



Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

A multilevel analysis of racial discipline disproportionality: A focus on student perceptions of academic engagement and disciplinary environment

Kristine E. Larson^{a,*}, Jessika H. Bottiani^b, Elise T. Pas^c, Joseph M. Kush^b,
Catherine P. Bradshaw^{b,c}

^a Notre Dame of Maryland University, USA

^b University of Virginia, Curry School of Education and Human Development, USA

^c Johns Hopkins University, Bloomberg School of Public Health, USA

ARTICLE INFO

Keywords:

Racial discipline disproportionality
Academic engagement
Disciplinary environment
School climate
Positive behavior support
Multilevel modeling

ABSTRACT

Excessive use of exclusionary school discipline with Black students is a persistent, systemic problem in U.S. schools with potential to affect students' perceptions of their school. For example, students may notice racial differences in out-of-school suspensions, which may relate to how academically engaged they feel and the extent to which they view the school's disciplinary environment as positive. The current study investigated school-level racial discipline disproportionality and observed classroom-level, positive behavior supports in relation to student perceptions of academic engagement and school disciplinary environment by fitting a series of three-level models, which included data on students ($N = 17,115$), classrooms ($J = 310$), and schools ($K = 53$). Two metrics of discipline disproportionality were used (i.e., the risk ratio and the risk difference) and moderation was examined through cross-level interactions. Results indicated that, regardless of race, students perceived the disciplinary environment as significantly less favorable in schools with greater racial discipline disproportionality when measured by the risk ratio, but not when measured by the risk difference. Using different disproportionality metrics in education research has important implications for policies and practices to identify and address the issue. How discipline disparities relate to the way that students perceive the disciplinary environment will likely inform intervention efforts for school psychologists.

Racial discipline disproportionality occurs when there are disparities in exclusionary discipline (e.g., out-of-school suspensions) between racial and ethnic groups. Despite increased concerns over the last decade, these disparities have persisted in schools across the United States (Losen & Gillespie, 2012; Losen & Martinez, 2013; U.S. Department of Education, 2018). For example, Black students, on average, are almost three times as likely to be suspended as White students (Skiba, Mediratta & Rausch, 2016; U.S. Department of Education, 2018). The use of exclusionary discipline has been associated with poor academic, behavioral, civic, and vocational outcomes (Fabelo et al., 2011; Hemphill, Toumbourou, Herrenkohl, McMorris, & Catalano, 2006; Kupchik & Catlaw, 2015; Noltemeyer, Marie, McLoughlin, & Vanderwood, 2015; Peguero & Bracy, 2015; Wald & Losen, 2003).

Schools' use of exclusionary and racially disproportionate discipline may also exert a deleterious effect for all students, regardless of whether they have been disciplined themselves, via its impact on school climate. Schools with positive climates demonstrate higher

* Corresponding author.

E-mail address: klarson2@ndm.edu (K.E. Larson).

<https://doi.org/10.1016/j.jsp.2019.09.003>

Received 11 June 2018; Received in revised form 10 July 2019; Accepted 12 September 2019

Available online 26 November 2019

0022-4405/ © 2019 Published by Elsevier Ltd on behalf of Society for the Study of School Psychology.

student attendance, academic achievement, and emotional health (Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013). Thus, it is critical to understand how school discipline practices, and racially disproportionate discipline in particular, relate to school climate. Prior research has shown that racially disproportionate school discipline is negatively associated with teacher-student relationship-focused dimensions of school climate. For example, research has shown that in schools characterized by higher racial discipline disproportionality, Black students have more negative perceptions of perceived fair treatment by adults at school (Bottiani, Bradshaw, & Mendelson, 2017). Greater racial discipline disproportionality in schools has similarly been linked to more negative perceptions of connectedness to adults among all students (Anyon, Zhang, & Hazel, 2016).

Given that school climate is a multi-faceted construct, it is important to further explore the extent to which racial discipline disproportionality is associated with other dimensions of climate, such as the disciplinary and academic environment of the school and classroom. To date, there has been little quantitative research examining how racial discipline disproportionality relates to students' views of their schools' discipline practices or academic engagement strategies; the previously mentioned studies (i.e., Anyon et al., 2016; Bottiani et al., 2017) did not examine these dimensions of climate, as they focused on student perceptions of equity, connectedness, school belonging, and adjustment. Student views of school discipline practices may inform efforts to prevent negative student behavior. For example, positive perceptions of the school disciplinary environment were related to lower rates of disruptions (Way, 2011). On the other hand, more school rules, higher perceived strictness, and more severe consequences have been linked with more disruptive behavior and defiance among students (Way, 2011). Further, students' connection to adults at school and sense of school belonging have been shown to be associated with racial discipline disproportionality (Bottiani et al., 2017), and since both factors have also been shown to be associated with academic engagement (Reyes, Brackett, Rivers, White, & Salovey, 2012), it is important to also explore how discipline disproportionality relates to student academic engagement.

The role classroom behavior support in these associations has also been overlooked in the literature examining the association between racial discipline disproportionality and school climate. This gap in the literature is noteworthy, as research indicates the classroom specifically may be the context in which the highest rates of exclusionary and disproportionate discipline transpire (Gion, McIntosh, & Horner, 2014; Smolkowski, Girvan, McIntosh, Nese, & Horner, 2016). Since students spend most of their time in classrooms during the school day and classrooms are where discipline factors often first arise, teacher practices may relate to student perceptions of the school disciplinary environment and academic engagement (Bradshaw, Pas, et al., 2018; Cook et al., 2018; Mitchell & Bradshaw, 2013) as well as to racial discipline disproportionality (Gion et al., 2014; Skiba, Chung, Trachok, & Hughes, 2014). Teachers' use of positive behavior supports in the classroom specifically is key to consider as it is an important predictor of equity in school discipline (Vincent & Tobin, 2011) and likely relates to the school-level disciplinary environment. As such, the current study leveraged a rigorous multilevel modeling approach to examine how student perceptions of the disciplinary environment and academic engagement are related to student race, classroom positive behavior support, and school-level racial discipline disproportionality, utilizing two commonly-used metrics of disproportionality (risk ratio and risk difference; Petrosino, Fronius, Goold, Losen, & Turner, 2017). As a secondary aim, we examined cross-level interactions of disproportionality with student race and classroom use of positive behavior supports. The overarching goal of this study was to deepen our understanding of the potential associations between classroom and school discipline practices and students' perceptions of discipline and engagement; this particular line of research is best achieved through a multilevel modeling approach, which allows for the simultaneous inclusion of these variables across the student, classroom, and school levels. As such, this line of research has important potential implications for future intervention research and practice in school psychology to address the discipline gap as well as policies relating to how we define, measure, and reduce disproportionality.

1. School climate perceptions: academic engagement and disciplinary environment

School climate refers to the quality and character of school life based on patterns of students', parents', and school personnel's experience of school life (National School Climate Center, 2019) and is often measured using self-reported surveys (Wang & Degol, 2016). One key domain of school climate is engagement, a multidimensional construct that includes a focus on academic engagement (Bradshaw, Waasdorp, Debnam, & Lindstrom Johnson, 2014b; U.S. Department of Education, 2009). Academic engagement reflects students' beliefs that they can do well in school, that their teachers want them to do their best, and that it is important to finish high school (Bradshaw et al., 2014a,b). Academic engagement has been linked to school completion (Li & Lerner, 2011; Wang & Fredricks, 2014) and thus of great interest, particularly in high schools. Academic engagement is also interconnected with behavior and thus school discipline, as less engaged students display conduct problems and violate school rules, often leading to discipline consequences (e.g., suspension; Bradshaw, O'Brennan, & McNeely, 2008; Carter, McGee, Taylor, & Williams, 2007). Yet many high schools struggle to engage students (National Research Council, 2003) and increasing student diversity may amplify this challenge (Pas, Cash, et al., 2015). Engagement may also be particularly important for low-income and marginalized students (Shernoff & Schmidt, 2008). Academic engagement has been found to be a protective factor for African American youth against involvement in violence (Voisin & Elsaesser, 2016), and given that decreased engagement at school and subsequent dropout have been linked to exclusionary discipline practices (Fabelo et al., 2011), it is critical to better understand how racial discipline disproportionality may relate to Black youth's academic engagement in school.

Student perceptions of their school's disciplinary environment is another aspect of climate that also may uniquely contribute to our understanding of important mechanisms to address and improve issues of racial discipline disproportionality. In providing their perceptions of the disciplinary environment, students may shed light upon inconsistency in their understanding of school personnel response to rules and suggest areas for school improvement. Extant research on school-wide positive behavior supports has demonstrated that improving the disciplinary environment with regard to the use of data, systems, and practices that promote positive

behavior is associated with reductions in the use of exclusionary discipline (e.g., office discipline referrals, in-school suspensions, and out-of-school suspensions; Childs, Kincaid, George, & Gage, 2016; Gage, Grasley-Boy, George, Childs, & Kincaid, 2019; Pas, Ryoo, Musci, & Bradshaw, 2019), but is more mixed about how it impacts racial disproportionality specifically (Bradshaw, Waasdorp, & Leaf, 2012; Vincent, Swain-Bradway, Tobin, & May 2011; Vincent & Tobin, 2011). However, a systematic literature review by Johnson (2009) found that students were more aware of school rules and perceived the rules to be fairer in schools with less violence. Moreover, in schools that were less violent, students had more positive relationships with their teachers and felt that the environment was orderly. Thus, perceptions of the school disciplinary environment may serve as correlates or a proxy for other outcomes of interest and may demonstrate racial discrepancies important to racial discipline disproportionality.

1.1. Racial discipline disproportionality

Racial discipline disproportionality, also called the *discipline gap* or *discipline disparities*, refers to the excess impact of exclusionary discipline on Black, Latinx, and indigenous/American Indian student groups (Skiba et al., 2016). Although Asian students are the least likely groups to receive school discipline of any kind (Morgan & Wright, 2018), research since the 1970s has mostly focused on the gap between Black and White students because of the prevalence of White cultural norms in school behavioral expectations (Girvan, 2019; McIntosh, Girvan, Horner, & Smolkowski, 2014). Specifically, the percent of Black students receiving suspensions grew by 10% between 1973 and 2010, but only increased by 2% for White students (Losen et al., 2015). Research has shown that being suspended from school is associated with a higher risk of delinquency (Ganao, Silvestre, & Glen, 2013), student dropout (Peguero & Bracy, 2015), and subsequent entry into the juvenile justice system (Fabelo et al., 2011; Pesta, 2018). The process by which school exclusion may push students into the justice system is commonly referred to as the “school-to-prison pipeline” (Mizel et al., 2016; Wald & Losen, 2003). Scholars theorize that the trajectory of the school-to-prison pipeline for Black youth may be interrupted if school exclusion of Black youth as a discipline practice is reduced (Coggs, Osher, & Columbi, 2013).

Although there is limited research on the specific causes of racial discipline disproportionality, racial bias and lack of cultural sensitivity and awareness of school staff (e.g., school-level administrators and classroom teachers) have been shown to escalate disciplinary encounters (Girvan, Gion, McIntosh, & Smolkowski, 2017; Kirk, 2009; Okonofua, Walton, & Eberhardt, 2016; Skiba, Chung, Trachok, Baker, & Hughes, 2014). Research suggests that relational rule violations or “soft offenses” (e.g., disrespect, insubordination, and defiance), rather than objective school rule violations (e.g., substance use on school property), may be a primary driver of excessive disciplinary action with Black students (Heilbrun, Cornell, & Lovegrove, 2015; Skiba et al., 2002). This research highlights the need for considering cultural context around expectations and behavior (i.e., how rules and violations are defined and subsequent disciplinary actions; Bondy, Ross, Galligane, & Hambacher, 2007).

1.2. Teachers' use of positive behavior supports in the classroom

Although empirical research on the impact of school-wide applications of positive behavior supports on racial discipline disproportionality is mixed (Bradshaw et al., 2012; Gage, Grasley-Boy, Peshak, George, Childs, & Kincaid, 2019; Vincent et al., 2011; Vincent & Tobin, 2011), teachers' student behavior support is an important factor in both student engagement and discipline outcomes. For example, having a teacher with an authoritative discipline style (i.e., balancing warmth and demandingness) is associated with lower student suspension risk and lower racial discipline disproportionality (Heilbrun, Cornell, & Konold, 2018; Huang & Cornell, 2018). Similarly, adults' active efforts to connect with students in their classrooms have been linked with more positive student behavior (Cash, Debnam, Waasdorp, Wahl, & Bradshaw, 2019; Cook et al., 2018). Actively communicating with students (e.g., implementation of the Greet-Stop-Prompt intervention; Cook et al., 2018) is associated with decreases in office referrals of Black male students. Furthermore, theory and empirical research suggests that Black students benefit from such warm-demanding relational styles (Ford & Sassi, 2014; Sandilos, Rimm-Kaufman, & Cohen, 2017). There is also a growing body of research highlighting the importance of specific teacher strategies, such as praising students and providing positive feedback to engage them, in connection with more positive student behavior (e.g., Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008; Sutherland, Wehby, & Copeland, 2000). Taken together, these findings suggest the promise of teachers' use of positive behavior supports in the classroom as a potential target for improving student perceptions of the school disciplinary environment and engagement.

1.3. Current study

Although a few multilevel studies have examined racial discipline disproportionality as a school-level factor associated with students' school climate perceptions (Anyon et al., 2016; Bottiani et al., 2017), prior studies have rarely examined academic engagement and positive disciplinary environment perceptions as specific outcomes of interest or accounted for teachers' use of positive behavior supports in the classroom simultaneously with school contextual factors. Engagement and perceptions of the environment are likely associated both with individual and ecological factors; thus, the utilization of multilevel modeling is an important method to capture the variability and associations of factors at multiple levels. Thus, the primary aim of this study was to determine whether: 1) racial disproportionality was associated with student perceptions of academic engagement and disciplinary environment, above and beyond positive behavior supports in the classroom; 2) teachers' use of positive behavior supports in the classroom was related to student perceptions of academic engagement and disciplinary environment, regardless of racial disproportionality; and 3) student perceptions of academic engagement and disciplinary environment differed among Black and White students. As a secondary aim, we were also interested in whether racial disproportionality moderated the association between the two outcomes (i.e., academic

engagement and disciplinary environment) and teacher's use of positive behavior support in the classroom and student race (i.e., racial disparities). We leveraged two disproportionality metrics (Petrosino et al., 2017), which is a novel contribution to the field.

A previous study by Bottiani et al. (2017) examined how school-level disproportionality related to students' perceptions of school equity, belonging, and externalizing problems. With the exception of the study by Bottiani et al. (2017), much of the disproportionality research has examined this construct as an outcome; however, we believe it is equally important to understand the potential impact of attending a school with racial disproportionality in exclusionary discipline. We extended the prior work by Bottiani et al. (2017) using data from the same dataset, but collected at a different time point and leveraging other data sources (e.g., classroom observations) and exploring additional facets of school climate and disproportionality. The classroom observational data allowed for the examination of positive behavior supports and added an additional (classroom) level to the multilevel modeling. In addition, this study included student-perceived academic engagement and school disciplinary environment as key outcomes of interest as compared to the earlier study by Bottiani et al. (2017) which examined belonging, equity, and student adjustment problems as outcomes. Another novel aspect of the current study is that we examined two disproportionality metrics (i.e., risk ratio and risk difference) to determine whether different indicators of disproportionality yield different findings (Girvan, McIntosh, & Smolkowski, 2019). Such a finding would have important implications for future research on disproportionality, and further inform policy and practice efforts aimed at identifying disproportionate schools with the goal of reducing racial discipline disproportionality. Specifically, we hypothesized that racial disproportionality would be negatively associated with student perceptions of academic engagement and the disciplinary environment, above and beyond positive behavior supports in the classroom, and that teachers' use of positive behavior supports in the classroom would be positively associated with student perceptions of academic engagement and disciplinary environment, regardless of racial disproportionality (Koth, Bradshaw, & Leaf, 2008). We expected that Black students' perceptions of the disciplinary environment would be less favorable than White students' perceptions overall.

To address the secondary aim, we incorporated cross-level interactions into our models to advance prior work linking disproportionality with less favorable perceptions of student connectedness to adults, sense of belonging, and fair treatment, particularly among Black students (Anyon et al., 2016; Bottiani et al., 2017). We hypothesized that racial discipline disproportionality would be more negatively associated with Black students' perceptions of these two constructs relative to their White peers' perceptions. Similarly, using cross-level interactions, we further examined whether the associations of racial discipline disproportionality and student-perceived academic engagement and disciplinary environment differed for Black students (relative to White students) and for students in classrooms where teachers used higher proportions of positive behavior supports.

2. Method

2.1. Participants

These data come from the Maryland Safe and Supportive Schools (MDS3) Project (see Bradshaw, Debnam, et al., 2014a). We restricted the sample specifically to include students who identified themselves as White or Black given the disparities historically found between these two groups, as well as researchers' recommendation to use "White" as the reference group (Skiba et al., 2011). Moreover, these two groups comprised 77% of the total sample and to optimize our analysis of disparities in student perceptions between Black and White students. Therefore, survey data from 17,115 students ($n = 6332$, 37% Black; $n = 10,782$, 63% White) in 310 classrooms and 53 high schools across 12 districts (i.e., half of the districts) in Maryland participating in a project focused on measuring and improving school climate were analyzed. Data were collected using a web-based survey from all students in grades 9–12 completed during the spring of 2014 (see Bradshaw et al., 2014b for details). The average school enrollment in the 53 schools

Table 1
Student and school demographic characteristics.

Student Characteristics ($N = 17,115$)	N (%)
Gender	
Male	8551 (50)
Race/Ethnicity	
Black/African American	6332 (37)
White/Caucasian	10,782 (63)
Classroom Characteristics ($N = 310$)	M (SD)
Positive Behavior Support	.66 (.28)
Percent of Males Present	51 (.17)
Percent of Students of Color Present	48 (.29)
School Characteristics ($N = 53$)	M (SD)
Total School Enrollment	1269.30 (461.99)
% Free and Reduced Meals (FARMS)	38.29 (17.84)
Risk Ratio for Black Students	3.47 (2.46)
% Suspension	15.93 (11.72)

was 1269.30 ($SD = 461.99$) with the percentage of students of color being 48.26% ($SD = 24.89$), and 38.29% ($SD = 17.84$) receiving free and reduced-price meals (FARMS). On average, the school-level suspension rate was 15.93% ($SD = 11.72$; See Table 1).

2.2. Procedure

School districts in Maryland were approached by the state's department of education to participate in the MDS3 Project (see Bradshaw, Debnam, et al., 2014b). After attending informal meetings that provided an overview of the project, interested principals signed commitment forms. A waiver of active parental consent and youth assent process was used to collect the non-identifiable student survey data. Using a standardized written procedure, school staff administered the anonymous, online self-report survey to students across approximately 25 classrooms in each school. Each student entered a classroom-identifying password to access the survey, which allowed for a link between student surveys and classroom observations. Surveys were administered by a specific content area throughout the school (mostly language arts) to ensure that the same student did not respond twice to the survey. In addition to the survey, approximately 25 classrooms in each school were also observed by trained observers (see details below). Language arts classes were prioritized for observations, followed by other core instructional classes (e.g., math, science, and social studies) at random, until the goal of observing 25 different teachers/classrooms per school was met. We prioritized core subject areas and began with language arts because it was required of all students, thus maximizing our catchment of the student enrollment. It also aligned this data source to the student survey collection approach. The collection and analysis of these data were approved by the Institutional Review Board at the researchers' institutions.

2.3. Measures

MDS3 Climate Survey. Researchers at the Johns Hopkins Center for the Prevention of Youth Violence and project partners developed the MDS3 School Climate Survey (Bradshaw et al., 2014b) to assess safety, engagement, and environment; the three domains of school climate proposed by the U.S. Department of Education. Prior psychometric work on the MDS3 School Climate Survey confirmed the factor structure as well as measurement invariance across various subgroups (e.g., gender, race/ethnicity, grade level; Lindstrom Johnson, Reichenberg, Waasdorp, Shukla, & Bradshaw, 2018; Shukla et al., 2019, in press). In the current study, we focused on the student perceptions of academic engagement and disciplinary environment aspects of this survey, which are further described below.

Academic engagement. Student perceptions of academic engagement were measured using the four items: “My teachers believe that I can do well in school,” “I believe that I can do well in school,” “My teachers always want me to do my best,” and “It is important to finish high school” (Cronbach's alpha [α] for the current sample was .78); α -values above 0.70 are considered acceptable (Cortina, 1993). Items were rated on a 4-point Likert-type scale from 1 for ‘Strongly Disagree’ to 4 for ‘Strongly Agree’, and averaged, such that higher scores indicated a higher level of student academic engagement.

Disciplinary environment. This five-item scale measured student perceptions of the school disciplinary environment, which included “At this school, students listen to the teachers,” “At this school, teachers can handle students who disrupt class,” “At this school, there are clear rules about student behavior,” “At this school, everyone knows what the school rules are,” and “At this school, students are rewarded for positive behavior” (5-item α from this sample = .74). Items were rated on a 4-point Likert-type scale from 1 for ‘Strongly Disagree’ to 4 for ‘Strongly Agree’, and averaged, such that higher scores indicated a more positive school disciplinary environment.

Student demographics. Students identified themselves by several demographic characteristics (e.g., race, gender, and achievement level). Given our interest in contrasting White and Black student perspectives, we restricted the analysis sample to only include students who self-identified as White or Black (coded 0 = White, 1 = Black); all other students were excluded. Youth also self-reported their gender (coded 0 = female, 1 = male). Students were asked their average grades on an ordinal grading scale from A to F (i.e., “I receive mostly ... As, Bs, Cs, Ds, or Fs”). Given the correlations between student achievement and engagement, a binary variable “Receives a B or higher” was created as a control variable to indicate above average achievement (i.e., Mostly As and Bs responses were coded 1 and Mostly Cs, Ds, and Fs were grouped and coded 0), which also enabled us to account for the skewed data on this ordinal scale. As a result, this binary achievement variable enabled for more meaningful interpretation of the coefficients (i.e., as compared to treating the variable as an ordinal or dummy-coded variable); further, the difference between students earning one letter grade higher than another may not have as much meaning as “students earning higher/above average grades” and students earning “average/below average grades”.

ASSIST Classroom Observation. Trained, independent observers assessed the classroom using the *Assessing School Settings: Interactions of Students and Teachers* (ASSIST; Rusby, Taylor, & Milchak, 2001), an observational measure that records tallies of specific teacher and student behaviors as well as global ratings of classroom social processes during a 15-min observation period. For the purpose of this study, tallies of teacher behaviors and student compositional data were analyzed. Each observer was required to reach an average of 80% inter-observer agreement with a master trainer across all teacher and student behavior tallies, during three classrooms observations prior to observing independently in study schools. The average training inter-observer agreement rate for these tallies was 87% when these data were collected in the spring of 2014. Inter-observer agreement rates were examined again during active data collection and also averaged 87% (see Pas, Cash, O'Brennan, Debnam, & Bradshaw, 2015 for further description of the ASSIST training and reliability).

Positive behavior support. This construct was a proportion variable indicating the proportion of positive behavior support verbalizations. This was generated by adding the *proactive behavior management* (e.g., teacher explained, reminded, commanded,

prompted, or modeled expected behavior) and *approvals* (e.g., teacher used verbal praise or made positive gestures like a “thumbs up” or a pat on the back) tallies as the numerator. The denominator added these proactive/positive tallies to reactive/negative tallies, which included *reactive classroom behavior management* (e.g., touch, gesture, proximity, verbal response to redirect inappropriate student behavior) and *disapproval* (e.g., teacher threatening a punishment for a behavior or using verbal criticism or other negative reactions to student behavior). Similar to studies that compare teacher's positive praise and reprimands (Reinke, Herman, & Newcomer, 2016), the positive tallies were divided by all positive and negative teacher communication tallies to calculate the proportion of positive behavior support, which was included as a classroom-level predictor. Intraclass correlation coefficients were calculated as a measure of reliability for each tally and were reported as high for proactive behavior management and approvals, moderate for reactive behavior management, and low for disapprovals given the high frequency of zero occurrences (see Pas et al., 2015).

Classroom contextual variables. Observers also tallied the total number of students in the classroom, as well as the number of students they observed to be male and White, separately. The percentage of students observed to be male in the class was calculated by dividing the number of students counted as male by the total number of students in the class. Similarly, to calculate the percent of students of color in the classroom, the total number of White students observed in the classroom was divided by the total number of students in the class. This percent was then subtracted from one to arrive at its inverse (e.g., one minus the proportion of White students in the classroom was equal the proportion of students of color in the classroom). Although operationalizing race using this method drew on independent observers' perceptions of students' visible racial position within the classroom based on phenotypic traits, this measure does not purport to assess *actual* student race/ethnicities. Instead, this approach conceptualizes race and racism as endemic and a form of marginalization of people of color based on phenotypic traits; consistent with critical race theory (Solórzano, 1997), it captures perceived student diversity. As such, the method used in the current study to identify percent of students of color in the classroom is consistent with extant research examining perceptions of race based on phenotypic skin color ascribed to individuals by others (Perreira, Wassink, & Harris, 2018).

Exclusionary discipline use. School-level discipline data were obtained from the Maryland State Department of Education, which annually collected school suspension data in the form of the suspension rate (i.e., total number of suspension incidents divided by total number of students enrolled at the school).

Racial discipline disproportionality. Racial discipline disproportionality has been measured using various methods (Petrosino et al., 2017). The measurement of racial discipline disproportionality in research, practice, and policy contexts has been debated and no one metric has emerged as the best indicator (Nishioka, Shigeoka, & Lolich, 2017). Rather, consideration of multiple metrics is recommended to best understand both practical and substantive aspects of racial discipline disproportionality (Girvan et al., 2019). For example, the *risk ratio* (also called the relative risk ratio) is frequently used given its face validity as an intuitive reflection of parity or disparity (1.0 indicates the same risk for both groups, 2.0 indicates a risk that is twice as high for Black students as White students, and so on). Although useful, risk ratios do not provide information about the absolute magnitude of the difference. According to Girvan et al. (2019), “schools with the same risk ratio can have very different overall levels of discipline” (p. 44). For example, a school with a risk ratio of 3.0 might have a risk among Black students of 30% and White students of 10%, or a risk among Black students of 3% and a risk among White students of 1%. In either case, the risk ratio would be the same (i.e., Black students' risk is three times that of White students). Determining the *risk difference*, or the absolute difference in risk on a percentage scale from 1 to 100, allows for more substantive assessment of the difference in the magnitude of the risk. In the prior example, the risk difference in the first scenario would be 20% (30% minus 10% = 20%) and in the second would be 2% (3% minus 1% = 2%), which has very different implications for intervention. Using the two measures in tandem provides a more well-rounded picture of disproportionality and its potential impact on school climate, as well as providing a more holistic view of a schools' disproportionality, which is necessary to determine which schools to prioritize for reform and intervention. Based on work by Girvan et al. (2019) and several other scholars in this area (e.g., Bottiani, Bradshaw, & Gregory, 2018; Nishioka et al., 2017), we opted to examine these two metrics of disproportionality rather than rely solely on one indicator.

The risk ratio and a risk difference were both based on the count of students, disaggregated by race, who received one or more out-of-school suspensions for each school in the sample; these data were obtained from the Office of Civil Rights data files for the school year corresponding with the survey data collection (i.e., 2013–2014). The ratio was calculated for each school by dividing the out-of-school suspension rate for Black students by the out-of-school suspension rate for White students (Petrosino et al., 2017), such that risk ratios exceeding 1.0 indicated higher suspension rates among Black students and risk ratios lower than 1.0 indicated higher suspension rates among White students. The risk difference was calculated for each school by subtracting the out-of-school suspension rate for White students from the out-of-school suspension rate for Black students, such that positive risk differences (i.e., greater than zero) indicated higher suspension rates among Black students whereas negative risk differences indicated higher suspension rates among White students.

School demographics. Demographic variables at the school level were obtained from the MSDE. These variables included the percent of students receiving free and reduced-price meals (FARMS) and the total enrollment of students in each school. Although school-level student race data were also available, these data were too highly correlated with the FARMS rate data for both to be included in the statistical models and race was already captured at the individual (student race), classroom (racial composition), and school (racial discipline disproportionality) levels.

2.4. Analyses

To account for the nested nature of the data, a three-level linear model was fit using Stata software (14.2; StataCorp, 2015) where

student variables were included at level 1, classroom variables were included at level 2, and school variables were included at level 3. Level 1 included student demographics (e.g., race, gender, and high academic achievement as measured by grades of B or above). Multilevel modeling allowed for the investigation of effects at both an individual level and two group (classroom and school) levels, while appropriately attributing unexplained variability at different levels. The proportion of teacher positive behavior support, percent of male students present, and the percent of students of color present were analyzed at level 2. Included at level 3 were the school percentage of students who received FARMs, the total enrollment in the school, schools' overall suspension rates, and Black-White risk ratios and risk differences of suspension. In total, eight models were analyzed. First, four models examined the *direct* effect of the associations of the variables of interest (i.e., student race, classroom positive behavior supports, and school racial discipline disproportionality) and control variables (e.g., classroom and school compositional variables in all models) on two distinct outcomes of interest (i.e., student perceptions of academic engagement and school disciplinary environment) and modeled with two metrics of disproportionality separately. Additionally, four models included *cross-level interaction effects* examining racial discipline disproportionality (two models, one for each metric) as a moderating factor in the association between student race and positive behavior support and the two outcomes (e.g., student perceptions of academic engagement and school disciplinary environment). Standardized coefficients were generated and can be interpreted as effect sizes, in which an effect of 0.20 is considered small (Cohen, 1992).

Missingness. Missingness of outcome data was relatively low, with only 5.4% ($n = 1167$) of students not responding on the survey for the outcome of academic engagement and 8% ($n = 1725$) of students not responding to questions regarding the disciplinary environment. At level 2, between 17 and 18% of data were missing from classroom observations. And at level 3, less than 2% of data were missing. Patterns of missingness in the data were further probed. Little's (1988) multivariate test of Missing Completely at Random (MCAR) indicated the data did not meet the assumptions of the MCAR missing data mechanism ($\chi^2_{(207)} = 6788.66, p\text{-value} < .001$). To mitigate the potential of bias in parameter estimates due to missing data, multiple imputation techniques were used. Specifically, multiple imputation was used to generate multiple datasets, each with plausible values filled in for the originally-missing values. The complete datasets were then analyzed, with estimates from each model pooled into a single set of results (see Enders, 2010 for a more thorough description). For all imputation models, the variables used were: academic engagement, disciplinary environment, male, Black, B or higher, teacher positive support, percent of male students in classroom, percent of students of color in classroom, percent of FARMs in school, percent suspended in school, and total enrollment of school. For all risk ratio models, the risk ratio was included, and for all risk difference models, the risk difference was included. Further, for cross-level interaction models, the appropriate interactions were included. Using the Blimp software (see Enders, Keller, & Levy, 2017; Keller & Enders, 2018), a fully conditional specification (i.e., chained equations) approach was used to impute values for the two fully unconditional three-level models, as well as the four direct effect three-level models. For each three-level model including cross-level effects and random slopes, a respective substantive model compatible fully conditional specification approach was used (SMC-FCS; Bartlett, Seaman, White, & Carpenter, 2015). This approach differs from the standard fully conditional specification approach, in that the outcome variable is imputed conditional on predictors and auxiliary variables. Both approaches, however, appropriately address the nested structure of the data. Specifically, all predictor variables to be used in the analyses, as well as significantly correlated auxiliary variables, were included in each imputation model. A total of $m = 20$ replicated data sets were produced (Enders, 2010; Rubin, 1987). Finally, analyses were run on each imputed dataset, with final parameter and standard error estimates pooled from each model. This approach produces less biased estimates than listwise deletion (Rubin, 1987).

Centering. The student self-report variable *Black* was group-mean centered relative to the level 2 classroom group mean, and observed classroom positive behavior support was group-mean centered relative to level 3 school group mean. This approach allows for an interpretable estimate of the cross-level interaction, by appropriately partitioning the within- and between-group effects. The *percent of males present* and the *percent of students of color present* were centered around the grand mean of level 3 schools, as these were main effects only. At level 3, suspension risk ratio and risk difference for Black students, overall suspension rate, the percent of students receiving FARMs, and total enrollment of the school were also centered around the grand mean.

Intraclass correlation coefficient (ICC). Improvements between the unconditional and final models were assessed. The ICC at the school level for the fully unconditional academic engagement model was 0.02, while the ICC at the classroom level was 0.03. For the fully unconditional disciplinary environment model, the school-level ICC was 0.01, while the ICC at the classroom level was 0.03. These estimates can be interpreted as the proportion of variance in the outcome than can be explained solely by the school a student attends, or the classroom within a school that a student attends, respectively. The percent between-school variance explained as well as the percent between-classroom variance explained were also calculated (see Tables 2 and 3). Estimates of standard errors have been shown to be biased in studies ignoring the clustering effect, and therefore violating the independence assumption (Kish, 1965; Raudenbush & Bryk, 2002; Tate & Wongbundhit, 1983), even when the estimated intraclass correlation is small (e.g., 0.02; Fitzmaurice, Laird, & Ware, 2004)). Further, Hox (1998) argued for the use of multilevel models in scenarios in which data arise from a natural hierarchical structure, such as pupils nested within schools. He states "standard statistical formulas will underestimate the sampling variance ... [unless] either the intraclass correlation is zero, or the cluster size is one.

3. Results

3.1. Academic engagement and risk ratio

Stata results for the direct effects models using risk ratio as the racial discipline disproportionality metric (see model 1a in Table 2) indicated there were no significant school-level predictors. At level 2, the percent of students of color in the classroom was

Table 2

Results for 3-level model examining the association among student, school, and classroom variables using risk ratio.

	Academic Engagement		Academic Engagement Cross-Level Interaction (1b)		Disciplinary Environment Direct Effect (2a)		Disciplinary Environment Cross-Level Interaction (2b)	
	Direct Effect (1a)							
	<i>Coeff.</i>	<i>p</i>	<i>Coeff.</i>	<i>p</i>	<i>Coeff.</i>	<i>p</i>	<i>Coeff.</i>	<i>p</i>
Level 1 (π; Student)								
Male	-0.01	0.13	-0.01	0.15	0.02**	0.002	0.02**	0.002
Black	0.03***	< 0.001	0.02*	0.03	0.02*	0.01	0.02*	0.03
B or Higher Grade	0.20***	< 0.001	0.20***	< 0.001	0.10***	< 0.001	0.10***	< 0.001
Level 2 (β; Classroom)								
Positive Behavior Support (PBS)	0.02	0.29	0.01	0.28	0.01	0.25	0.01	0.17
Percent of Male Students	-0.02	0.08	-0.02	0.08	-0.02*	0.03	-0.02*	0.045
Percent of Students of Color	0.04*	0.02	0.04*	0.02	0.02	0.41	0.01	0.47
Level 3 (γ; School)								
Percent of FARMS	0.00	0.84	0.00	0.83	-0.13***	< 0.001	-0.13***	< 0.001
Black-White Risk Ratio (RR)	-0.03	0.09	-0.03	0.08	-0.05*	0.03	-0.05*	0.03
Percent of Students Suspended	0.00	0.85	0.00	0.86	-0.03	0.24	-0.03	0.28
Total Enrollment	-0.02	0.36	-0.02	0.36	-0.07**	0.008	-0.07*	0.01
Cross-Level Interaction								
Black x RR	N/A	N/A	-0.01	0.32	N/A	N/A	-0.02	0.06
PBS x RR	N/A	N/A	0.00	0.73	N/A	N/A	0.00	0.84
School-level ICC	0.010		0.025		0.025		0.067	
Classroom-level ICC	0.023		0.058		0.039		0.031	
Between-school Variance % Reduction	23.5		24.2		13.2		11.2	
Between-classroom Variance % Reduction	34.0		33.7		55.4		54.9	

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Coefficients are standardized; Significant effects are bolded; N/A = not applicable (i.e., interaction was not modeled). Total Enrollment variable was divided by 100 to aid in interpretation. Percent reduction represents the percent between-school or between-classroom variance explained as compared to the fully unconditional model.

positively associated with academic engagement ($\beta = 0.04$; $p = .02$). In other words, students in classrooms with a higher percentage of students of color reported significantly higher levels of academic engagement. Lastly, at level 1, student race ($\pi = 0.03$; $p < .001$) and grades ($\pi = 0.20$; $p < .001$) were significantly related to student perceptions of academic engagement. Specifically, Black students reported significantly higher perceptions of academic engagement than White students, and students receiving higher report card grades (i.e., B or better) reported significantly more academic engagement than students who received lower report card grades (C or lower). Cross-level interactions between racial discipline disproportionality, as measured by school-level risk ratio, interacted with student race (student level) and positive behavior support (classroom level) were not significantly related to student perceptions of academic engagement. All above-noted findings were robust to the additions of these cross-level interactions (see model 1b in Table 2).

3.2. Academic engagement and risk difference

Results for direct effect models using the risk difference as the measure of racial discipline disproportionality (see model 3a in Table 3) indicated there were no significant predictors at either the school or classroom level. At level 1, student grades ($\pi = 0.20$; $p < .001$) were significantly related to student perceptions of academic engagement, such that students receiving higher report card grades reported more favorable perceptions of academic engagement than students receiving lower grades. The cross-level interaction between racial discipline disproportionality, as measured by school-level risk difference, and student race was significantly related to student perceptions of academic engagement ($\gamma = 0.03$; $p = .01$). There is an overall negative effect of larger risk differences, but it appears that White students are most impacted by these gaps. In schools with smaller gaps, Black students reported less favorable perceptions of academic engagement than White students, but higher engagement than White students in schools with a larger risk difference (see Fig. 1). The highest engagement ratings provided were by White students in schools with small racial difference gaps. There was no significant cross-level interaction between risk difference and positive behavior support (see model 3b in Table 3).

Table 3

Results for 3-level model examining the association among student, school, and classroom variables using risk difference.

	Academic Engagement Direct Effect (3a)		Academic Engagement Cross- Level Interaction (3b)		Disciplinary Environment Direct Effect (4a)		Disciplinary Environment Cross- Level Interaction (4b)	
	<i>Coeff.</i>	<i>p</i>	<i>Coeff.</i>	<i>p</i>	<i>Coeff.</i>	<i>p</i>	<i>Coeff.</i>	<i>p</i>
Level 1 (π; Student)								
Male	-0.01	0.13	-0.01	0.16	0.02**	0.002	0.03**	0.001
Black	0.03***	< 0.001	0.02	0.09	0.02*	0.01	0.02	0.09
B or Higher Grade	0.20***	< 0.001	0.20***	< 0.001	0.09***	< 0.001	0.10***	< 0.001
Level 2 (β; Classroom)								
Positive Behavior Support (PBS)	0.01	0.35	0.01	0.33	0.01	0.29	0.01	0.19
Percent of Male Students	-0.2	0.09	-0.02	0.09	-0.02*	0.03	-0.02	0.05
Percent of Students of Color	0.03	0.09	0.03	0.10	0.01	0.74	0.00	0.82
Level 3 (γ; School)								
Percent of FARMS	-0.01	0.78	-0.01	0.80	-0.13**	0.001	-0.13**	0.001
Black-White Risk Difference (RD)	0.04	0.15	0.04	0.15	0.04	0.35	0.04	0.32
Percent of Students Suspended	0.04	0.30	0.04	0.30	0	0.98	0.00	0.93
Total Enrollment	-0.02	0.44	-0.02	0.45	-0.07*	0.01	-0.06*	0.02
Cross-Level Interaction								
Black x RD	N/A	N/A	0.03*	0.01	N/A	N/A	0.02	0.16
PBS x RD	N/A	N/A	0.01	0.61	N/A	N/A	0.00	0.81
School-level ICC	0.010		0.025		0.027		0.073	
Classroom-level ICC	0.024		0.057		0.041		0.036	
Between-school Variance % Reduction	23.0		24.0		13.5		11.4	
Between-classroom Variance % Reduction	33.4		32.9		52.5		50.3	

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Coefficients are standardized; Significant effects are bolded; N/A = not applicable (i.e., interaction was not modeled). Total Enrollment variable was divided by 100 to aid in interpretation. Percent reduction represents the percent between-school variance explained as compared to the fully unconditional model.

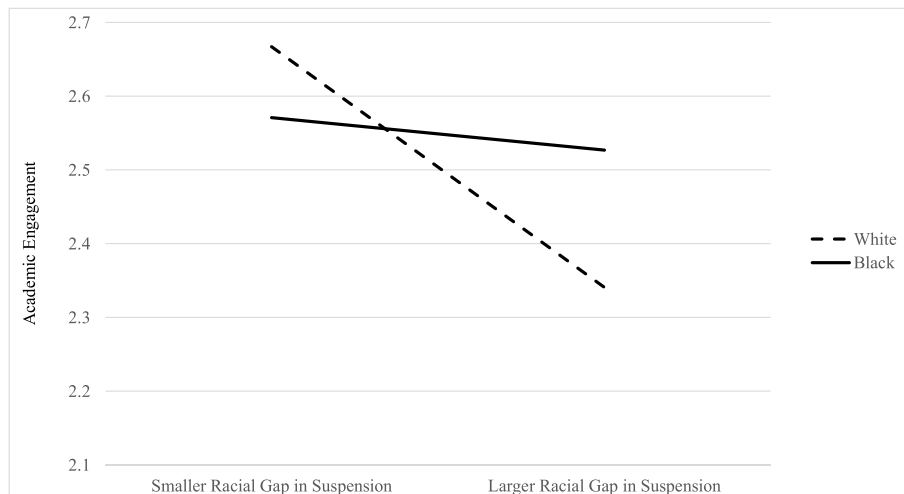


Fig. 1. Cross-Level Interaction of Racial Discipline Disproportionality Measured by Risk Difference and Student Race (Black vs. White) on Academic Engagement.

3.3. Disciplinary environment and risk ratio

The results of the direct effects model using risk ratio as the measure of racial discipline disproportionality (see model 2a in Table 2) indicated that discipline disproportionality was significantly and negatively associated with student perceptions of the

school disciplinary environment ($\gamma = -0.05$; $p = .03$). Similarly, the percentage of students receiving FARMs ($\gamma = -0.13$; $p < .001$) and total enrollment ($\gamma = -0.07$; $p = .008$) were also significantly and negatively related. In other words, students rated the disciplinary environment less favorably in schools with higher discipline disproportionality, more students receiving FARMs, and a larger student enrollment. Worth noting is that the beta values for FARMs are rather small, but reflect the decrease in ratings of the environment for each additional 1% of students receiving FARMs. At the larger scale (e.g., comparing schools that have 20% more students receiving FARMs), the magnitude of this impact would be more sizable. At level 2, the percent of males present ($\beta = -0.03$; $p = .03$) was significantly associated with perceptions of the disciplinary environment, such that students in classrooms with higher proportions of males present reported a poorer school disciplinary environment. Gender ($\pi = 0.02$; $p = .002$), student race ($\pi = 0.02$; $p = .01$), and grades ($\pi = 0.10$; $p < .001$) were also significantly related to student perceptions of the school disciplinary environment at level 1. Specifically, male students, Black students, and students receiving higher report card grades (i.e., Bs or better) reported significantly more favorable perceptions of the school disciplinary environment than females, White students, and students with C grades or worse. Cross-level analyses indicated no significant cross-level interactions. Findings from the disciplinary environment direct effects model were robust to the additions of the cross-level interaction variables.

3.4. Disciplinary environment and risk difference

Results for direct effect models using the risk difference as the measure of racial discipline disproportionality (see model 4a in Table 3) mirrored the above findings with the exception that when modeled as risk difference, the racial disproportionality finding no longer held. Thus, this model again indicated that both the percentage of students receiving FARMs ($\gamma = -0.13$; $p = .001$) and total enrollment ($\gamma = -0.07$; $p = .02$) were significantly and negatively related to student perceptions of school disciplinary environment. At the classroom level, the percent of males present ($\beta = -0.02$; $p = .03$) was significantly associated with perceptions of the disciplinary environment. Gender ($\pi = 0.02$; $p = .002$), grades ($\pi = 0.09$; $p < .001$), and student race ($\pi = 0.02$; $p = .01$) were significantly related to student perceptions of the school disciplinary environment at level 1. The interpretations of these findings are the same as the above section. Cross-level interactions between racial discipline disproportionality, as measured by school-level risk difference, and student race and positive behavior support were not related to student perceptions of the disciplinary environment (see model 4b in Table 3).

4. Discussion

The current study sought to understand how students' perceptions of academic engagement and the school disciplinary environment were associated with teacher's use of positive behavior support in the classroom and school-level racial discipline disproportionality, while accounting for multiple other classroom and school contextual variables. Two measures of disproportionality, the risk ratio and the risk difference, were leveraged to add rigor and methodological robustness to the study. We were also interested in how the perceptions of these constructs varied by student race for Black and White students. Below we interpret the findings from the multilevel analyses and conclude by discussing some of the implications for school psychology research and practice.

4.1. School climate perceptions: academic engagement and disciplinary environment

We found four small, significant associations involving the student race variable. First, Black students reported significantly greater academic engagement than White students (both in general and in schools with greater risk difference). Although this was not an original hypothesis, this finding supports earlier research on the academic achievement paradox (Bottiani, Bradshaw, & Mendelson, 2014; 2016; Shernoff & Schmidt, 2008), whereby Black students report consistently higher on self-reports of academic engagement than White students (Bottiani, Bradshaw, & Mendelson, 2014, 2016; Shernoff & Schmidt, 2008). Second, students in classrooms with a higher percentage of students of color reported significantly and slightly higher levels of academic engagement, which may relate to the student-level race finding. Third, Black students reported significantly and slightly more positive perceptions of the school disciplinary environment than White students overall, but the cross-level interaction between student race and the racial discipline disproportionality at the school level indicate there is more nuance to this finding.

In understanding these findings, we draw upon prior literature on the benefits of diverse classrooms, particularly for students of color (Ayscue, Frankenberg, & Siegel-Hawley, 2017). Specifically, not only do racially diverse learning environments have positive impacts on academic achievement for students of all races, but students of color achieve at higher levels in racially diverse schools than in segregated schools (Ayscue et al., 2017; Reardon, 2016). Regarding the interaction finding, when student race was interacted with the risk difference indicator, the results suggested that there is higher academic engagement for students in schools with low racial discipline disproportionality. However, we found that White students may have been more sensitive to the racial disproportionality than Black students, such that their report of academic engagement was significantly lower than Black students' in schools with higher racial discipline disproportionality. In fact, our analyses suggested that Black students had somewhat comparable ratings of academic engagement in both racially disproportional and non-disproportional settings. This finding may be because teachers need more support implementing culturally responsive classroom management. Specifically, teachers may stop their instruction frequently to manage student behaviors, particularly in classrooms with a higher percent of students of color. This constant interruption to instruction may contribute to White students feeling less engaged. This was the only significant cross-level interaction.

We also found that male students reported significantly and slightly more favorable perceptions of the school disciplinary environment than females, but that students in classrooms with higher proportion of males present reported poorer ratings of the school

disciplinary environment. Given the literature demonstrating that males are referred and suspended more than females (Skiba, Chung et al., 2014), it may be that the males are more attuned to rules because they are reminded more frequently about the rules by way of more frequent referrals for themselves and same-gender peers. This finding supports earlier research suggesting that there are gender differences around perceptions of punishment whereby males perceived punishment as more appropriate than females (Kelder, McNamara, Carlson, & Lynn, 1991).

In addition, the results suggested that students receiving higher report card grades (i.e., B or better) reported significantly and moderately more academic engagement and more positive perceptions of the school disciplinary environment than students who received lower report card grades (C or lower). The first finding, that students receiving higher report card grades are more engaged academically follows logically. Students who are academically engaged tend to also have more positive behavior and be grouped with other students in classes who are similarly behaved because of ability grouping and tracking and systems (Veldman & Sanford, 1984). In particular, Language Arts classes (which comprise a large portion of this data), may be to some extent “tracked” by reading level. Thus, the amount of time spent in classrooms with high-performing peers may account for the more favorable perceptions of the disciplinary by students with higher grades. Results also indicated that students in schools with more students receiving FARMS and a larger student enrollment rated the disciplinary environment *less* favorably. This finding adds to literature on varying discipline practices by context (Noltemeyer & McLoughlin, 2010). It may be that it may be more difficult to maintain a positive discipline environment in larger schools and in schools where the economic background may vary from the middle-class context of the majority of the teaching force.

Our hypotheses that teachers' use of positive behavior supports would moderate the association between disproportionate school discipline and students' perceptions of the two school climate domains was not supported. Prior research has characterized the broader classroom disciplinary environment as including students' perceptions of each other's behavior (e.g., whether they listen), their teachers' ability to manage disruptive behavior, and the structure of the school environment in terms of having rules and the extent to which students are rewarded for positive behavior (Farmer, Reinke, & Brooks, 2014). These domains are well aligned with aspects of Positive Behavioral Interventions and Supports (PBIS; Farmer et al., 2014; Horner, Sugai, & Anderson, 2010). Therefore, it is surprising that students in the classrooms of teachers who use more positive interactions in their classroom did not have better perceptions of the school's disciplinary environment, particularly since students spend most of the day in the classroom. Moreover, we expected that positive teacher-student classroom interactions would buffer the association between high rates of disproportionality and student perceptions of the disciplinary environment, but this hypothesis was not supported. It may be that the proportion of positive interactions alone is insufficient to buffer this association or that the frequency of these interactions were not enough to offset this association between disproportionality and student perceptions of the disciplinary environment. It also may be that negative interactions (i.e., reprimands, disapprovals) are necessary to examine explicitly. These findings highlight the need for more research into other protective factors and/or teacher practices that may buffer this association.

4.2. Racial discipline disproportionality

The findings suggest that in schools characterized by higher levels of racial discipline disproportionality as measured by the risk ratio, both Black and White students reported less favorable perceptions school disciplinary environment. These results are consistent with and strengthen conclusions from prior research suggesting that racial gaps in discipline practices affect students' school experiences broadly (Anyon et al., 2016; Bottiani et al., 2017) and demonstrate the relevance of examining student perceptions of school discipline in relation to racial disciplinary disproportionality. Yet it is important to note that the association between racial discipline disproportionality and perceived school disciplinary environment was not replicated when we used the risk difference metric to operationalize racial discipline disproportionality. One substantive difference in the two metrics to keep in mind is that the risk ratio is more of an indicator of disparity, whereas the risk difference may be more of an indicator of high frequency of exclusion among Black students. It may be that inequity itself, rather than frequency, is more of a driver of students' perceptions of the disciplinary environment. Although we cannot draw firm conclusions as to why these metrics yielded different results, the discrepancy itself is suggestive of the importance of utilizing multiple disproportionality metrics in education research to ensure valid and precise inferences are made.

In terms of racial discipline disproportionality as a cross-level moderator, we found that the risk difference metric moderated the association between Black student race and academic engagement, but did not find a moderation effect in this association when using the risk ratio metric. Furthermore, this finding was in the opposite direction of what we had hypothesized, such that White students' academic engagement was more negatively associated with racial discipline disproportionality than Black students' (i.e., Black students' report of academic engagement appeared less sensitive to racial discipline disproportionality than White students, although the trend was negative for both). Keeping in mind that risk difference may indicate higher frequency of exclusion of Black students, rather than disparity as indicated by the ratio metric, this finding is consistent with prior research that suggests high rates of exclusionary discipline negatively affect students' perceptions of school climate (Mitchell & Bradshaw, 2013). Although it is more intuitive that Black students' academic engagement would be more related to racial discipline disproportionality than White students', the pattern of association was negative for both; it may simply be that Black students' academic engagement is more resilient to these effects than White students'.

No differences in the association between racial discipline disproportionality and student perceptions of the disciplinary environment were found between Black and White students, for either the risk ratio difference or the risk difference metric, suggesting racial discipline disproportionality relates to students' perceptions of positive disciplinary environment similarly regardless of race.

4.3. Strengths and limitations

This study offers several methodological strengths that contribute to the research literature. Specifically, multilevel modeling was used to not only account for the nested nature of the data (e.g., students clustered in classrooms within schools), but also to leverage relevant factors across three levels to better understand these student perceptual outcomes. Moreover, this study drew from a large relatively diverse sample of students, classrooms, and high schools, and also pooled from multiple data collection methods (i.e., student self-reported surveys, observational measures conducted by external observers, two metrics of disproportionality, and administrative datasets). To avoid bias in parameter estimates, multiple imputation techniques were used instead of listwise deletion of missing data, helping to strengthen the analyses and results. Moreover, the standardized coefficients eases interpretation of the results, specifically in terms of the magnitude of effect sizes. Even with these and other methodological strengths, there are some limitations to consider.

One measurement limitation to consider is related to the measurement of race at different levels (i.e., individual, classroom). Race was measured via self-report at the student-level. However, at the classroom level, it was measured during observations, such that the trained observers counted the number of ostensibly “White” students present in the classroom. The approach taken by observers was to assess the degree to which *visible* racial/ethnic minorities were present by counting apparently White students in proportion to the total number of students. One potential concern regarding this approach is that it does not confirm with students their actual self-reported racial/ethnic identity. Unfortunately, we were not able to link student demographics with classroom observations given the anonymous nature of both the survey and observational measure. Disruption of classroom activities during observational periods to ask students or teachers to self-identify was not feasible. Although this is an imperfect measure, we considered it pertinent to our research questions to assess the potential for bias exposure by ascertaining whether a student is *perceived* by others to be a racial/ethnic minority (i.e., because such bias is based on others’ perceptions). We strongly caution against any inferences regarding observed race/ethnicity in this study as a within-student identity characteristic; instead, observed proportion of students of color represents an ecological context variable that represents the classroom’s potential for exposure to bias in this study. Similarly, in this study we compared disproportionality metrics between Black and White students as the mean percentage of Asian and Hispanic students in the sample of schools was small for Asian and Hispanic students (only 4% and 6% respectively). However, future research could compare Black students with students from other backgrounds.

Relatedly, we found that students in classrooms with a higher percentage of students of color reported significantly higher levels of academic engagement; however, this finding was only significant in the model that used the risk ratio indicator of disproportionality, and not the model that used the risk difference indicator. This inconsistency suggests the need for further research to confirm this finding. Although not necessarily a limitation per se, it does highlight the need for researchers to consider using multiple indicators of disproportionality in their work. Moreover, though statistically significant, our effect sizes, reported by the standardized coefficients at each level, are mostly in the small range, with the exception of the grades finding, which was moderate. However, it is important to note that the effect size indicates change per each one percent increase in students of color or males, and thus, the effect size, for example, of 0.04 for percent of students of color is reflective of change when there is 1% more students of color in the classroom. If comparing classrooms differing by 10%, this effect is actually 0.40 (and moderate in size).

Further, we dichotomized some of the ordinal variables (e.g., report card grades) to aid in interpretation; however, more precise measurement of these indicators may prove helpful in future studies. It is possible the reduction in variance resulting from the dichotomization resulted in small estimated effect sizes; therefore, more precise measurement may be warranted. It is also possible that the positive behavior supports variable may not fully capture teachers’ positive classroom management practices. More research is needed to determine under what circumstances (if any) using positive behavior supports in the classroom is likely to moderate the effects of disproportionality. These findings are consistent, however, with extant research that suggests that PBIS alone is not associated with reductions in disproportionality (Bradshaw et al., 2012; Vincent, Swain-Bradway et al., 2011; Vincent & Tobin, 2011); in fact, recent research suggests that reductions in disproportionality are more likely to occur when PBIS is combined with professional development and coaching of teachers in their use of culturally-responsive practices (Bradshaw et al., 2018). Future research may also consider other teacher characteristics in relation to student and classroom factors, such as teacher ethnicity and racial/ethnic match with students. Moreover, additional disproportionality metrics should be considered (e.g., individual suspension risk) in future studies. Unfortunately, however, we did not have data on individual risk.

Other limitations include the cross-sectional study design; therefore, we cannot make inferences about the directionality of the associations between constructs. Additionally, we only examined the perceptions of White and Black students in a single mid-Atlantic state, which limits generalizability to states with different racial/ethnic compositions (e.g., states differ in different population diversity). Future research should consider examining disproportionality in samples including other ethnic groups and using data from other states, particularly states in other geographic locations that show other groups (e.g., Latinx students) are disproportionately represented in exclusionary discipline. Similarly, these analyses are pertinent to students in high schools (grades 9–12) and may not generalize to middle or elementary school students.

4.4. Implications and conclusions

Student engagement has become a central focus for programs aiming to improve conditions for learning in the United States, mainly because high engagement is linked with other several positive student outcomes, including achievement and key social, emotional, and behavioral competencies needed for a successful transition into adulthood (Aspen Institute, 2017; Wang & Fredricks, 2014). Similarly, student behavior and the use of exclusionary discipline in schools are an acute concern, particularly with respect to

how exclusionary disciplinary actions are applied disproportionately by race, and given how steeply rates of suspension among Black students have risen over the past several decades (Losen et al., 2015).

These issues are of particular interest to school psychologists, as professional organizations have clearly outlined the role that school psychologists should play in addressing them. For example, the *American Psychological Association (APA) Division 16 (2012)* emphasizes “the enhancement of the status of children, youth, and adults as learners and productive citizens in schools and other settings” (p.1). Further, the *National Association of School Psychologists (NASP, 2010)* emphasizes that psychologists should play an integral role in school-wide practices to promote learning and preventive and responsive services, as well as student-level services to address academic and social skills. Thus, the issue of racial discipline disproportionality should be of central focus both in research and practice, and has increasingly become of interest as evidenced by the formation of statements and committees regarding social justice in schools (NASP, 2019). These issues are also very timely, given the increasing diversity of U.S. public schools, and the growing interest in school climate, as emphasized in the *Every Student Succeeds Act (ESSA, 2015, p. 1177)*.

Unfortunately, the findings of this study suggest that more work is needed in operationalizing the measurement of the construct of school-level racial discipline disproportionality for research purposes. Specifically, these results suggested that the association between disproportionality and student perceptions varies based on the metric used to operationalize racial discipline disproportionality. We also observed conflicting results for student perceptions of the school disciplinary environment (i.e., risk ratio, but not risk difference, was associated with a positive disciplinary environment). On the other hand, there was no direct association between disproportionality, as measured by either metric, and academic engagement, which is an area also in need of additional research. Although recent work has provided guidance to schools and districts on the measurement of racial discipline disproportionality (e.g., Girvan et al., 2019), there is less guidance for researchers who are interested in the effects of racial discipline disproportionality on school climate or the best way to measure it as an environmental feature of the school in multilevel models. Future research could explore the gamut of disproportionality metrics as they relate to student, staff, and observationally-assessed school climate indicators to provide a foundation for which metric is best to use.

The findings presented here highlight the importance of examining factors at multiple levels, and confirm the complexity and difficulty schools have encountered in impacting academic engagement and the school disciplinary environments. There is research demonstrating the promise of school-wide approaches, which address overall reliance on reactive disciplinary strategies like office discipline referrals and suspensions, to improve school climate (Bradshaw, Koth, Thornton, & Leaf, 2009; Horner et al., 2010), but also evidence that more needs to be done to address racial discipline disproportionality (Bottiani et al., 2018). Specifically, more may be needed at the school level to ensure that rules are constructed in a culturally-responsive manner, but schools may also need to provide further opportunities for students to engage in academic and social activities (Lewallen, Hunt, Potts-Datema, Zaza, & Giles, 2015; Zimmerman et al., 2018). To leverage school-wide strategies, more training for pre- and in-service teachers in positive, evidence-based behavior interventions and culturally responsive practices may also be warranted (Bradshaw et al., 2018). Additional research is needed not only to further explore the interaction of these factors at the classroom level in relation to teacher practices (Larson, Pas, Bradshaw, Rosenberg, & Day-Vines, 2018), but also the rigorous testing of coaching and professional development models that aim to improve positive teacher-student classroom interactions and classroom management (see for example, Bradshaw et al., 2018). Moreover, teacher educators will likely want to train pre-service educators in these skills prior to their entry into classrooms.

The student-level findings solidify that disproportionality is not just a discipline, or race, issue. Students perceive academic engagement and the school environment differently based on demographics (race and gender) as well as their achievement. This is important for schools to recognize, and for school psychologists to encourage teams to consider in data-based decision making. The examination of a *wide range* of data (i.e., not just achievement or behavior but also climate) by different student groups (and not just race) is imperative to ensuring that every student succeeds in our schools. In the new guidelines from ESSA, states and thus schools were encouraged to think beyond the most common metrics of accountability, and this study re-affirms the importance of doing so.

Acknowledgement

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305H150027 (PI: C. Bradshaw) and the National Institute of Justice (2014-CK-BX-0005), to the University of Virginia. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences, the U.S. Department of Education, or the National Institute of Justice. This work was also funded in part by a grant from the William T. Grant Foundation (PI: Jessika Bottiani). We would like to thank the Maryland State Department of Education and Sheppard Pratt Health System for their support of this research through the Maryland Safe and Supportive Schools Project.

References

- American Psychological Association Division 16. (2010). *Division 16 By-laws*. Retrieved from <https://apadivision16.org/wp-content/uploads/2016/11/Division-16-Current-Bylaws.pdf>.
- Anyon, Y., Zhang, D., & Hazel, C. (2016). Race, exclusionary discipline, and connectedness to adults in secondary schools. *American Journal of Community Psychology, 57*(3–4), 342–352.
- Aspen Institute National Commission on Social, Emotional, and Academic Development. (2017). *The evidence base for how we learn: Supporting students' social, emotional, and academic development*. Retrieved from https://assets.aspeninstitute.org/content/uploads/2018/03/FINAL_CDS-Evidence-Base.pdf.
- Ayscue, J., Frankenberg, E., & Siegel-Hawley, G. (2017). The complementary benefits of racial and socioeconomic diversity in schools. *Education, 36*, 197–238.
- Bartlett, J. W., Seaman, S. R., White, I. R., & Carpenter, J. R. (2015). Multiple imputation of covariates by fully conditional specification: Accommodating the

- substantive model. *Statistical Methods in Medical Research*, 24(4), 462–487.
- Bondy, E., Ross, D. D., Galligane, C., & Hambacher, E. (2007). Creating environments of success and resilience: Culturally responsive classroom management and more. *Urban Education*, 42(4), 326–348.
- Bottiani, J. H., Bradshaw, C. P., & Gregory, A. (2018). Nudging the gap: Introduction to the special issue “closing in on discipline disproportionality. *School Psychology Review*, 47(2), 109–117.
- Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2014). Promoting an equitable and supportive school climate in high schools: The role of school organizational health and staff burnout. *Journal of School Psychology*, 52, 567–582. <https://doi.org/10.1016/j.jsp.2014.09.003>.
- Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. M. (2016). Racial inequality in Black and White high school students' perceptions of school support: An examination of race in context. *Journal of Youth and Adolescence*, 45(6), 1176–1191. <https://doi.org/10.1007/s10964-015-0411-0>.
- Bottiani, J., Bradshaw, C., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and White adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, 109(4), 532–545. <https://doi.org/10.1037/edu0000155>.
- Bradshaw, C. P., Debnam, K. J., Lindstrom Johnson, S., Pas, E., Hershfeldt, P., Alexander, A., et al. (2014a). Maryland's evolving system of social, emotional, and behavioral interventions in public schools: The Maryland Safe and Supportive Schools Project. *Adolescent Psychiatry*, 4(3), 194–206. <https://doi.org/10.2174/221067660403140912163120>.
- Bradshaw, C. P., Koth, C. W., Thornton, L. A., & Leaf, P. J. (2009). Altering school climate through school-wide positive behavioral interventions and supports: Findings from a group-randomized effectiveness trial. *Prevention Science*, 10(2), 100.
- Bradshaw, C. P., O'Brennan, L. M., & McNeely, C. A. (2008). Core competencies and the prevention of school failure and early school leaving. In N. G. Guerra, C. P. Bradshaw, & Guest (Vol. Eds.), *Core competencies to prevent problem behaviors and promote positive youth development*. Vol. 122. *Core competencies to prevent problem behaviors and promote positive youth development* (pp. 19–32). New Directions for Child and Adolescent Development. <https://doi.org/10.1002/cd.226>.
- Bradshaw, C. P., Pas, E. T., Bottiani, J. H., Debnam, K. J., Reinke, W., Herman, K., et al. (2018). *Promoting cultural responsiveness and student engagement through Double Check coaching of classroom teachers: An efficacy study*. *School Psychology Review*.
- Bradshaw, C. P., Waasdorp, T. E., Debnam, K. J., & Lindstrom Johnson, S. (2014b). Measuring school climate in high schools: A focus on safety, engagement, and the environment. *Journal of School Health*, 84(9), 593–604.
- Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics*, 130(5), e1136–e1145. <https://doi.org/10.1542/peds.2012-0243>.
- Carter, M., McGee, R., Taylor, B., & Williams, S. (2007). Health outcomes in adolescence: Associations with family, friends and school engagement. *Journal of Adolescence*, 30, 51–62.
- Cash, A. H., Debnam, K. J., Waasdorp, T. E., Wahl, M., & Bradshaw, C. P. (2019). Adult and student interactions in nonclassroom settings. *Journal of Educational Psychology*, 111, 104–117. <https://doi.org/10.1037/edu0000275>.
- Childs, K. E., Kincaid, D., George, H. P., & Gage, N. A. (2016). The relationship between school-wide implementation of positive behavior intervention and supports and student discipline outcomes. *Journal of Positive Behavior Interventions*, 18(2), 89–99.
- Cogshall, J. G., Osher, D., & Colombi, G. (2013). Enhancing educators' capacity to stop the school-to-prison pipeline. *Family Court Review*, 51(3), 435–444.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- Cook, C. R., Duong, M. T., McIntosh, K., Fiat, A. E., Larson, M., & Pullmann, M. D. (2018). Addressing discipline disparities for Black male students: Linking malleable root causes to feasible and effective practices. *School Psychology Review*, 47(2), 135. <https://doi.org/10.17105/SPR-2017-0026.V47-2>.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.
- Enders, C. K., Keller, B. T., & Levy, R. (2017). A chained equations imputation approach for multi-level data with categorical and continuous variables. *Psychological Methods*, 23(2), 298–317.
- ESSA. (2015). *Every student succeeds Act of 2015*. Pub. L. No. 114-95 § 114 Stat (2015-2016).
- Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. New York: Council of State Governments Justice Center.
- Farmer, T. W., Reinke, W. M., & Brooks, D. S. (2014). Managing classrooms and challenging behavior: Theoretical considerations and critical issues. *Journal of Emotional and Behavioral Disorders*, 22(2), 67–73.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: John Wiley & Sons, Inc.
- Ford, A. C., & Sassi, K. (2014). Authority in cross-racial teaching and learning (re) considering the transferability of warm demander approaches. *Urban Education*, 49, 39–74. <https://doi.org/10.1177/0042085912464790>.
- Gage, N. A., Grasley-Boy, N., George, H. P., Childs, K., & Kincaid, D. (2019). A quasi-experimental design analysis of the effects of School-Wide Positive Behavior Interventions and Supports on discipline in Florida. *Journal of Positive Behavior Interventions*, 21, 50–61. <https://doi.org/10.1177/1098300718768208>.
- Ganao, J. S., Silvestre, F. S., & Glen, J. W. (2013). Assessing the differential impact of contextual factors on school suspension for Black and White Students. *The Journal of Negro Education*, 82(4), 393–407. <https://doi.org/10.7709/jnegroeducation.82.4.0393>.
- Gion, C. M., McIntosh, K., & Horner, R. (2014). *Patterns of minor office discipline referrals in schools using SWIS*. Retrieved from https://www.pbis.org/Common/Cms/files/pbisresources/EvalBrief_May2014.pdf.
- Girvan, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2017). The relative contribution of subjective office referrals to racial disproportionality in school discipline. *School Psychology Quarterly*, 32(3), 392.
- Girvan, E. J., McIntosh, K., & Smolkowski, K. (2019). Tail, tusk, and trunk: What different metrics reveal about racial disproportionality in school discipline. *Educational Psychologist*, Advance online publication. <https://doi.org/10.1080/00461520.2018.1537125>.
- Heilbrun, A., Cornell, D., & Konold, T. (2018). Authoritative school climate and suspension rates in middle schools: Implications for reducing the racial disparity in school discipline. *Journal of School Violence*, 17(3), 324. <https://doi.org/10.1080/15388220.2017.1368395>.
- Heilbrun, A., Cornell, D., & Lovegrove, P. (2015). Principal attitudes regarding zero tolerance and racial disparities in school suspensions. *Psychology in the Schools*, 52(5), 489–499.
- 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.
- Horner, R. H., Sugai, G., & Anderson, C. M. (2010). Examining the evidence base for schoolwide positive behavior support. *Focus on Exceptional Children*, 42(8), 1–14.
- Hox, J. (1998). Multilevel modeling: When and why. In I. Balderjahn, R. Mathar, & M. Schader (Eds.), *Classification, data analysis, and data highways* (pp. 147–154). New York: Springer Verlag.
- Huang, F. L., & Cornell, D. (2018). The relationship of school climate with out-of-school suspensions. *Children and Youth Services Review*, 94, 378–389. <https://doi.org/10.1016/j.childyouth.2018.08.013>.
- Johnson, S. L. (2009). Improving the school environment to reduce school violence: A review of the literature. *Journal of School Health*, 79(10), 451–465.
- Kelder, L. R., McNamara, J. R., Carlson, B., & Lynn, S. J. (1991). Perceptions of physical punishment: The relation to childhood and adolescent experiences. *Journal of Interpersonal Violence*, 6(4), 432–445.
- Keller, B. T., & Enders, C. K. (2018). *Blimp user's manual*. Version 1.1 (Los Angeles, CA).
- Kirk, D. (2009). Unraveling the contextual effects on student suspension and juvenile arrest: The independent influences of school, neighborhood, and family social controls. *Criminology*, 47, 479–520.
- Kish, L. (1965). *Survey sampling*. New York, NY: Wiley.
- Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology*, 100, 96.
- Kupchik, A., & Catlaw, T. J. (2015). Discipline and participation: The long-term effects of suspension and school security on the political and civic engagement of youth. *Youth & Society*, 47(1), 95–124.

- Larson, K. E., Pas, E. T., Bradshaw, C. P., Rosenberg, M. S., & Day-Vines, N. L. (2018). *Examining how proactive and culturally responsive classroom management relate to student behavior: Implications for measurement and practice*. School Psychology Review.
- Lewallen, T. C., Hunt, H., Potts-Datema, W., Zaza, S., & Giles, W. (2015). The whole school, whole community, whole child model: A new approach for improving educational attainment and healthy development for students. *Journal of School Health, 85*(11), 729–739.
- Li, Y., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Psychology, 47*, 233.
- Lindstrom Johnson, S., Reichenberg, R., Waasdorp, T., Shukla, K., & Bradshaw, C. P. (2018). *Improving the measurement of school climate: A novel application of item response theory*. Manuscript submitted for publication.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202.
- Losen, D. J., & Gillespie, J. (2012). *Opportunities suspended: The disparate impact of disciplinary exclusion from school*. Retrieved from <http://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/upcoming-cerr-research/>.
- Losen, D. J., Hodson, C. L., Keith, I. I., Michael, A., Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap? K-12 racial disparities in school discipline. The center for civil rights remedies*. Retrieved from https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federalreports/are-we-closing-the-school-discipline-gap/AreWeClosingTheSchoolDisciplineGap_FINAL221.pdf.
- Losen, D. J., & Martinez, T. E. (2013). *Out of school & off track: The overuse of suspensions in American middle and high schools*. Los Angeles, CA: The UCLA Center for Civil Rights Remedies at The Civil Rights Project. Retrieved from <https://files.eric.ed.gov/fulltext/ED541735.pdf>.
- McIntosh, K., Girvan, E. J., Horner, R., & Smolkowski, K. (2014). Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *Journal of Applied Research on Children*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2598212.
- Mitchell, M. M., & Bradshaw, C. P. (2013). Examining classroom influences on student perceptions of school climate: The role of classroom management and exclusionary discipline strategies. *Journal of School Psychology, 51*(5), 599–610.
- Mizel, M. L., Miles, J. N., Pedersen, E. R., Tucker, J. S., Ewing, B. A., & D'Amico, E. J. (2016). To educate or to incarcerate: Factors in disproportionality in school discipline. *Children and Youth Services Review, 70*, 102–111.
- Morgan, M. A., & Wright, J. P. (2018). Beyond Black and white: Suspension disparities for hispanic, asian, and white youth. *Criminal Justice Review, 43*(4), 377–398.
- National Association of School Psychologists (NASP) (2010). *Model for comprehensive and integrated school psychological services*. Retrieved from <https://www.nasponline.org/standards-and-certification/nasp-standards-revision>.
- National Association of School Psychologists (NASP) (2019). *Social justice: NASP guidance*. Retrieved from <http://www.nasponline.org/resources-and-publications/resources/diversity/social-justice>.
- National Research Council and Institute of Medicine of the National Academies. (2003). *Engaging schools: Fostering high school students' motivation to learn*. Washington, DC: The National Academies Press. <https://www.nap.edu/read/10421/chapter/2>.
- National School Climate Center. (2019). What is school climate? Retrieved from <https://www.schoolclimate.org/about/our-approach/what-is-school-climate>.
- Nishioka, V., Shigeoka, S., & Lohich, E. (2017). *School discipline data indicators: A guide for districts and schools (rel 2017–240)*. National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest Washington, DC: US Department of Education, Institute of Education Sciences. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Noltmeyer, A. L., Marie, R., McLaughlin, C., & Vanderwood, M. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review, 44*(2), 224. <https://doi-org.ezp.slu.edu/10.17105/spr-14-0008.1>.
- Noltmeyer, A., & McLaughlin, C. S. (2010). Patterns of exclusionary discipline by school typology, ethnicity, and their interaction. *Penn GSE Perspectives on Urban Education, 7*, 27–40.
- Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A vicious cycle: A social–psychological account of extreme racial disparities in school discipline. *Perspectives on Psychological Science, 11*(3), 381–398.
- Pas, E. T., Cash, A. H., O'Brennan, L., Debnam, K. J., & Bradshaw, C. P. (2015). Profiles of classroom behavior in high schools: Associations with teacher behavior management strategies and classroom composition. *Journal of School Psychology, 53*(2), 137–148.
- Pas, E. T., Ryoo, J. H., Musci, R., & Bradshaw, C. P. (2019). The effects of a state-wide scale-up of school-wide positive behavior intervention and supports on behavioral and academic outcomes: A quasi-experimental examination. *Journal of School Psychology, 73*, 41–55. <https://doi.org/10.1016/j.jsp.2019.03.001>.
- Peguero, A. A., & Bracy, N. L. (2015). School order, justice, and education: Climate, discipline practices, and dropping out. *Journal of Research on Adolescence, 25*(3), 412–426.
- Perreira, K. M., Wassink, J., & Harris, K. M. (2018). Beyond race/ethnicity: Skin color, gender, and the health of young adults in the United States. *Population Research and Policy Review, 1*–29.
- Pesta, R. (2018). *Labeling and the differential impact of school discipline on negative life outcomes: Assessing ethno-racial variation in the school-to-prison pipeline*. Crime & Delinquency0011128717749223.
- Petrosino, A., Fronius, T., Gool, C. C., Losen, D. J., & Turner, H. M. (2017). *Analyzing student-level disciplinary data: A Guide for districts. REL 2017-263*. Regional Educational Laboratory Northeast & Islands.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models*. Newbury Park, CA: Sage.
- Reardon, S. F. (2016). School segregation and racial academic achievement gaps. *The Journal of Social Sciences, 2*(5), 34–57.
- Reinke, W. M., Herman, K. C., & Newcomer, L. (2016). The Brief Student–Teacher Classroom Interaction Observation: Using dynamic indicators of behaviors in the classroom to predict outcomes and inform practice. *Assessment for Effective Intervention, 42*, 32–42.
- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*(3), 700–712. <https://doi.org/10.1037/a0027268>.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Rusby, J. C., Taylor, T., & Milchak, C. (2001). *Assessing school settings: Interactions of students and teachers (ASSIST) observation system* (Unpublished manual).
- Sandilos, L. E., Rimm-Kaufman, S. E., & Cohen, J. J. (2017). Warmth and demand: The relation between students' perceptions of the classroom environment and achievement growth. *Child Development, 88*(4), 1321–1337.
- Shermoff, D. J., & Schmidt, J. A. (2008). Further evidence of an engagement–achievement paradox among US high school students. *Journal of Youth and Adolescence, 37*(5), 564–580.
- Shukla, K., Waasdorp, T., Lindstrom Johnson, S., Nguyen, A., Orozco Solis, M. G., Colungar, C., et al. (2019). Does school climate mean the same thing in the U.S. As in Mexico? A focus on measurement invariance. *Journal of Psychoeducational Assessment* (in press).
- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children, 35*, 351–380.
- Skiba, R. J., Chung, C.-G., Trachok, M., Baker, T. L. A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal, 51*, 640–670.
- Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of african American and latino disproportionality in school discipline. *School Psychology Review, 40*(1), 85–107.
- Skiba, R. J., Mediratta, K., & Rausch, M. K. (Eds.). (2016). *Inequality in school discipline: Research and practice to reduce disparities*. Springer.
- Solórzano, D. G. (1997). Images and words that wound: Critical race theory, racial stereotyping, and teacher education. *Teacher Education Quarterly, 24*(3), 5–19.
- StataCorp (2015). *Stata statistical software: Release 14*. College Station, TX: StataCorp LP.
- Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8*(1), 2–8. <https://doi.org/10.1177/10634266000800101>.
- Tate, R. L., & Wongbunhit, Y. (1983). Random versus nonrandom coefficient models for multilevel analysis. *Journal of Educational Statistics, 8*(2), 103–120.
- Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357–385.

- US Department of Education (2009). Safe and supportive schools model. Available at: <http://safesupportiveschools.ed.gov/index.php?id=33>.
- U.S. Department of Education (2018). *2015-2016 Civil rights data collection: School climate and safety*. Washington, DC: Office for Civil Rights. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/docs/school-climate-and-safety.pdf>.
- Veldman, D. J., & Sanford, J. P. (1984). The influence of class ability level on student achievement and classroom behavior. *American Educational Research Journal*, *21*(3), 629–644.
- Vincent, C. G., Swain-Bradway, J., Tobin, T. J., & May, S. (2011). Disciplinary referrals for culturally and linguistically diverse students with and without disabilities: Patterns resulting from school-wide positive behavior support. *Exceptionality*, *19*, 175–190. <https://doi.org/10.1080/09362835.2011.579936>.
- Vincent, C. G., & Tobin, T. J. (2011). The relationship between implementation of school-wide positive behavior support (SWPBS) and disciplinary exclusion of students from various ethnic backgrounds with and without disabilities. *Journal of Emotional and Behavioral Disorders*, *19*(4), 217–232. <https://doi.org/10.1177/1063426610377329>.
- Voisin, D. R., & Elsaesser, C. (2016). Brief report: The protective effects of school engagement for African American adolescent males. *Journal of Health Psychology*, *21*(4), 573–576. <https://doi.org/10.1177/1359105314531607>.
- Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development*, *2003*(99), 9–16. <https://doi.org/10.1002/yd.51>.
- Wang, M. T., & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, *28*(2), 315–352.
- Wang, M.-T., & Fredricks, J. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Development*, *85*(2), 722–737. <https://doi.org/10.1111/cdev.12138>.
- Way, S. M. (2011). School discipline and disruptive classroom behavior: The moderating effects of student perceptions. *The Sociological Quarterly*, *52*(3), 346–375.
- Zimmerman, M. A., Eisman, A. B., Reischl, T. M., Morrel-Samuels, S., Stoddard, S., Miller, A. L., et al. (2018). Youth Empowerment Solutions: Evaluation of an after-school program to engage middle school students in community change. *Health Education & Behavior*, *45*, 20–31.