunch	−0.096*	0.040	−0.198	−0.36, −0.035	0.011	0.056	0.015	−0.139, 0.17
Fall Rating of DV	0.538***	0.038	0.517	0.446, 0.588	0.699***	0.049	0.705	0.608, 0.803
Intercept	1.541	0.163	0.091	−0.127, 0.309	1.321*	0.220	0.172	0.012, 0.333
R2			.377				.557	
Δ R^2			.002				.003	

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without ISS sum to the model including ISS sum.

Table 4. In-School Suspension Models Part 2 (n = 788)

	Self-Report: Depression (Spring)				Teacher-Report: Concentration Problems (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
ISS Sum	0.015	0.103	0.007	−0.086, 0.100	0.078***	0.011	0.148	0.106, 0.190
Female	0.849**	0.319	0.159	0.042, 0.276	−0.160*	0.073	−0.123	−0.232, −0.014
Black	0.168	0.514	0.031	−0.157, 0.22	0.291**	0.093	0.223	0.084, 0.362
Other Race	0.257	0.874	0.048	−0.272, 0.369	−0.046	0.160	−0.036	−0.276, 0.205
Free or Reduced Lunch	−0.021	0.489	−0.004	−0.183, 0.175	0.028	0.074	0.021	−0.09, 0.133
Fall Rating of DV	0.509***	0.045	0.500	0.413, 0.586	0.624***	0.034	0.616	0.551, 0.68
Intercept	2.902	0.673	0.018	−0.196, 0.232	0.394***	0.119	−0.512	−0.668, −0.355
R2			.330				.670	
Δ R^2			.000				.017	

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without ISS sum to the model including ISS sum.

Table 5. In-School Suspension Models Part 3 (n = 788)

	Teacher-Report: Disruptive Behavior (Spring)				Teacher-Report: Prosocial Behavior (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
ISS Sum	0.071***	0.017	0.203	0.108, 0.298	−0.065***	0.013	−0.154	−0.215, −0.094
Female	−0.068*	0.033	−0.079	−0.155, −0.003	0.111*	0.049	0.107	0.015, 0.200
Black	0.169**	0.056	0.197	0.068, 0.326	−0.176*	0.083	−0.170	−0.327, −0.013
Other Race	−0.037	0.078	−0.043	−0.222, 0.136	0.159	0.118	0.154	−0.069, 0.377
Free or Reduced Lunch	−0.018	0.040	−0.022	−0.114, 0.071	0.088	0.064	0.085	−0.035, 0.206
Fall Rating of DV	0.645***	0.054	0.589	0.492, 0.687	0.667***	0.037	0.617	0.549, 0.685
Intercept	0.684	0.100	−0.076	−0.242, 0.089	2.085***	0.177	0.541	0.389, 0.693
R2		.710				.653		
Δ R^2		.025				.017		

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without ISS sum to the model including ISS sum.

Table 6. In-School Suspension Models Part 4 (n = 788)

	Teacher-Report: Emotion Dysregulation (Spring)				Teacher-Report: Internalization (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
ISS Sum	0.068***	0.016	0.147	0.077, 0.217	0.016	0.010	0.049	−0.010, 0.108
Female	−0.029	0.055	−0.026	−0.121, 0.069	−0.021	0.037	−0.026	−0.118, 0.065
Black	0.186**	0.069	0.163	0.044, 0.282	−0.004	0.058	−0.005	−0.150, 0.140
Other Race	−0.134	0.116	−0.118	−0.316, 0.081	−0.056	0.131	−0.072	−0.398, 0.255
Free or Reduced Lunch	−0.069	0.051	−0.060	−0.148, 0.027	−0.071	0.040	−0.090	−0.189, 0.008
Fall Rating of DV	0.649***	0.036	0.600	0.535, 0.665	0.545***	0.040	0.542	0.464, 0.620
Intercept	0.649**	0.096	−0.236	−0.395, −0.077	0.502***	0.079	−0.603	−0.779, −0.428
R2		.664				.609		
Δ R^2		.016				.002		

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without ISS sum to the model including ISS sum.

Table 7. Out-of-School Suspension Models Part 1 (n = 788)

	Self-Report: Behavioral Engagement (Spring)				Self-Report: Self-Efficacy (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
OSS Sum	−0.025	0.015	−0.056	−0.123, 0.012	−0.007	0.021	−0.011	−0.073, 0.052
Female	0.077*	0.031	0.159	0.034, 0.284	−0.019	0.033	−0.027	−0.116, 0.063
Black	−0.035	0.048	−0.072	−0.267, 0.123	0.018	0.081	0.025	−0.197, 0.247
Other Race	−0.049	0.086	−0.100	−0.448, 0.247	−0.139	0.102	−0.194	−0.472, 0.084
Free or Reduced Lunch	−0.092*	0.040	−0.191	−0.352, −0.029	0.008	0.057	0.011	−0.144, 0.166
Fall Rating of DV	0.538***	0.039	0.517	0.444, 0.590	0.704***	0.049	0.710	0.612, 0.807
Intercept	1.531	0.167	0.062	−0.151, 0.275	1.260	0.215	0.131	−0.033, 0.295
R2		.378				.554		
Δ R^2		.003				.000		

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without OSS sum to the model including OSS sum.

Table 8. Out-of-School Suspension Models Part 2 (n = 788)

	Self-Report: Depression (Spring)				Teacher-Report: Concentration Problems (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
OSS Sum	0.211	0.186	0.043	−0.032, 0.118	0.118***	0.028	0.099	0.054, 0.145
Female	0.893**	0.296	0.167	0.059, 0.276	−0.161*	0.071	−0.123	−0.23, −0.017
Black	0.116	0.518	0.022	−0.168, 0.212	0.310***	0.090	0.237	0.101, 0.373
Other Race	0.258	0.877	0.048	−0.274, 0.37	−0.043	0.170	−0.033	−0.288, 0.223
Free or Reduced Lunch	−0.069	0.485	−0.013	−0.191, 0.165	0.020	0.074	0.015	−0.096, 0.126
Fall Rating of DV	0.506***	0.044	0.497	0.412, 0.581	0.648***	0.034	0.639	0.574, 0.705
Intercept	2.862	0.592	0.023	−0.167, 0.213	0.499***	0.113	−0.396	−0.542, −0.249
R2		.332				.662		
Δ R^2		.002				.009		

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without OSS sum to the model including OSS sum.

Table 9. Out-of-School Suspension Models Part 3 (n = 788)

	Teacher-Report: Disruptive Behavior (Spring)				Teacher-Report: Prosocial Behavior (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
OSS Sum	0.119***	0.032	0.153	0.073, 0.232	−0.156***	0.023	−0.164	−0.212, −0.117
Female	−0.072	0.034	−0.084	−0.161, −0.007	0.111*	0.047	0.107	0.018, 0.196
Black	0.180**	0.055	0.211	0.084, 0.337	−0.178*	0.075	−0.172	−0.314, −0.029
Other Race	−0.031	0.087	−0.037	−0.237, 0.164	0.162	0.122	0.156	−0.075, 0.387
Free or Reduced Lunch	−0.025	0.042	−0.030	−0.125, 0.066	0.101	0.064	0.098	−0.024, 0.219
Fall Rating of DV	0.680***	0.053	0.622	0.526, 0.718	0.664***	0.036	0.614	0.548, 0.679
Intercept	0.774	0.104	0.088	−0.051, 0.227	1.983***	0.160	0.420	0.268, 0.573
R2		.700				.656		
ΔR^2		.015				.020		

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without OSS sum to the model including OSS sum.

Table 10. Out-of-School Suspension Models Part 4 (n = 788)

	Teacher-Report: Emotion Dysregulation (Spring)				Teacher-Report: Internalization (Spring)			
Variable	*b*	*Cluster Robust SE*	β†	β† 95% CI	*b*	*Cluster Robust SE*	β†	β† 95% CI
OSS Sum	0.114**	0.040	0.109	0.033, 0.185	0.028	0.024	0.049	−0.01, 0.108
Female	−0.036	0.054	−0.032	−0.126, 0.062	−0.022	0.037	−0.026	−0.118, 0.065
Black	0.201**	0.067	0.176	0.06, 0.292	7.64E-06	0.058	−0.005	−0.15, 0.14
Other Race	−0.128	0.122	−0.112	−0.321, 0.097	−0.056	0.131	−0.072	−0.398, 0.255
Free or Reduced Lunch	−0.074	0.053	−0.065	−0.156, 0.026	−0.072	0.040	−0.090	−0.189, 0.008
Fall Rating of DV	0.668***	0.035	0.618	0.554, 0.682	0.545***	0.040	0.542	0.464, 0.62
Intercept	0.758	0.092	−0.114	−0.257, 0.028	0.533***	0.073	−0.603	−0.779, −0.428
R2		.656				.608		
ΔR^2		.008				.001		

Notes. *p≤.05 **p ≤ .01 ***p≤.001; †standardized beta weight; all regression models account for teacher fixed effects; ΔR^2 is the change from the model without OSS sum to the model including OSS sum.

This is particularly problematic given the widespread reliance on these methods (U.S. Department of Education, 2018) and that their application is often justified by the notion that their use will improve student behavior and reduce violence (Lamont et al., 2013). Within a predominately Black sample, results indicate that exposure to ISS during the school year is associated with more concentration problems, greater levels of disruptive behavior and emotion dysregulation, and less prosocial behavior based on data collected at the end of the school year controlling for ratings in these domains at the beginning of the school year. Further, exposure to OSS during the school year was associated with more concentration problems, disruptive behavior, and emotion dysregulation, and less prosocial behavior based on end of year ratings, controlling for fall ratings. However, in contrast to previous cross-sectional findings, depressive symptoms and engagement were not associated with ISS or OSS exposure. Also, we did not find evidence supporting the hypothesized link between self-efficacy and exposure to suspension.

With regard to the non-significant findings associated with depression and internalization, given previous findings in cross-sectional studies, this may be indicative of a process where although depressed students may be more likely to be suspended, suspension exposure may not be an important cause of depression. The absence of an association between suspension and engagement is surprising given that it has been hypothesized as a core feature of conceptual models describing how exclusionary discipline leads to long term negative outcomes (Skiba et al., 2014). The lack of association may be related to the limited scope of the measure which is only inclusive of academic engagement. Given the non-significant findings associated with self-efficacy, it may be that suspensions do not inhibit mastery experiences to the extent expected, but this warrants further examination given that this is the first study to assess this relation. Alternatively, the absence of an association with self-efficacy may be related to limitations of the measure, which demonstrated low to moderate internal consistency.

Additionally, although most of the teacher measures yielded associations, none of the student reported measures had significant findings. In recognizing this, it is important to note that the only overlapping constructs, depression and internalization, had consistent findings across raters. Future studies on suspension and psychological/behavioral outcomes could be strengthened by the use of multiple raters across all constructs. In contrast, there was uniformity across results associated with both in-school and out-of-school suspension. This is an important finding given that ISS is often viewed as a more benign disciplinary approach. Ultimately, it appears that both forms of suspension serve to decrease learning opportunities and remove students from prosocial contexts. As such, it is not surprising that they have similar associations with negative outcomes.

 DOI: 10.1080/2372966X.2020.1856628

A notable set of findings is that across models that account for suspension frequency, sex, and free or reduced lunch status, compared to White students, Black students were rated as being more disruptive, having more concentration problems, being more emotionally dysregulated, and displaying fewer prosocial behaviors. These findings are consistent with existing evidence that racial prejudice plays an important role in biased perceptions of behavior problems (Okonofua & Eberhardt, 2015) and negative emotions (Halberstadt et al., 2018) in Black youth, on the part of educational professionals.

Our findings on the association of exclusionary discipline with negative student outcomes, taken together with previous research reviewing behavioral and academic effects associated with exclusionary discipline, provide additional evidence for the deleterious impact of suspension on student well-being. Simply put, these practices do not result in the intended outcomes associated with their application and a good deal of evidence suggests that their use results in harm to students with the greatest level of need. In spite of this evidence, the reliance on suspension persists due to zero-tolerance policies that were enacted based on the false premise that harsher punishments make schools safer (Curran, 2016; Martinez, 2009). In addition, support for zero-tolerance policies on the part of school administrators is associated with the notion that school discipline problems are best addressed by removing at-risk students rather than supporting them (Skiba & Edl, 2004).

Any discussion regarding the harms associated with in-school and out-of-school suspension must acknowledge the disproportionate application to students of color, particularly Black children, and those living in poverty (Skiba et al., 2014). Given the pervasive racial disparities in achievement (Quintana et al., 2012) and incarceration (Hetey & Eberhardt, 2014), there is a critical need to identify and address factors associated with interpersonal and structural racism. The social reproduction of racial inequality will continue in the absence of such efforts. If suspension truly causes harm in the manner suggested by this and other studies, it is incumbent upon educational systems to curb their reliance on such practices to protect the civil rights of all students.

A number of alternative universal discipline strategies are currently under study. Positive Behavioral Interventions and Supports (PBIS) emphasizes the systematic use of positive reinforcement and proactive disciplinary measures such as explicitly teaching behavioral expectations that are consistent across settings. Gage et al. (2018) conducted a meta-analysis of large scale experimental studies of PBIS and found a significant effect related reducing the use of suspension. Restorative practices are based on the premise that the use of structured conflict resolution activities will foster positive relationships and limit the need for and use of punitive disciplinary approaches. Several recent studies have engaged in the large scale evaluation of restorative practices in schools with scholarship demonstrating significant reductions in suspension associated with implementation (Anyon et al., 2016; Gregory et al., 2018). However, although both approaches have a strong impact on suspension use, few studies of restorative practices have demonstrated reductions in racial disparities (e.g. Augustine et al., 2018) and findings associated with PBIS are mixed (McIntosh et al., 2018). Ultimately, universal approaches that do not specifically account for sources of structural and interpersonal racism may not be effective in undoing inequalities in suspension. As Carter et al. (2017) note, "the roots of racial inequality in our schools and our society are many centuries deep. Eliminating disciplinary disparities, or for that matter any inequity in our educational system, will require an ongoing awareness of how those disparities are produced and a steadfast commitment to finally bringing them to an end" (p. 225). This argument is further bolstered by study results suggesting that teachers provide poorer behavioral ratings for Black students, even after controlling for free and reduced lunch status and the frequency of disciplinary infractions.

Efforts are in progress to evaluate and implement promising targeted approaches to addressing disproportionate discipline (e.g., Bradshaw et al., 2018; Cook et al., 2018; Gion et al., 2020). As targeted approaches become validated, resources must be devoted to widespread dissemination and educator training activities. However, there is a need to acknowledge that in addition to having no utility in achieving intended outcomes (Lamont et al., 2013), suspension practices result in significant harm, often to students with the greatest level of need. In addition, although the overall use of suspension has decreased over time and these declines appear to coincide with a greater awareness of the problem and applicable requirements associated with the Every Students Succeed Act, overrepresentation and disparities based on race and disability status persist (Harper et al., 2019). This study serves to lend further support to the critical narrative arguing that suspension practices must not be used in school unless students display a clear and present safety threat. The in-depth scientific study of iatrogenic practices and the dissemination of associated knowledge is critical to the process of changing associated behaviors.

LIMITATIONS

This study has several limitations. First, although the study relies on two raters of psychological and behavioral outcomes to evaluate student dispositions in response to suspension, the study does not include observational measures of student behavior. In addition, given that discipline variables are sourced from district level administrative data alone, we were not able to evaluate the accuracy of the

suspension data independently. Another limitation was that although there are multiple types of engagement relevant to school settings, we only evaluated the impact of academic engagement on study outcomes. Also, given that baseline measures were administered in October of each school year, it is possible that a small proportion of students received suspensions prior to collection of baseline data.

Finally, students with suspensions are different from those who do not receive them in meaningful ways and may be more likely to be suspended due to the presence of problems at baseline. In addition, students who have been suspended may be perceived more negatively than other students and may be more likely to have poor student-teacher relationships which would also negatively impact teacher perceptions of behavior. Methodologically, the ideal approach for studying the impact of suspension would be to do so experimentally where individuals are randomly assigned to receive suspension or not in response to behavioral infractions, but such an approach would be ethically untenable. This study represents an alternative method for studying such impacts given our inability to use gold standard experimental methods to evaluate the consequences of suspension. Although this study represents an incremental increase in methodological rigor from single time-point cross sectional studies because we are able to control for baseline ratings prior to exposure, the absence of an experimental design limits our ability to make inferences about the causal relation between the suspension and the outcomes under study. Future non-experimental studies may be able to utilize propensity score matching to address some of these concerns.

CONCLUSION

Given the number of students who receive suspensions every year, it is one of the most common responses to problem behavior in students. Thus, it is critical to understand the impact of these practices on our youth. This need is underscored by overwhelming evidence that nonwhite students, and Black students in particular, are more likely to receive suspensions and that they are applied in an educational landscape that presents significant opportunity gaps for the very same student populations. A large body of research has demonstrated negative impacts associated with academic outcomes, but the behavioral, cognitive, and emotional impacts are relatively understudied. This study contributes to our knowledge about the association of both in-school and out-of-school suspension with negative outcomes such as concentration problems, poor emotion regulation, and disruptive behavior within a largely Black sample of middle school students. In addition, study findings on the relation between race and teacher ratings are consistent with previous research that underscores the grave consequences of racial bias on educator judgement and expectations.

School psychologists have a professional and moral obligation to safeguard student well-being and advocate for the needs of Black students. The results of this study further emphasize the responsibility of practitioners, trainers, and researchers to develop, evaluate, support, and implement systems-level interventions that reduce the reliance on suspensions, particularly for Black students. School psychologists should apply their unique knowledge and skills to redirect energies towards culturally responsive evidence-based strategies that build on existing strengths, foster prosocial behaviors, and promote school connectedness. Study findings also underscore the need for evidence-based strategies to reduce racial bias among educators. In order for schools to meet the social, emotional, and academic needs of Black youth, school psychologists must take decisive action to address both structural and interpersonal manifestations of racism.

NOTE

1. Key study measures were not introduced until the second year of the study so the first cohort of data were removed from the sample.

DISCLOSURE

The authors have no conflicts of interest to report.

FUNDING

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A130143 to the University of Missouri (PI: Keith C. Herman). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

REFERENCES

Anyon, Y., Gregory, A., Stone, S., Farrar, J., Jenson, J. M., Mcqueen, J., Downing, B., Greer, E.,& Simmons, J. (2016). Restorative interventions and school discipline sanctions in a large urban school district. *American Educational Research Journal*, *53*(6), 1663–1697. https://doi.org/10.3102/0002831216675719.

Arcia, E. (2006). Achievement and enrollment status of suspended students. *Education and Urban Society*, *38*(3), 359–369. https://doi.org/10.1177/0013124506286947.

Augustine, C. H., Engberg, J., Grimm, G. E., Lee, E., Wang, E. L., Christianson, K., & Joseph, A. A. (2018). Can restorative practices improve school climate and curb suspensions. In *An evaluation of the impact of restorative practices in a mid-sized urban school district*. (pp. 1–112). Rand.

Balfanz, R., Byrnes, V., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and conse-

quences of being suspended in the ninth grade. *Journal of Applied Research on Children*, *5*(2), 13.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.

Bradshaw, C. P., Pas, E. T., Bottiani, J. H., Debnam, K. J., Reinke, W. M., Herman, K. C., & Rosenberg, M. S. (2018). Promoting cultural responsivity and student engagement through double check coaching of classroom teachers: An efficacy study. *School Psychology Review*, *47*(2), 118–134. https://doi.org/10.17105/SPR-2017-0119.V47-2

Capaldi, D. M., Chamberlain, P., & Patterson, G. R. (1997). Ineffective discipline and conduct problems in males: Association, late adolescent outcomes, and prevention. *Aggression and Violent Behavior*, *2*(4), 343–353. https://doi.org/10.1016/S1359-1789(97)00020-7

Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C., & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*, *100*(3), 525–534. https://doi.org/10.1037/0022-0663.100.3.525

Carter, P. L., Skiba, R., Arredondo, M. I., & Pollock, M. (2017). You can't fix what you don't look at: Acknowledging race in addressing racial discipline disparities. *Urban Education*, *52*(2), 207–235. https://doi.org/10.1177/0042085916660350

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum.

Cook, C. R., Duong, M. T., McIntosh, K., Fiat, A. E., Larson, M., Pullmann, M. D., & McGinnis, J. (2018). Addressing discipline disparities for black male students: Linking malleable root causes to feasible and effective practices. *School Psychology Review*, *47*(2), 135–152. https://doi.org/10.17105/SPR-2017-0026.V47-2

Cook, T. D., Habib, F. N., Phillips, M., Settersten, R. A., Shagle, S. C., & Degirmencioglu, S. M. (1999). Comer's school development program in Prince George's County, Maryland: A theory-based evaluation. *American Educational Research Journal*, *36*(3), 543–597. https://doi.org/10.3102/00028312036003599

Cook, T. D., Murphy, R. F., & Hunt, H. D. (2000). Comer's School Development Program in Chicago: A theory-based evaluation. *American Educational Research Journal*, *37*(2), 535–597. https://doi.org/10.3102/00028312037002535

Crutchfield, J., Phillippo, K. L., & Frey, A. (2020). Structural racism in schools: A view through the lens of the national school social work practice model. *Children & Schools*, *42*(3), 187–193.

Cuellar, A. E., & Markowitz, S. (2015). School suspension and the school-to-prison pipeline. *International Review of Law and Economics*, *43*, 98–106. https://doi.org/10.1016/j.irle.2015.06.001

Curran, F. C. (2016). Estimating the effect of state zero tolerance laws on exclusionary discipline, racial discipline gaps, and student behavior. *Educational Evaluation and Policy Analysis*, *38*(4), 647–668. https://doi.org/10.3102/0162373716652728

Davis, J. E., & Jordan, W. J. (1994). The effects of school context, structure, and experiences on Black males in middle and high school. *The Journal of Negro Education*, *63*(4), 570–587. https://doi.org/10.2307/2967296

Dodge, K. A., Malone, P. S., Lansford, J. E., Miller, S., Pettit, G. S., & Bates, J. E. (2009). A dynamic cascade model of the development of substance-use onset. *Monographs of the Society for Research in Child Development*, *74*(3), vii–119. https://doi.org/10.1111/j.1540-5834.2009.00528.x

Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*(1), 222–217. https://doi.org/10.1186/2193-1801-2-222

Enders, C. (2010). *Applied missing data analysis*. The Guildford Press.

Fazel, M., Hoagwood, K., Stephan, S., & Ford, T. (2014). Mental health interventions in schools 1: Mental health interventions in schools in high-income countries. *Lancet Psychiatry*, *1*(5), 377–387. https://doi.org/10.1016/S2215-0366(14)70312-8

Fenning, P., & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline the role of school policy. *Urban Education*, *42*(6), 536–559. https://doi.org/10.1177/0042085907305039

Flament, M. F., Cohen, D., Choquet, M., Jeammet, P., & Ledoux, S. (2001). Phenomenology, psychosocial correlates, and treatment seeking in major depression and dysthymia of adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(9), 1070–1078. https://doi.org/10.1097/00004583-200109000-00016

Ford, T., Parker, C., Salim, J., Goodman, R., Logan, S., & Henley, W. (2018). The relationship between exclusion from school and mental health: A secondary analysis of the British Child and Adolescent Mental Health Surveys 2004 and 2007. *Psychological Medicine*, *48*(4), 629–641. https://doi.org/10.1017/S003329171700215X

Gage, N. A., Whitford, D. K., & Katsiyannis, A. (2018). A review of schoolwide positive behavior interventions and supports as a framework for reducing disciplinary exclusions. *The Journal of Special Education*, *52*(3), 142–151. https://doi.org/10.1177/0022466918767847

Gion, C., McIntosh, K., & Falcon, S. (2020). Effects of a multifaceted classroom intervention on racial disproportionality. *School Psychology Review* https://doi.org/10.1080/2372966X.2020.1788906.

Gregory, A., Huang, F. L., Anyon, Y., Greer, E., & Downing, B. (2018). An examination of restorative interventions and racial equity in out-of-school suspensions. *School Psychology Review*, *47*(2), 167–182. https://doi.org/10.17105/SPR-2017-0073.V47-2

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap two sides of the same coin? *Educational Researcher*, *39*(1), 59–68. https://doi.org/10.3102/0013189X09357621

Halberstadt, A. G., Castro, V. L., Chu, Q., Lozada, F. T., & Sims, C. M. (2018). Preservice teachers' racialized emotion recognition, anger bias, and hostility attributions. *Contemporary Educational Psychology*, *54*, 125–138. https://doi.org/10.1016/j.cedpsych.2018.06.004

Harper, K., Ryberg, R., & Temkin, D. (2019). *Black students and students with disabilities remain more likely to receive out-of-school suspensions, despite overall declines*. https://www.childtrends.org/publications/black-students-disabilities-out-of-school-suspensions

Hemphill, K. C., Reinke, W. M., Dong, N., & Bradshaw, C. P. (2020). Can effective classroom behavior management increase student achievement in middle school? Findings from a group randomized trial. *Journal of Educational Psychology*, Advance online publication. https://doi.org/10.1037/edu0000641

Hemphill, S. A., Toumbourou, J. W., Herrenkohl, T. I., McMorris, B. J., & Catalano, R. F. (2006). The effect of school suspensions and arrests on subsequent adolescent antisocial behavior in Australia and the United States. *Journal of*

Adolescent Health, *39*(5), 736–744. https://doi.org/10.1016/j.jadohealth.2006.05.010

Herman, K. C., Reinke, W. M., Dong, N., & Bradshaw, C. P. (2020). Can effective classroom behavior management increase student achievement in middle school? Findings from a group randomized trial. *Journal of Educational Psychology*, Advance online publication. https://doi.org/10.1037/edu0000641

Hetey, R. C., & Eberhardt, J. L. (2014). Racial disparities in incarceration increase acceptance of punitive policies. *Psychological Science*, *25*(10), 1949–1954. https://doi.org/10.1177/0956797614540307

Hinojosa, M. S. (2008). Black-white differences in school suspension: Effect of student beliefs about teachers. *Sociological Spectrum*, *28*(2), 175–193. https://doi.org/10.1080/02732170701796429

Huang, F. L. (2016a). Do Black students misbehave more? Investigating the differential involvement hypothesis and out-of-school suspensions. *The Journal of Educational Research*, *111*(3), 284–211. https://doi.org/10.1080/00220671.2016.1253538

Huang, F. L. (2016b). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education*, *84*(1), 175–196. https://doi.org/10.1080/00220973.2014.952397

Huang, F. L. (2020). Prior problem behaviors do not account for the racial suspension gap. *Educational Researcher*, *49*(7), 493–502. https://doi.org/10.3102/0013189X20932474

Hughes, T., Raines, T., & Malone, C. (2020). School pathways to the juvenile justice system. *Policy Insights from the Behavioral and Brain Sciences*, *7*(1), 72–79. https://doi.org/10.1177/2372732219897093

Institute of Education Sciences. (2017). *What works clearinghouse procedures handbook*. US Department of Education. https://ies.ed.gov/ncee/wwc/docs/referenceresources/wwc_procedures_handbook_v4.pdf.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2009). Teacher observation of classroom adaptation—checklist: Development and factor structure. *Measurement and Evaluation in Counseling and Development*, *42*(1), 15–30. https://doi.org/10.1177/0748175609333560

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lamont, J. H., Devore, C. D., Allison, M., Ancona, R., Barnett, S. E., Gunther, R., Holmes, B., Minier, M., & Okamoto, J. K. (2013). Out-of-school suspension and expulsion. *Pediatrics*, *131*(3), e1000–e1007.

Lansford, J. E., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2016). A public health perspective on school dropout and adult outcomes: A prospective study of risk and protective factors from age 5 to 27 years. *The Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, *58*(6), 652–658. https://doi.org/10.1016/j.jadohealth.2016.01.014

Lee, T., Cornell, D., Gregory, A., & Fan, X. (2011). High suspension schools and dropout rates for black and white students. *Education and Treatment of Children*, *34*(2), 167–192. https://doi.org/10.1353/etc.2011.0014

Losen, D. J., & Whitaker, A. (2017). *Lost instruction: The disparate impact of the School Discipline Gap in California*. Civil Rights Project-Proyecto Derechos Civiles.

Martinez, S. (2009). A system gone berserk: How are zero-tolerance policies really affecting schools?*Preventing School Failure: Alternative Education for Children and Youth*, *53*(3), 153–158. https://doi.org/10.3200/PSFL.53.3.153-158

McIntosh, K., Gion, C., & Bastable, E. (2018). *Do schools implementing SWPBIS have decreased racial and ethnic disproportionality in school discipline?* PBIS Evaluation Brief. OSEP National Technical Assistance Center on Positive Behavioral Interventions and Supports.

McNeely, C. A., Nonnemaker, J. M., & Blum, R. W. (2002). Promoting school connectedness: Evidence from the national longitudinal study of adolescent health. *The Journal of School Health*, *72*(4), 138–146. https://doi.org/10.1111/j.1746-1561.2002.tb06533.x

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

Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly*, *90*(4), 1003–1018.

Noltemeyer, A. L., Ward, R. M., & Mcloughlin, C. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review*, *44*(2), 224–240. https://doi.org/10.17105/spr-14-0008.1

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science*, *26*(5), 617–624. https://doi.org/10.1177/0956797615570365

Pachter, L. M., Bernstein, B. A., Szalacha, L. A., & Coll, C. G. (2010). Perceived racism and discrimination in children and youths: An exploratory study. *Health & Social Work*, *35*(1), 61–69. https://doi.org/10.1093/hsw/35.1.61

Paterson, G., DeBaryshe, B., & Ramsey, E. (1989). A developmental perspective on antisocial behaviour. *American Psychologist*, *44*, 329–335.

Quintana, S., Boykin, A. W., Fuligni, A., Graham, S., Ortiz, S., & Worrell, F. C. (2012). *Ethnic and racial disparities in education: Psychology's contributions to understanding and reducing disparities*. Report to the APA Presidential Task Force on Educational Disparities.

Reid, K. (2005). The causes, views and traits of school absenteeism and truancy: An analytical review. *Research in Education*, *74*(1), 59–82. https://doi.org/10.7227/RIE.74.6

Reinke, W. M., & Walker, H. M. (2006). Deviant peer effects in education. In K. A. Dodge, T. J. Dishion, & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions* (pp. 122–140). The Guilford Press.

Rushton, J. L., Forcier, M., & Schectman, R. M. (2002). Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(2), 199–205. https://doi.org/10.1097/00004583-200202000-00014

Scaramella, L. V., & Leve, L. D. (2004). Clarifying parent-child reciprocities during early childhood: The early childhood coercion model. *Clinical Child and Family Psychology Review*, *7*(2), 89–107. https://doi.org/10.1023/b:ccfp.0000030287.13160.a3

Simson, D. (2013). Exclusion, punishment, racism, and our schools: A critical race theory perspective on school discipline. *UCLA Law Review*, *61*, 506.

Skiba, R. J., Arredondo, M. I., & Williams, N. T. (2014). More than a metaphor: The contribution of exclusionary discipline to a school-to-prison pipeline. *Equity & Excellence in Education*, *47*(4), 546–564.

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*(4), 317–342. https://doi.org/10.1023/A:1021320817372

Skiba, R. J., & Edl, H. (2004). *The disciplinary practices survey: How do Indiana's principals feel about discipline*. Center for Evaluation & Education Policy Web Site. Children Left behind Policy Briefs, Supplementary Analysis.

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, *100*(4), 765–780. https://doi.org/10.1037/a0012840

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, *69*(3), 493–525. https://doi.org/10.1177/0013164408323233

Thomas, O. N., Caldwell, C. H., Faison, N., & Jackson, J. S. (2009). Promoting academic achievement: The role of racial identity in buffering perceptions of teacher discrimination on academic achievement among African American and Caribbean Black adolescents. *Journal of Educational Psychology*, *101*(2), 420–431. https://doi.org/10.1037/a0014578

Toldson, I. A., McGee, T., & Lemmons, B. P. (2015). Reducing suspensions by improving academic engagement among school-age Black males. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 107–117). Teachers College Press.

US Department of Education. (2018). *2015-16 civil rights data collection: School climate and safety*. US Government Printing Office.

United States Government Accountability Office. (2016). *Better use of information could help agencies identify disparities and address racial discrimination* (GAO Publication No. GAO-16-345). https://www.gao.gov/assets/680/676745.pdf.

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology*, *31*(2), 125–141. https://doi.org/10.1016/j.cedpsych.2005.03.002

Way, S. M. (2011). School discipline and disruptive classroom behavior: The moderating effects of student perceptions. *The Sociological Quarterly*, *52*(3), 346–375. https://doi.org/10.1111/j.1533-8525.2011.01210.x

Welner, K. G., & Carter, P. L. (2013). Achievement gaps arise from opportunity gaps. In P. L. Carter & K. G. Welner (Eds.) *Closing the Opportunity Gap: What America Must Do to Give Every Child an Even Chance* (pp. 1–10). Oxford University Press.

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

AUTHOR BIOGRAPHICAL STATEMENTS

Daniel R. Cohen, Ph.D., M.P.H., is an assistant professor of school psychology at the University of Alabama and the Codirector of the Center for Interconnected Behavioral and Mental Health Systems. He completed his doctoral degree in school psychology at the University of Missouri and his master of public health degree at Johns Hopkins University. His research interests include educational equity, school mental health, and health disparities.

Crystal Lewis has a master's degree in public policy from the University of Minnesota, Twin Cities. She is currently Director of Data Strategy at the Missouri Prevention Science Institute and focuses her time on building capacity around reproducible research, data stewardship, and creating products to integrate research and practice.

Colleen L. Eddy, M.A., is a sixth-year doctoral candidate in the Counseling Psychology program at the University of Missouri, Columbia. She is a predoctoral intern with Devereux Children's Behavioral Health Services. Her research interests are reducing stress and improving coping for teachers and systemic interventions for youth mental health.

Lauren Henry, Ph.D., is a Pediatric Mental Health Institute (PMHI) Psychology Fellow at Children's Hospital Colorado. She completed her Doctorate in School Psychology at the University of Missouri- Columbia. Her areas of research and clinical interest include prevention and treatment of child emotional and behavioral disorders, child motivation, behavioral consultation models, and systemic levels of support for youth mental health.

Caroline Hodgson, Ph.D., is a Pediatric Acute Treatment Psychology Fellow at Nationwide Children's Hospital. She received her doctoral degree in school psychology at the University of Missouri. Her research focuses on harnessing innovative measurement strategies and intervention modalities to identify and support youth with internalizing difficulties and enhance the science of youth suicide prevention.

Francis L. Huang, Ph.D., is an associate professor in the Statistics, Measurement, and Evaluation in Education program in the College of Education at the University of Missouri and the Methodology Director of the Missouri Prevention Science Institute. His research focuses on both methodological (e.g., analysis of clustered data) and substantive (e.g., school climate, bullying, disparities in disciplinary sanctions) areas of interest.

Wendy M. Reinke, Ph.D., is a Professor in the Department of Educational, School, and Counseling Psychology at the University of Missouri and codirector of the Missouri Prevention Science Institute and PI on the National Center for Rural School Mental Health. She has special interest in the prevention and treatment of youth mental health and implementation of evidence-based interventions (EBIs) in communities and schools. She has been PI or Co-PI on over $30 million in federal research grants. Further, she is currently the lead investigator of a six school district-wide mental health project that has developed a web-based assessment and reporting system to identify students at risk and provide appropriate supports. She has coauthored seven books and 11 chapters related to prevention of social emotional and behavior problems in youth and over 90 peer-reviewed publications.

Keith Herman, Ph.D., is a Codirector of the Missouri Prevention Science Institute and a Distinguished Curator's Professor of school and counseling psychology at the University of Missouri. He has authored 120 peer-reviewed publications and four books on topics related to promoting youth mental health and well-being.