

BRIEF REPORT

In Pursuit of Equity: Discipline Disproportionality and SWPBIS Implementation in Urban Schools

Brittany Zakszeski, Laura Rutherford, Kamontá Heidelberg, and Lisa Thomas
Devereux Center for Effective Schools, King of Prussia, Pennsylvania, United States

School-wide positive behavioral interventions and supports (SWPBIS) frameworks have been suggested as a promising approach to reducing disproportionality in school discipline practices; however, few studies have tested this potential, and findings have been mixed. In our sample of 27 urban schools and approximately 15,000 students, risk difference and ratio trajectories for office discipline referrals (ODRs) across 3 years of Tier I implementation evidenced sustained or, in some cases, heightened disproportionality for students identified as Black/African American or Latinx/Hispanic (relative to White; in a study subsample), identified as male (relative to a female), enrolled in each school's upper three grades (relative to lowest three grades), and not participating in special education (relative to participating in special education). Tier I implementation fidelity minimally contributed to expounding trajectories and outcomes across schools. These findings highlight the importance of equity-focused implementation and advanced-tier supports in the service of social justice within schools.

Impact and Implications Statement

Schools may overrepresent various student demographic groups (e.g., according to race, gender, grade, and special education status) in ODRs. High-fidelity implementation of Tier I SWPBIS provides an important foundation for school improvement but may be insufficient in singlehandedly reducing discipline disproportionality. Equity-focused implementation and advanced-tier supports may be necessary to achieve equity in school discipline practices.

Keywords: school discipline, school-wide positive behavioral interventions and supports, multitiered systems of support, disproportionality, equity

Among the most well-documented educational disparities in the United States is the overrepresentation of students from racially minoritized groups in school discipline practices (Gregory et al., 2017). Students who are Black/African American, Latinx/Hispanic, and American Indian commonly receive office discipline referrals (ODRs) and exclusionary discipline at far higher rates than their White peers (e.g., Anyon et al., 2014; Losen et al., 2015; Skiba et al., 2011). For example, across the 2015–2016 school year, approximately 2.7 million students in the United States received at least one out-of-school suspension; Black/African American students accounted for 39% of these students, despite only accounting for 16% of all enrolled students (U.S. Department of Education, 2018). Frequent, ineffective use of exclusionary discipline practices may be especially prevalent in

underresourced schools in urban centers (Putnam et al., 2009), where Black/African American and Latinx/Hispanic students are commonly enrolled in the United States. Research connects receipt of exclusionary discipline to varied negative outcomes including disengagement from school, lower academic achievement, and future involvement in the juvenile justice system (e.g., Cartledge & Kourea, 2008), perpetuating the adverse experiences of historically disadvantaged groups.

Other types of discipline disproportionality, beyond racial disproportionality, may endure in schools. Growing evidence suggests that students who are male and who participate in special education receive ODRs and exclusionary discipline more frequently than their counterparts (U.S. Government Accountability Office, 2018). Additionally, ODRs and exclusionary discipline are dramatically increasing in middle and high schools (e.g., Losen & Martinez, 2013), suggesting older students may be likewise overrepresented in the discipline system.

Brittany Zakszeski  <https://orcid.org/0000-0002-0503-7879>

Kamontá Heidelberg  <https://orcid.org/0000-0003-1388-3191>

Kamontá Heidelberg is now at the University at Buffalo.

Correspondence concerning this article should be addressed to Brittany Zakszeski, Devereux Center for Effective Schools, 2012 Renaissance Blvd., King of Prussia, PA 19406, United States. Email: Brittany.Zakszeski@devereux.org

Addressing Discipline Disproportionality Through School-Wide Positive Behavioral Interventions and Supports

Multitiered systems of support (MTSS) frameworks are a promising approach to reducing disparities in education, with school-wide positive behavioral interventions and supports (SWPBIS),

specifically, being well matched to the goals of reducing exclusionary discipline and disproportionality in the discipline (Gregory et al., 2017; McIntosh, Gion, et al., 2018). A substantial body of literature supports the use of SWPBIS in promoting educationally significant academic, behavioral, and organizational outcomes (Lee & Gage, 2020) and identifies SWPBIS as a highly scalable approach to school reform (Horner et al., 2017).

High-fidelity SWPBIS implementation may promote equitable outcomes through a number of mechanisms targeting students and staff, as depicted in Figure 1. Within this logic model, we make multiple assumptions regarding potential sources of discipline inequities for students who are Black/African American, Latinx/Hispanic, male, in upper grades, and/or participating in special education, compared to their counterparts. First, we assume that staff's implicit biases (related to race, expressed gender, and disability; e.g., Girvan et al., 2017) and varying behavior tolerance levels across grade levels may partially account for disparities. This suggests that minimizing vulnerable discipline-related decision points may contribute to decreasing disparities (McIntosh, Ellwood, et al., 2018). Second, we assume that students and staff have lived experiences influenced by some shared and some divergent cultural socialization processes, such that they may enter the schooling environment with different learning histories and mental models for expected behavior norms (e.g., Carter et al., 2017). This suggests that minimizing ambiguity in behavioral expectations and providing explicit instruction on behavioral expectations may contribute to decreasing disparities. Finally, in alignment with cumulative risk theory (Rutter, 1979), we assume that students—especially those in high-poverty, urban centers—who are Black/African American, Latinx/Hispanic, male, in upper grades, and/or participating in special education are more likely to have accumulated risk factors as a result of their experiences in a society with imbalanced power and privilege, such that they may be more predisposed to engage in challenging behavior at school

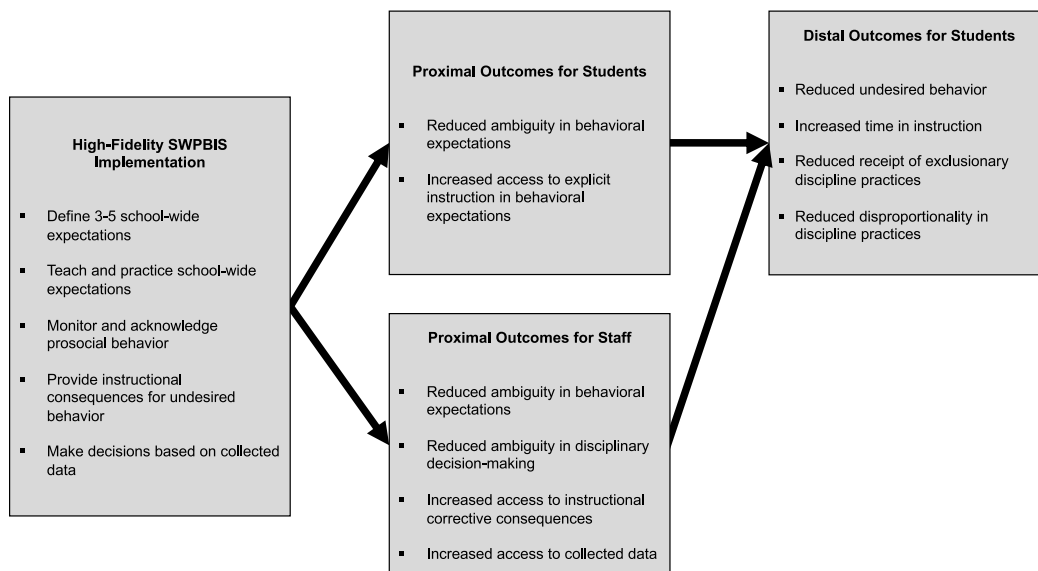
(e.g., Appleyard et al., 2005; Ashworth & Humphrey, 2020). This, too, suggests that providing explicit instruction on behavioral expectations, including acknowledgments for prosocial behavior and instructional consequence alternatives to exclusionary discipline for undesired behavior, may contribute to decreasing disparities.

A review of SWPBIS in relation to exclusionary discipline (Gage et al., 2018), synthesizing 4 studies across 90 schools, identified significant overall reductions in suspensions but no such reductions in ODRs. Findings have been inconsistent regarding outcomes for students most at risk for exclusionary discipline, with some studies suggesting positive effects and others identifying heightened disproportionality following SWPBIS implementation (Vincent et al., 2016). Generally, research in this area is lacking. Extant studies focus almost exclusively on racial disproportionality. Moreover, few studies examine the impact of SWPBIS implementation on discipline disproportionality, with mixed findings, and even fewer studies describe trends in disproportionality across years of SWPBIS implementation. Research is particularly needed to examine disproportionality in urban schools, which are commonly underresourced institutions serving students at risk for adverse outcomes.

The Current Study

To begin to explore this area, we designed this study to examine trajectories of ODR disproportionality according to students' race, gender, grade, and participation in special education across 3 years of SWPBIS implementation in an urban school district. We considered the extent to which schools evidenced disproportionality across years of SWPBIS implementation, attending specifically to the role of Tier I implementation fidelity in explaining variance in trajectory endpoints and slopes. This line of inquiry offers promise in informing comprehensive school equity endeavors.

Figure 1
Logic Model for Effects of SWPBIS on Discipline Equity



Method

Setting and Participants

This study involved secondary analysis of implementation and outcome data collected through ongoing SWPBIS initiatives. Institutional Review Board approval was obtained from both the school district and the authors' organization. Data were drawn from a large, urban district in the mid-Atlantic United States. Schools were included in the study if they (a) were enrolled in a SWPBIS training cohort between 2015 and 2018, (b) did not achieve a score of 70% or above on the Benchmarks of Quality (BoQ; Kincaid et al., 2010) prior to the 2016–2017 academic year, (c) completed the BoQ in the spring of 2019, and (d) entered ODR data into the School-Wide Information System™ (SWIS; May et al., 2020) between 2016 and 2019. Twenty-seven schools serving approximately 14,965 students in kindergarten through eighth grade, as described in Table 1, met these criteria and were included in analyses.

Measures

Disciplinary Outcomes

Office Discipline Referrals. ODRs offer an efficient means of measuring externalizing behavior (Irvin et al., 2006) and evidence moderate reliability and validity (Pas et al., 2011). Using standardized paper ODR forms, school personnel documented the following information regarding behavioral incidents: Student name, grade, date, time, referring staff, problem behavior, location, persons involved, probable motivation, and administrative decision. Personnel coded the “problem behavior” according to operationally defined SWIS behavior categories (Educational and Community Supports, n.d.), which included “minor” infractions (i.e., lower-level behaviors managed within the classroom) and “major” infractions (i.e., higher-level behaviors managed with the support of other school-based personnel). A designated staff member entered information from ODR forms into the SWIS web-based application (May et al., 2020). School personnel’s

Table 1

Demographics of Participating Schools (N = 27)

	<i>n</i> , schools		%, schools
Grades served			
K–5	3		11.11
K–6	3		11.11
K–8	18		66.67
5–8	1		3.70
6–8	2		7.40
PBIS cohort year			
2015	7		25.93
2016	4		14.81
2017	7		25.93
2018	9		33.33
	<i>n</i>	%, sample	<i>M</i> (<i>SD</i> , range), school
Student enrollment, 2018–2019	14,965		554 (258, 257–1,387)
Gender			
Male	7,881	52.66	293 (133, 135–714)
Female	7,084	47.34	262 (127, 119–673)
Grade			
Kindergarten	1,588	10.61	59 (35, 0–153)
First grade	1,677	11.21	62 (40, 0–169)
Second grade	1,693	11.31	63 (39, 0–164)
Third grade	1,704	11.39	63 (39, 0–161)
Fourth grade	1,691	11.30	63 (42, 0–181)
Fifth grade	1,679	11.22	62 (38, 0–182)
Sixth grade	1,960	13.10	73 (58, 0–271)
Seventh grade	1,488	9.94	55 (59, 0–271)
Eighth grade	1,485	9.92	55 (55, 0–229)
Characteristics			
CEP economically disadvantaged rate			99.97% (99.18%–100%)
Participating in special education	2,409	16.10	89 (45, 37–221)
English language learners	1,714	11.45	63 (80, 0–310)
Race/ethnicity			
American Indian	34	.23	1 (2, 0–8)
Asian	650	4.34	24 (47, 0–192)
Black/African American	9,262	61.89	343 (158, 42–627)
Latinx/Hispanic	3,077	20.56	114 (136, 5–436)
Multiracial	929	6.21	34 (28, 4–125)
Pacific Islander	14	.09	1 (2, 0–7)
White	999	6.68	37 (65, 1–262)

Note. CEP = Community Eligibility Provision.

accuracy and reliability in completing ODR forms were supported by building-level PBIS teams. Data entry staff's accuracy and reliability in entering ODR data into SWIS were promoted through building-specific systems as well as data integrity system alerts built into the SWIS platform. District PBIS coaches directly trained and consulted PBIS teams and data entry staff in these areas.

For the current study, ODR data were pulled from SWIS across the four marking periods of each study year (2016–2017, 2017–2018, 2018–2019). These data encompassed 41,912 referrals [14,442 minor infractions (34.46%), 27,470 major infractions (65.54%)]. Minor infractions were most commonly documented for minor disruptions ($n = 4,687$), minor defiance ($n = 3,136$), and inappropriate physical contact ($n = 2,377$). Major infractions were most commonly documented for physical aggression ($n = 5,215$), fighting ($n = 5,071$), and skipping class ($n = 3,334$).

ODR data were used to calculate risk indices, risk ratios, and, in some cases, risk differences for student groups hypothesized to be at risk. We recognize that each available disproportionality evaluation approach is imperfect and yields a limited description of the problem (Girvan et al., 2019). For example, risk ratios are unable to illuminate the absolute magnitudes of risk differences; however, risk ratios also provide intuitive and meaningful calculations of risk, and relative magnitudes of differences, that account for differences in enrollment, a critical consideration in our data set of schools of varying sizes (at fall 2018, $M = 554$ students, $SD = 253$, range: 257–1,387). Moreover, risk ratios have been selected as a choice metric in educational guidelines (e.g., Bollmer et al., 2014) and legislation (e.g., *Assistance to States for the Education of Children with Disabilities*, 34 CFR § 300.647 2018). Accordingly, we analyzed *risk ratios* as outcomes in study analyses examining differences between large student subgroups [i.e., gender, grade, and individualized education plan (IEP) status]. Given risk ratios' instability with the small target or reference groups (e.g., inability to be calculated when a reference group has a risk index of 0) in the context of our sample's limited racial diversity (Girvan et al., 2019), we descriptively analyzed *risk differences* in comparing outcomes across student racial groups.

Risk Indices. Risk indices were calculated by dividing the number of students in a demographic group who received one or more ODRs within the marking period by the total number of students in the demographic group enrolled in the school during that marking period (Boneshefski & Runge, 2014). For example, if 20 Black/African American students received one or more ODRs within a marking period and 100 Black/African American students were enrolled in the school during that marking period, the risk index for Black/African American students during that marking period would be .2, meaning 20% of Black/African American students received one or more ODRs during that marking period (Girvan et al., 2019).

Risk Ratios. Risk ratios for students identified as Black/African American, identified as Latinx/Hispanic, identified as male, enrolled in one of the upper three grades of the school, and with an IEP, respectively, were calculated by dividing the risk index for that demographic group and marking period by the risk index for the reference group (i.e., identified as White, identified as female, enrolled in one of the lowest three grades of the school, and did not have an IEP, respectively) during that marking period (Boneshefski & Runge, 2014). For example, given a risk index for Black/African American students of .2 and a risk index for White students of .15, the risk ratio would be 1.33, meaning Black/African American students were 1.33 times more

likely than their White peers to have received one or more ODRs during that marking period (Girvan et al., 2019).

Risk Differences. Risk differences for students identified as Black/African American or Latinx/Hispanic, respectively, were calculated by subtracting the risk index for the reference group (i.e., identified as White) from the risk index for that demographic group during that marking period (Girvan et al., 2019). For example, given a risk index for Black/African American students of .2 and a risk index for White students of .15, the risk difference would be .05, meaning that Black/African American students had a .05 higher risk of receiving an ODR than White students (Girvan et al., 2019).

Student Demographic Groups

Student race (*American Indian, Asian, Black/African American, Latinx/Hispanic, Pacific Islander, White*), gender (*male, female*),¹ grade (*K, 1, 2, 3, 4, 5, 6, 7, 8*), and IEP status (*yes, no*) were linked to each ODR entered into SWIS. Data entry staff at each school used a school information system (SIS) information to create and maintain student records in SWIS. We used these variables to compute risk indices, ratios, and differences.

Implementation Fidelity

SWPBIS teams completed the BoQ (Kincaid et al., 2010) during the spring of 2019. The BoQ is a 53-item instrument used by teams to evaluate Tier I implementation fidelity. Teams reviewed each BoQ item and its rating options, with maximum values ranging from 1 to 3, and reached a consensus on the appropriate rating for each item. The BoQ has been found to evidence strong internal consistency, test-retest reliability, interrater reliability, and concurrent validity with other fidelity instruments (e.g., Cohen et al., 2007). Possible total scores range from 0 to 100 points, or 0% to 100%. The average BoQ score across the 27 schools was 78.56% ($SD = 9.93%$; range, 53%–93%), which exceeds the 70% criterion signifying high implementation (e.g., Kincaid et al., 2007). We included each school's total BoQ score (grand-mean-centered) as a predictor in analyses.

Covariates

To account for potential differences in implementation and outcomes according to (a) the timing of schools' enrollment into SWPBIS training and coaching as well as (b) the grades served by each school, we included as covariates each school's cohort year (dummy codes for 2016, 2017, and 2018, analyzed alongside 2015 as the intercept) as well as highest grades served (dummy codes for *grade 5* and *grade 6* analyzed alongside *grade 8* as the intercept).

Procedures

Upon enrollment into a SWPBIS cohort, each school developed a PBIS team and was matched with a coach from a district PBIS team that received training and consultation from the authors' institution, a center providing schools with systems-level training, consultation, and technical assistance housed within a national, not-for-profit behavioral health organization. District PBIS coaches generally

¹ The school district form provided two response options for gender, which we recognize does not begin to encompass all possible gender identities.

supported between six and eight schools at varied stages of implementation each year, namely by participating in each school team’s PBIS meetings (held at least monthly), supporting procedure and product development, and facilitating cohort PBIS trainings (held at least three times per year in each of the first 2 years of implementation). In their first year of cohort enrollment, each school’s PBIS team participated in three full-day trainings that followed the state’s prescribed training sequence and culminated in the development of manualized Tier I procedures and products. In their second year of cohort enrollment, each team participated in three full-day trainings related to SWIS, Team-Initiated Problem Solving (Todd et al., 2013), and school-wide interventions, respectively. In later years, schools participated in PBIS trainings on an as-needed basis given staff turnover or building-specific needs. It is important to note that the district did not adopt an intentional focus on equity in PBIS (e.g., completion of fidelity assessments tailored to evaluate elements of cultural responsiveness, representation of family and community members on teams, instruction on examining disproportionality in data, and training on measures to address disproportionality) into PBIS training and coaching systems until the 2019–2020 school year, the year following this study’s completion.

Data Analysis

In the case of race, we only descriptively summarized risk differences and ratios for select schools because few schools in the sample had adequate racial diversity within their student populations to enable a meaningful inferential analysis of racial disproportionality. For all other variables, we modeled outcomes longitudinally using an endpoint intercept within HLM 8.0 (Raudenbush et al., 2019) to examine risk ratios across years of SWPBIS implementation. Outcome data were risk ratios for 12 measurement occasions (4 marking periods across 3 years) across 27 schools. Visual inspections of

graphed data and curve estimation analyses identified trajectories for the grade as best characterized by linear terms and trajectories for gender and IEP status as best characterized by logarithmic terms; thereby, logarithmic terms were included to account for curvature in trajectory slopes as indicated.

Results

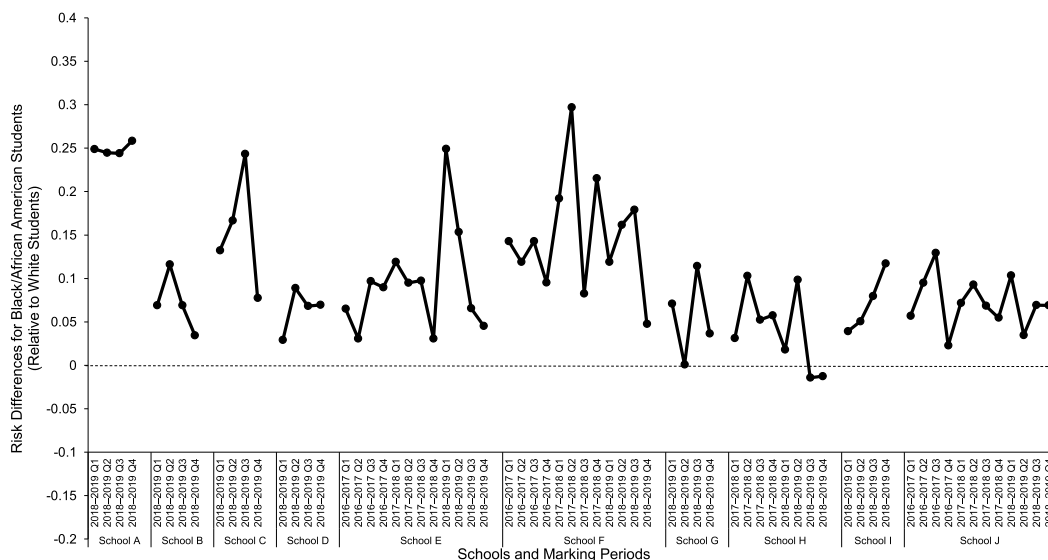
Preliminary analyses revealed that, across schools, students identified as Black/African American (*M*, risk ratio = 6.13), identified as Latinx/Hispanic (*M*, risk ratio = 2.18), identified as male (*M*, risk ratio = 2.12), and enrolled in each building’s upper three grades (*M*, risk ratio = 4.71) had higher risk indices than their counterparts at the outset of SWPBIS training and coaching. By contrast, students with IEPs had similar risk indices to students without IEPs (*M*, risk ratio = .99), yet we maintained our focus on students with IEPs given the potential for outcomes to differ across schools and time, particularly in relation to SWPBIS implementation.

Risk Trajectories According to Race

Most schools in this sample served a majority Black/African American (*n* = 19; for these 19, *M* = 82.58%, *SD* = 14.89%; range, 50.06%–95.87%) or Latinx/Hispanic student population (*n* = 3; for these 3, *M* = 65.01%, *SD* = 4.36%; range, 60.15%–68.58%). Per the guidance of Bollmer et al. (2014), we computed risk ratios and differences for 10 schools that served at least 10 students in each racial subgroup (*M* = 120, *SD* = 105.72, range: 11–303). The remaining 17 schools served fewer than 10 White students each year and were excluded from summaries and graphs.

Across the 68 occasions for these 10 schools, no White students received an ODR during 18 (26.47%) of these periods; thereby, risk ratios could not be computed for at least one occasion for 9 of the 10

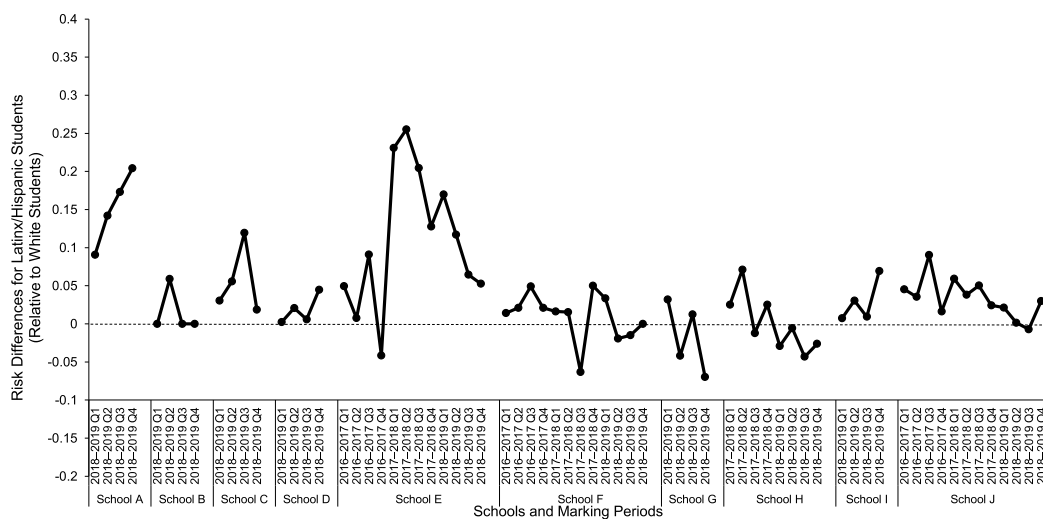
Figure 2
Risk Differences Across Marking Periods for Black/African American Students in a School Subsample (n = 10)



Note. Schools are ordered from lowest to highest Spring 2019 BoQ score (School A = 54%, B = 64%, C = 70%, D = 78%, E = 83%, F = 83%, G = 84%, H = 85%, I = 85%, J = 88%).

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 3
Risk Differences Across Marking Periods for Latinx/Hispanic Students in a School Subsample ($n = 10$)



Note. Schools are ordered from lowest to highest Spring 2019 BoQ score (School A = 54%, B = 64%, C = 70%, D = 78%, E = 83%, F = 83%, G = 84%, H = 85%, I = 85%, J = 88%).

schools. Thus, risk ratios proved to be a problematic approach to analyzing racial disproportionality with this subsample (see Girvan et al., 2019), and we graphed and interpreted risk differences in considering trends across time. Trajectories for Black/African American students are modeled in Figure 2 and for Latinx/Hispanic students are modeled in Figure 3. In general, schools with greater Tier I implementation fidelity, as evidenced by higher Spring 2019 BoQ scores, demonstrated (a) somewhat lower variability in risk difference across time for both Black/African American and Latinx/Hispanic students as well as (b) slightly more equitable outcomes for Latinx/Hispanic students compared to White students. Overall, however, risk difference for these racially minoritized students is generally sustained across SWPBIS implementation.

It should be noted that risk differences are not interpreted in the same way as risk ratios and cannot be compared to the values described below for students according to gender, grade, and IEP status. Risk ratios at the final measurement occasion for both Black/African American students ($M = 5.74$, median = 2.60, $SD = 6.96$, range: .65–23.94) and Latinx/Hispanic students ($M = 3.64$, median = 1.76, $SD = 5.14$, range: .26–17.27) suggested sustained significant disproportionality when applying many risk ratio thresholds, such as Pennsylvania's risk ratio threshold of 2.5 for disciplinary outcomes (Pennsylvania Department of Education, 2020).

Risk Trajectories According to Gender, Grade, and IEP Status

The overall expected risk ratio at spring 2019 was 2.17 for male students, 3.94 for students in each school's upper three grades, and .84 for students with IEPs. Risk ratio outcomes significantly varied across schools (male students, $\chi^2[26] = 108.71$, $p < .001$; students in upper grades, $\chi^2[23] = 219.86$, $p < .001$; students with IEPs, $\chi^2[24] = 199.63$, $p < .001$). No significant change in risk ratios across time was found for male students (linear, $t[26] = -.81$, $p = .42$; logarithmic, $t[26] = .64$, $p = .53$), students in upper grades (linear, $t[23] =$

.58, $p = .57$), or students with IEPs (linear, $t[24] = -1.45$, $p = .16$; logarithmic, $t[24] = 1.21$, $p = .24$). Moreover, change in risk ratios did not significantly vary across schools for male students (linear, $\chi^2[26] = 33.29$, $p = .15$; logarithmic, $\chi^2[26] = 38.24$, $p = .06$) but did significantly vary across schools for students in upper grades (linear, $\chi^2[23] = 94.31$, $p < .001$) and students with IEPs (linear, $\chi^2[24] = 46.25$, $p = .004$; logarithmic, $\chi^2[24] = 67.94$, $p < .001$).

Controlling for school cohort and grades served, BoQ scores significantly predicted spring 2019 risk ratios for male students ($t[20] = 2.45$, $p = .02$) but not for students in upper grades ($t[17] = -.01$, $p = .996$) or students with IEPs ($t[18] = 1.77$, $p = .09$). Further, BoQ scores did not significantly predict change in risk ratios across time for any group (male students, linear, $t[20] = -1.46$, $p = .16$; logarithmic, $t[20] = 1.96$, $p = .07$; students in upper grades, linear, $t[17] = -.33$, $p = .75$; students with IEPs, linear, $t[18] = -.41$, $p = .69$; logarithmic, $t[18] = 1.34$, $p = .20$). Thus, having higher spring 2019 BoQ scores was associated with male students' greater overrepresentation in spring 2019 ODRs, but no other significant associations between BoQ scores and risk ratio outcomes or trends were found (see Table 2).

Discussion

The current study underscores the importance of addressing overrepresentation and variables associated with sustained disproportionality over time. This study, with schools in a large, urban district, documented patterns of ODR overrepresentation that were sustained across years of SWPBIS implementation. Specifically, a small number of schools with racially heterogeneous student populations evidenced an overrepresentation of Black/African American and Latinx/Hispanic students in ODRs that, for the most part, sustained across years of SWPBIS implementation. Likewise, a larger sample of schools evidenced an overrepresentation of male students and students in upper elementary and middle school grades in ODRs that sustained across time. Tier I implementation fidelity,

Table 2
Models for Risk by Gender, Grade, and IEP Status

Fixed effects	Risk ratio, male (relative to female) students		Risk ratio, students in upper three grades (relative to students in lower three grades)		Risk ratio, students with (relative to students without) IEPs	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Intercept						
Intercept	2.56***	.41	2.97	2.57	.95*	.39
Slope, linear	-.05	.70	.03	.13	-.01	.06
Slope, logarithmic	.49	.41			.18	.34
2016–2017 cohort						
Intercept	-.44	.70	.72	3.43	1.49*	.63
Slope, linear	.05	.09	.04	.18	-.12	.11
Slope, logarithmic	-.66	.69			1.03	.56
2017–2018 cohort						
Intercept	.08	.57	1.66	3.23	-.18	.54
Slope, linear	-.08	.12	.05	.18	-.13	.10
Slope, logarithmic	.49	.72			.34	.52
2018–2019 cohort						
Intercept	-.27	.57	1.50	3.25	.10	.56
Slope, linear	-.56	.66	-.03	.24	-.04	.33
Slope, logarithmic	1.65	2.2			-.11	1.13
Highest grade, 5						
Intercept	-.83	.76	2.82	3.98	-1.52*	.70
Slope, linear	.05	.10	.14	.21	.05	.12
Slope, logarithmic	-1.02	.75			-1.08	.63
Highest grade, 6						
Intercept	-.70	.78	-3.09	3.51	-.11	.69
Slope, linear	.26	.14	-.67***	.20	.15	.15
Slope, logarithmic	-1.85	.89			-.06	.72
Spring 2019 BoQ						
Intercept	.06*	.03	-.0001	.13	.04	.02
Slope, linear	-.01	.004	-.002	.01	-.002	.01
Slope, logarithmic	.05	.03			.03	.02
Random effects	Variance component	χ^2	Variance component	χ^2	Variance component	χ^2
Intercept	.12	20.02	21.64	172.17***	.66	96.87***
Slope, linear	.001	19.77	.05	79.84***	.02	53.43**
Slope, cubic						
Slope, logarithmic	.002	18.23			.44	39.17**

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

as measured by the BoQ, was found to be minimally relevant in elucidating risk ratio trajectories, and, in fact, higher fidelity predicted male students' greater overrepresentation in ODRs.

Results reveal persistent inequitable outcomes for students, which must be addressed as a social justice imperative. Rates of student behavior were not directly assessed in this study, such that the extent of staff implicit bias (e.g., toward racial/ethnic minoritized and/or male students) or varying tolerance levels (e.g., toward students in upper grades) cannot be inferred. Given, for example, prior research suggesting that discipline disproportionality is not explained by racial differences in student behavior (Bradshaw et al., 2010), it is likely that factors related to students and staff (e.g., perceptions, attitudes, and culturally bound socialization processes), as well as their school contexts (e.g., policies and procedures), are intersecting to produce inequitable outcomes. Although MTSS and SWPBIS, specifically, have been identified as comprehensive frameworks with potential for reducing discipline disproportionality (e.g., Tobin & Vincent, 2011), others before us have noted that, without a focus on the educational and relational interplay of societal constructs (e.g., race, gender, and gender roles; Carter et al., 2017) and

“culturally conscious implementation” (Gregory et al., 2017, p. 254), these frameworks may intensify disproportionality for minoritized and otherwise disadvantaged groups. It is beyond the scope of this article to comprehensively describe what adopting this focus and type of implementation would entail. We advise readers to review Gregory et al.'s (2017) principles related to prevention and intervention for increasing equity in school discipline (e.g., bias-aware classrooms, the inclusion of student and family voice, reintegration of students following absence or conflict) as well as McIntosh, Ellwood, et al.'s (2018) four-step process for leveraging data-based decision making in pursuit of discipline equity.

Two factors must be considered in contextualizing study findings. First, we caution the use of ODRs as a school reform evaluation approach. Although this metric has been found to demonstrate psychometric evidence (e.g., Pas et al., 2011) and has been employed in randomized controlled trials of SWPBIS implementation (see Lee & Gage, 2020), the reliability of ODRs may be threatened by many factors, among them systems-level initiatives. For example, we have seen schools infrequently or inconsistently using ODRs prior to adopting SWPBIS—such as with staff completing ODRs

contingent upon (a) the perception that office disciplinary action is “warranted” rather than (b) the direct observation of an operationally defined behavior of concern. Upon adoption of SWPBIS and, for example, the Team-Initiated Problem Solving model (Todd et al., 2011, 2013), the function of ODRs in a setting may change (i.e., from coordinating discipline to recording behavior for problem-solving), and schools may increase their consistency in completing ODRs. In relation to the current study, when more ODRs are completed (e.g., upon high-fidelity SWPBIS implementation), there may be more opportunities for patterns of disproportionality to emerge. Thus, disproportionality might not have truly increased or sustained across years of SWPBIS implementation; instead, previously undocumented levels of disproportionality might have been accentuated with better record-keeping.

Second, this study focused exclusively on Tier I implementation: Participating schools were not implementing advanced tiers with fidelity during the collection of data. The success of an MTSS framework is predicated on targeted and intensive supports layered upon universal supports for students demonstrating greater needs, and these advanced-tier supports may be necessary to realizing truly equitable outcomes in schools. For example, the current study identified that students with IEPs were less likely to receive ODRs than their peers without IEPs, which deviates from prior research (e.g., U.S. Government Accountability Office, 2018). Although special education services cannot be equated with advanced-tier supports, students in special education may be more likely to receive differentiated supports for their academic and behavioral needs, which, in turn, may result in fewer behaviors of concern and ODRs. This phenomenon, thereby, might approximate what could be achieved through the provision of well-matched advanced-tier interventions. It is certainly problematic, however, when a Tier I system is differentially effective for students of varying demographics—this may suggest a misalignment of the system with the values and needs of all students.

In addition to the aforementioned limitations (i.e., the flawed nature of ODRs in evaluating school reform, the limited focus on universal tier implementation), other constraints limit the extent to which study findings can be used to inform future research, practice, and policy. Perhaps most notably, we had a limited sample size (i.e., $n = 27$ level-2 units), which resulted in underpowered multi-level analyses. In addition, few schools in our sample served racially diverse student populations, which precluded our application of inferential statistics to examine data trends according to race. Moreover, our measurement of Tier I implementation relied upon school teams’ self-assessment overseen by their district PBIS coach, as schools consistently completed the BoQ assessment across years but were subject to various other fidelity assessments completed by external evaluators according to changing guidelines and requirements of the state PBIS network. Our findings may have been masked by our analytic foci on (a) overall ODRs, rather than minor and major infractions separately or objective and subjective infractions separately (e.g., Girvan et al., 2017); (b) overall BoQ scores, rather than specific scale or item scores; and (c) BoQ scores at one point in time, rather than differences in scores across time. These and other limitations implore the need for additional research examining trends in discipline disproportionality across schools’ implementation of SWPBIS. Future research would help clarify the extent to which and conditions under which high-fidelity SWPBIS

implementation may be associated with improved equity in discipline practices.

In sum, schools may be overrepresenting student demographic groups in ODRs that extend beyond race to include students who are male, in schools’ upper grades, and not participating in special education. Results of this study suggest Tier I SWPBIS practices may be inadequate in reducing discipline disproportionality without equity-focused (e.g., culturally responsive) implementation and the provision of advanced-tier interventions.

References

- Assistance to States for the Education of Children with Disabilities, 34 CFR § 300.647. (2018). *Determining significant disproportionality*.
- Anyon, Y., Jenson, J., Altschul, I., Farrar, J., McQueen, J., Greer, E., Downing, B., & Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review, 44*, 379–386. <https://doi.org/10.1016/j.childyouth.2014.06.025>
- Appleyard, K., Egeland, B., van Dulmen, M. H. M., & Alan Sroufe, L. (2005). When more is not better: The role of cumulative risk in child behavior outcomes. *The Journal of Child Psychology and Psychiatry, and Allied Disciplines, 46*, 235–245. <https://doi.org/10.1111/j.1469-7610.2004.00351.x>
- Ashworth, E., & Humphrey, N. (2020). More than the sum of its parts: Cumulative risk effects on school functioning in middle childhood. *British Journal of Educational Psychology, 90*, 43–61. <https://doi.org/10.1111/bjep.12260>
- Bollmer, J. M., Bethel, J. W., Munk, T. E., & Bitterman, A. R. (2014). *Methods for assessing racial/ethnic disproportionality in special education* [Technical report]. Westat.
- Boneshefski, M. J., & Runge, T. J. (2014). Addressing disproportionate discipline practices within a school-wide positive behavioral interventions and supports framework: A practical guide for calculating and using disproportionality rates. *Journal of Positive Behavior Interventions, 16*, 149–158. <https://doi.org/10.1177/1098300713484064>
- Bradshaw, C. P., Mitchell, M. M., O’Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *Journal of Educational Psychology, 102*, 508–520. <https://doi.org/10.1037/a0018450>
- Carter, P. L., Skiba, R., Arredondo, M. I., & Pollock, M. (2017). You can’t fix what you don’t look at: Acknowledging race in addressing racial discipline disparities. *Urban Education, 52*, 207–235. <https://doi.org/10.1177/0042085916660350>
- Cartledge, G., & Kourea, L. (2008). Culturally responsive classrooms for culturally diverse students with and at risk for disabilities. *Exceptional Children, 74*, 351–371. <https://doi.org/10.1177/001440290807400305>
- Cohen, R., Kincaid, D., & Childs, K. E. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the Benchmarks of Quality. *Journal of Positive Behavior Interventions, 9*, 203–213. <https://doi.org/10.1177/10983007070090040301>
- Educational and Community Supports. (n.d.). *SWIS office referral form definitions*. <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjj2bXMka3sAhXUmHIEHaa8DIUQFjABegQIBBAC&url=https%3A%2F%2Fwww.pbisapps.org%2Fresources%2Fswis%2520Publications%2Freferral%2520Category%2520Definitions.rtf&usq=AOvVaw3uskpdG1EOcP1VK8yIyvD3>
- Gage, N. A., Whitford, D. K., & Katsiyannis, A. (2018). A review of schoolwide positive behavior interventions and supports as a framework for reducing disciplinary exclusions. *The Journal of Special Education, 52*, 142–151. <https://doi.org/10.1177/0022466918767847>
- 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, 392–404. <https://doi.org/10.1037/spq0000178>
- Girvan, E. J., McIntosh, K., & Smolkowski, K. (2019). Tail, tusk, and trunk: What different metrics reveal about racial disproportionality in school discipline. *Educational Psychologist*, 54, 40–59. <https://doi.org/10.1080/00461520.2018.1537125>
- Gregory, A., Skiba, R. J., & Mediratta, K. (2017). Eliminating disparities in school discipline: A framework for intervention. *Review of Research in Education*, 41, 253–278. <https://doi.org/10.3102/0091732X17690499>
- Horner, R. H., Sugai, G., & Fixsen, D. L. (2017). Implementing effective educational practices at scales of social importance. *Clinical Child and Family Psychology Review*, 20, 25–35. <https://doi.org/10.1007/s10567-017-0224-7>
- Irvin, L. K., Horner, R. H., Ingram, K., Todd, A. W., Sugai, G., Sampson, N. K., & Boland, J. B. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle schools: An empirical evaluation of validity. *Journal of Positive Behavior Interventions*, 8, 10–23. <https://doi.org/10.1177/10983007060080010301>
- Kincaid, D., Childs, K., Blase, K. A., & Wallace, F. (2007). Identifying barriers and facilitators in implementing schoolwide positive behavior support. *Journal of Positive Behavior Interventions*, 9, 174–184. <https://doi.org/10.1177/10983007070090030501>
- Kincaid, D., Childs, K., & George, H. (2010). *School-wide benchmarks of quality* [Instrument]. University of South Florida.
- Lee, A., & Gage, N. (2020). Updating and expanding systematic reviews and meta-analyses on the effects of school-wide positive behavior interventions and supports. *Psychology in the Schools*, 57, 783–804. <https://doi.org/10.1002/pits.22336>
- Losen, D., Hodson, C., Keith, M. A., II, Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?* [Report]. The Center for Civil Rights Remedies. https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/are-we-closing-the-school-discipline-gap/AreWeClosingTheSchoolDisciplineGap_FINAL221.pdf
- Losen, D. J., & Martinez, T. E. (2013). *Out of school and off track: The overuse of suspension in American middle and high schools*. The Civil Rights Project at UCLA, The Center for Civil Rights Remedies.
- May, S., Ard, W., Todd, A., Horner, R., Glasgow, A., Sugai, G., & Sprague, J. (2020). *School-wide information system* (version 7.4.5 b109) [Software]. University of Oregon.
- McIntosh, K., Ellwood, K., McCall, L., & Girvan, E. J. (2018). Using discipline data to enhance equity in school discipline. *Intervention in School and Clinic*, 53, 146–152. <https://doi.org/10.1177/1053451217702130>
- McIntosh, K., Gion, C., & Bastable, E. (2018). *Do schools implementing SWPBIS have decreased racial and ethnic disproportionality in school discipline?* [Research brief]. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports.
- Pas, E. T., Bradshaw, C., & Mitchell, M. M. (2011). Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychology in the Schools*, 48, 541–555. <https://doi.org/10.1002/pits.20577>
- Pennsylvania Department of Education. (2020). *Significant disproportionality* [Technical guide]. https://www.pattan.net/getmedia/fe2ac034-aeaa-4ab0-ba22-cc74269e4f87/Sig_Dispro_Overview_CalcFAQ20320
- Putnam, R., McCart, A., Griggs, P., & Choi, J. H. (2009). Implementation of school-wide positive behavior support in urban settings. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of positive behavior support* (pp. 443–463). Springer. https://doi.org/10.1007/978-0-387-09632-2_19
- Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2019). *HLM 8.0 for Windows* [Software manual]. Scientific Software International, Inc.
- Rutter, M. (1979). Maternal deprivation, 1972–1978: New findings, new concepts, new approaches. *Child Development*, 50, 283–305. <https://doi.org/10.2307/1129404>
- Skiba, R. J., Horner, R. H., Chung, C., Rausch, M. K., May, S., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, 40, 85–107. <https://doi.org/10.1080/02796015.2011.12087730>
- Tobin, T. J., & Vincent, C. G. (2011). Strategies for preventing disproportionate exclusions of African American students. *Preventing School Failure*, 55, 192–201. <https://doi.org/10.1080/1045988X.2010.532520>
- Todd, A. W., Horner, R. H., Newton, J. S., Algozzine, R. F., Algozzine, K. M., & Frank, J. L. (2011). Effects of team-initiated problem solving on decision making by schoolwide behavior support teams. *Journal of Applied School Psychology*, 27, 42–59. <https://doi.org/10.1080/15377903.2011.540510>
- Todd, A. W., Newton, J. S., Algozzine, K., Horner, R. H., & Algozzine, B. (2013). *Team-Initiated Problem Solving (TIPS II) training manual*. University of Oregon, Educational and Community Supports.
- U.S. Department of Education. (2018). *Office for civil rights, civil rights data collection, 2015–16*. <https://www2.ed.gov/about/offices/ist/ocr/docs/schoolclimate-and-safety.pdf>
- U.S. Government Accountability Office. (2018). *Discipline disparities for Black students, boys, and students with disabilities* [Report]. <https://www.gao.gov/assets/700/690828.pdf>
- Vincent, C. G., English, J., Girvan, E. J., Sprague, J. R., & McCab, T. (2016). School-wide Positive and Restorative Discipline (SWPRD): Integrating school-wide positive behavior interventions and supports and restorative discipline. In R. Skiba, K. Mediratta, & K. M. Rausch (Eds.), *Inequality in school discipline: Research and practice to reduce disparities* (pp. 115–134). Palgrave MacMillan.

Received June 19, 2020

Revision received December 15, 2020

Accepted December 23, 2020 ■