



Suspended Again: The Racialized Consequences of a 9th Grade Suspension on Future Suspension Patterns

Benjamin W. Fisher^{1,2}  · Stephanie A. Wiley³ · Anne McGlynn-Wright⁴

Accepted: 14 May 2021

© Springer Science+Business Media, LLC, part of Springer Nature 2021, corrected publication 2021

Abstract

Although prior research has linked school-based punishment to a series of negative consequences, little is known about how being punished in school predicts future school-based punishment. To address this, the current study examines the extent to which being suspended in 9th grade predicts subsequent suspensions within the same school. Using stereotype congruence theory as a framework, we examine differences by race (black versus white) and household income. The data are drawn from three cohorts of four-wave annual administrative data from a large urban school district in the Midwestern USA ($N = 11,006$). Findings indicate that being suspended in 9th grade is associated with higher odds of subsequent suspension and a greater number of subsequent suspensions, but not a greater number of days per suspension. Black students suspended in 9th grade were particularly likely to experience more subsequent suspensions. Further, these racial differences are not driven by household income measures. These findings indicate that racially disparate school punishment practices have cascading effects for black students.

Keywords Suspension · School discipline · Punishment · Discipline gap · Educational equity · Racial disparities

U.S. public schools saw a precipitous rise in exclusionary discipline as a form of school-based punishment in the 1980s and 1990s, buoyed by rising fear of school violence and federal monies premised on the adoption of crime control measures (Hirschfield, 2008; Kupchik, 2010; Simon, 2007). Exclusionary discipline, which temporarily or permanently removes children from school or class, has been shown to incur damage on students' educational trajectories, leading to disinvestment in education, poor grades, and higher rates of school pushout (e.g., Arcia, 2006; Skiba, Arredondo, et al., 2014). These collateral consequences of exclusionary discipline may redirect some youth toward the criminal justice system, with increased likelihood of being arrested

and incarcerated later in life (Hemez et al., 2020; Monahan et al., 2014; Mowen & Brent, 2016; Wolf & Kupchik, 2017).

In the past decade, scholars, activists, and policymakers have increasingly drawn attention to the negative consequences of exclusionary discipline and further underscored the disproportionate punishment of youth of color. The latest national-level data found that although black students constituted only 16% of the population of students enrolled in public schools, they were 39% of the students who received at least one out-of-school suspension (USDE Office of Civil Rights, 2019). By contrast, white students constituted 49% of enrolled students, but were only 32% of the students who received an out-of-school suspension (USDE Office of Civil Rights, 2019). Even after controlling for individual- and school-level characteristics, black, Latinx, and Indigenous students bear the brunt of disciplinary practices relative to their white peers (Pearman et al., 2019; Skiba et al., 2011). Despite a substantial body of empirical literature showing that youth of color are disproportionately suspended and expelled, it is unknown how these disparities unfold over time to affect their experiences of future school discipline.

Research has recently begun to examine the impact that school discipline has on future school punishments (Wiley et al., 2020). In a process known as secondary sanctioning,

✉ Benjamin W. Fisher
bwfisher@fsu.edu

¹ College of Criminology & Criminal Justice, Florida State University, Tallahassee, FL, USA

² Department of Criminal Justice, University of Louisville, Louisville, KY, USA

³ Simon Fraser University, Burnaby, Canada

⁴ Newcomb Institute, Tulane University, New Orleans, LA, USA

initial punishments may lead to subsequent punitive responses because of increased surveillance and reactivity (Liberman et al., 2014; Rios, 2011; Sampson & Laub, 1997). Teachers and school administrators may have heightened awareness of and increasingly punitive responses to youth who have been deemed “troublemakers”. However, these patterns are also highly racialized. Qualitative research suggests that a secondary sanctioning pattern is particularly pernicious for black boys—who are often singled out for their behavior and thus elicit more stringent scrutiny in the future (Ferguson, 2000; Gansen, 2020). Research within the criminal justice system further finds long-term criminal justice consequences for black youth who are exposed to police through police-initiated stops, while white youth experience no such long-term criminal justice responses (McGlynn-Wright et al., 2020). These findings suggest that the combination of race and punishment interact to create differential subsequent punishments.

Building on this literature, the purpose of the current study is to examine black–white differences in the extent to which being suspended in 9th grade predicts being suspended in subsequent years within the same school. This study contributes to the literature in two ways. First, while the existing research highlights the negative impact that exclusionary discipline has on educational outcomes and justice system involvement, we examine an understudied outcome: subsequent school punishment. If youth who experience a single exclusionary discipline event are more likely to receive subsequent punishment, this could compound the risk for future negative outcomes (Mowen & Brent, 2016; Sampson & Laub, 1997; Wiley et al. 2020). Highlighting the cascading consequences of school punishment is useful for mitigating both the short- and long-term negative impacts of exclusionary discipline for all youth. Second, we build on the widely acknowledged finding that black youth disproportionately experience school discipline by examining whether, once suspended, patterns of future suspension vary for white and black students. If black youth are already disadvantaged due to differential disciplinary practices, they may be at increased risk of receiving additional school punishment, which mortgages future achievements. This focus on black–white differences further emphasizes the need to address the discriminatory nature of school discipline practices.

Race and Patterns of School Discipline

Although the education system is seen as a cornerstone of social and economic advancement, it has also become an institution of punishment, control, and oppression (Annamma, 2017; Morris, 2016)—governed by many of the same principles that guide the criminal justice system

(Rios, 2011; Shedd, 2015; Simon, 2007). Beyond the consequences that exclusionary discipline practices have for all students, the punitive turn in school discipline has exacerbated many existing educational inequities, particularly among schools in low-income areas. In addition to grappling with overcrowding, teacher and student retention, poor academic performance, and technological issues, impoverished school districts must also contend with a limited number of resources for dealing with students. A lack of counselors or restorative practices to address student behaviors perceived as non-normative means that exclusionary discipline and even justice system involvement are often the most efficient available responses (Ferguson, 2000; Kupchik, 2010). In already disadvantaged schools, then, reliance on punitive responses contributes to a cyclical pattern whereby academic progress is affected, funding is further reduced, and criminalizing responses are increasingly used (Kozol, 1991; Simon, 2007).

One consequence of the overreliance on exclusionary discipline in poorer schools is that students of color are disproportionately affected by these practices. In fact, the empirical research on school discipline consistently shows that, even after controlling for socioeconomic factors, schools administer significantly more exclusionary discipline to black students (within schools) and in schools with larger proportions of black students (between schools; Edwards, 2016; Jacobsen et al., 2019; Payne & Welch, 2010; Petras et al., 2011; Skiba et al., 2011). One explanation for the racialized pattern of school discipline is that schools serving more non-white students also tend to have fewer resources. Indeed, class and race are deeply intertwined, and this relationship is particularly noticeable within the educational sphere (Kozol, 1991). Above and beyond the relationship between race and class, however, racism may uniquely impact schools’ use of formal social control. Exploring the unique effect of race in educational inequality, Lewis and Diamond (2015) observed a racially diverse, well-funded high school in the Midwest United States, noting that it maintained a two-tier academic system—one of Advanced Placement classes primarily serving white students and general education courses primarily serving black students. When it came to discipline, the authors found that suspensions tended to be leveled more at black students than white students. Similarly, Skiba and colleagues (2002) assessed the disciplinary records of over 11,000 students in 19 middle schools across several midwestern cities. After accounting for socioeconomic status, the researchers reported that black youth were still more likely than their white counterparts to be disciplined for the same behavior.

These findings have prompted the need for explanations of racialized school discipline practices that are independent of socioeconomic status. Although it may be tempting to explain the racial gap in school discipline by pointing to

differential racial patterns of student behaviors, this argument has little empirical support. In an analysis of nationally representative data that controlled for prior assessments of student behavior, Huang (2020) found that differences in student behaviors could not fully account for the existence of the nationwide racial discipline gap. Related research has found that racial discipline gaps are primarily driven by characteristics of school administrators and their philosophies about discipline, not by characteristics of students or their behaviors that led to their experience of school discipline (Skiba, Arredondo, et al., 2014; Skiba, Chung, et al., 2014). While there is clear evidence that student behaviors are associated with patterns of student discipline (Huang, 2020), they are insufficient in explaining the racial gap in school discipline.

Moving beyond a focus on student behavior, one prominent explanatory framework is racial stereotyping (Edwards, 2016; Ferguson, 2000; Skiba et al., 2002, 2011). According to this framework, teachers and school administrators are more likely to view students of color as problematic due to conditioned beliefs about the types of behaviors youth of color participate in. In the next section, we use stereotype congruence theory as a framework for understanding the relationship between student race and patterns of exclusionary discipline over time. Although this study does not empirically examine stereotypes or student behaviors that might lead to stereotype congruence, this theoretical framework may be useful for directing future research that investigates this phenomenon or on-the-ground interventions that seek to interrupt racialized patterns of school discipline.

Stereotype Congruence and Stereotypes of Black Criminality

In the United States, race plays a central and enduring role in everyday life from policy development to interpersonal interactions, and schools are no exception (Omi & Winant, 2014). Stereotype congruence theory offers one way of understanding racial differences in patterns of school punishment. This theory suggests that identical behaviors of black and white individuals will result in different interpretations based on the extent to which the behavior is stereotype congruent or divergent (e.g., Eberhardt et al., 2006; Goff et al., 2008; Valentino et al., 2002). Stereotypes are assumptions that distinguishable groups hold specific traits, and can influence individuals' interpretations of ambiguous behavior (Hilton & von Hippel, 1996) and even influence memory processes (Rothbart et al., 1979). Stereotype congruence is when the behavior of an individual is seen as consistent with prevailing stereotypes. One prevailing stereotype throughout the history of the United States is that of black criminality (Muhammad, 2010), which holds that black individuals are

more likely to engage in crime. In the context of the educational setting, stereotype congruence might appear when teachers and school administrators punish students on the basis of preconceived behavioral expectations; if school personnel believe that black students are more disruptive or delinquent than their white peers, then black students are more likely to receive punishments even if their behavior is similar to white students'.

Experimental research with educators provides support for stereotype congruence theory as it relates to teachers' expectations of black students. For example, one experiment tracked the eye movements of preschool teachers watching videos of a group of children playing; when the teachers were prompted to expect challenging behaviors from the children—even though no challenging behaviors were actually present—their eyes tracked more on the black children, expecting the black children to misbehave in a manner congruent with stereotypes of black criminality (Gilliam et al., 2016). In this case, it is likely that educators' stereotypes linking blackness to misbehavior directed their scrutiny and attention toward black students when they were primed to expect misbehavior. Another experiment found that pre-service teachers who read a vignette about a defiant student were more likely to expect future problem behaviors when the vignette was about a black student rather than a white student (Kunesh & Noltemeyer, 2019).

In addition to contextualizing teachers' perceptions of student behaviors, stereotype congruence theory is helpful for understanding how teachers respond to behaviors differently on the basis of student race. In one factorial survey experiment, teachers read vignettes about student misbehavior that were varied by (a) student race (using stereotypically black and white names), and (b) the number of infractions the students had committed (Okonofua & Eberhardt, 2015). Teachers recommended more severe discipline for black students than white students when there were multiple infractions, suggesting that teachers viewed black students as more likely to continue to have infractions and in need of punishment to deter future misbehavior.

Observational research corroborates this trend. For example, Gansen's (2020) ethnographic observations in three preschools showed that black boys were the most likely recipients of punitive disciplinary measures, including time-outs and physical restraint. These responses, Gansen (2020) argues, identified the "bad" and "good" kids, which reproduced societal stereotypes and hierarchies in the classroom. Other research shows that schools are more likely to punish black students, but give therapy or medication to their white peers (Ramey, 2018). Together, these studies' findings provide support for stereotype congruence by showing that teachers expect black students to behave in ways consistent with stereotypes that link blackness to criminality, but also by demonstrating that teachers believe harsh discipline is required to curb the

perceived misbehavior of black students more so than white students.

There is reason to believe that initial incidents of exclusionary discipline more frequently lead to subsequent discipline among black students. When school personnel draw upon common cultural scripts linking blackness to violence or disruptive behavior, they may view black youth who have already been suspended as stereotype congruent. That is, when a black student is suspended it may signal to school personnel that the student has done something congruent with the stereotype of black criminality, leading to increased surveillance and continued punishment during the student's time in the school. Ferguson's (2000) ethnographic fieldwork among black elementary school boys documents these stereotyping patterns. Minor misbehavior, when initiated by black boys who had already experienced disciplinary infractions, was rarely ignored by teachers. Indeed, each incident was used as "an occasion for a teacher to (re)mark the identity of a boy as disruptive" (Ferguson, 2000, p. 92).

Current Study

Prior research has made three parallel phenomena clear. First, black students are far more likely to experience school-based punishment than their white peers. Second, educators often view black students in ways consistent with stereotypes of criminality. Third, experiencing punishment is likely to lead to subsequent punishment, particularly for black youth. An examination of suspension patterns across multiple years is necessary to understand how to mitigate associated harms for all students, while our focus on racial disparities uncovers additional ways in which black youth are disproportionately impacted by school punishments. As such, the current study addresses this gap in the literature by examining the use of suspensions within a longitudinal sample of students from a large, diverse, urban school district. This study is guided by the following research question: Is being suspended in 9th grade associated with future suspensions differently for white and black students? Additionally, because class is a salient construct distinct from (but often highly correlated with) race, we investigate a parallel research question: Is being suspended in 9th grade associated with future suspensions differently for students from low-income families? Together, answering these research questions will allow us to better understand differential patterns of within-school disparities in punishment.

Method

Data

The data used in this study come from the administrative records of a single large urban school district in the Midwest

United States. In the original panel data set, information about students in kindergarten through 12th grade were included for six consecutive years (i.e., 2010–11 through 2015–16) in such a way that students' data could be linked over time. This included data from 149,494 students nested in 115 schools. For the purposes of this study, we restricted the sample to the three cohorts of students that had data for each of their four years of high school. Because we were interested in students' suspension patterns during their time within the same school, we removed observations from students who moved schools. Additionally, given our interest in black-white differences, we removed observations from students who were not identified as white or black in the dataset. This yielded a final analytic sample of 11,006 students in 21 schools who provided four waves of data representing their 9th–12th grade years.

Measures

Dependent Variables

Three dependent variables were used in this study to examine the robustness of our results across distinct operationalizations of follow-up discipline. First, to examine the likelihood of a student ever being suspended after 9th grade, we used a dichotomous measure of whether a student was suspended in 10th, 11th, or 12th grade (0 = No, 1 = Yes). Second, to examine the frequency of suspension, we used a count of the number of times a student was suspended in grades 10, 11, and 12. This measure is more sensitive than the likelihood of being suspended, allowing for the identification of the total number of suspensions after 9th grade rather than a binary indicator of whether a suspension occurred. Third, to measure the average length of suspension, we divided the total number of days suspended by the total number of suspensions to calculate the mean number of days per suspension in grades 10, 11, and 12. School personnel have discretion in the length of suspensions that they administer to students, and administer longer suspensions for behaviors they consider more severe. As such, this measure of the number of days per suspension captures the severity of punishment.

Independent Variable

The focal independent variable in this study was a dichotomous measure of whether a student was suspended in 9th grade (0 = No, 1 = Yes).

Moderators

We included two potential moderators to examine whether the relationship between being suspended in 9th grade and

suspension outcomes in future years depended on student characteristics. First, we incorporated a measure of student race to examine racial disparities in school discipline (0 = white, 1 = black). Second, we included a measure of whether a student was eligible for free or reduced-price lunch (FRPL) in 9th grade (0 = No, 1 = Yes). FRPL eligibility is based on household income and is thus an indicator of poverty, which allows us to determine whether race and socioeconomic status exert independent effects on subsequent exclusionary discipline. We note that although these moderators are student-level measures, we do not take them to be explanatory variables in and of themselves. Rather, we understand them as indicators of which students are likely to experience the benefits and harms of the prevailing structures of racism and classism in the school environment (Holland, 2008).

Control Variables

We included control variables at both the student and school level. To adjust for possible year-to-year differences in practices around suspension, we included a measure of the cohort to which each student belonged (1 = 2010–11, 2 = 2011–12, 3 = 2012–13). Student race and FRPL eligibility were used as control variables in models where they were not used as moderators. At the school level, we included racial composition measured as the percent of the student body that was identified as black in the administrative data. We also included a measure of school-level poverty operationalized as the percent of the student body that was FRPL eligible.

Data Analysis

To assess the relationships between being suspended in 9th grade and the three suspension outcomes, we estimated a series of multilevel models to account for the non-independence of observations (i.e., students nested within schools). We first estimated null models to assess the appropriateness of using a multilevel modeling strategy. Results indicated a substantial proportion of the variability in the outcome was attributable to school-level characteristics, making multilevel modeling an appropriate choice. Specifically, the intraclass correlation coefficient (ICC)—a measure of the proportion of total variability that occurs between clusters—for the dichotomous measure of suspension after 9th grade was 0.14; the ICC for the count of suspensions after 9th grade was 0.09; and the ICC for the number of days per suspension after 9th grade was 0.12. Next, we added all the independent variables to the models to assess the relationship between being suspended in 9th grade and each of the three outcomes. Finally, we examined race and FRPL eligibility as moderators. For the dichotomous measure of suspension after 9th grade, we used a logit link function. For

the count measure of suspension after 9th grade, we used a log link function. For the continuous measure of the number of days per suspension after 9th grade, we used an identity link function. All student-level control variables were group-mean-centered and all school-level control variables were grand-mean centered. Because we were interested in within-school estimates (i.e., differential experiences of suspension within the same school), all models used fixed slopes and random intercepts.

To assess the extent to which the relationship between being suspended in 9th grade and each of the three outcomes was moderated by student race and FRPL eligibility, we created multiplicative interaction terms between the measures of 9th grade suspension and race and FRPL eligibility, respectively. These interactions were examined in separate models. Because the *p*-values associated with the regression coefficients of interaction terms are untrustworthy in non-linear models, we followed the guidance of Mize (2019) and examined the predicted probabilities and counts associated with each interaction and the statistical significance of the second differences. These second differences refer to, for example, the difference between (a) the black-white discrepancy in the predicted probability of being suspended after 9th grade for students who *were not* suspended in 9th grade, and (b) the black-white discrepancy in the predicted probability of being suspended after 9th grade for students who *were* suspended in 9th grade. The statistical significance of this second difference was used to determine the statistical significance of the interaction.

We modeled the outcomes in three additional ways to assess the extent to which our findings were sensitive to the choices made in our analyses. First, rather than allowing for random intercepts, we used fixed intercepts so that our models represented a school fixed effects analysis. Second, we group-mean-centered the variables that constituted the multiplicative interaction term. Third, we used single-level logistic, negative binomial, and ordinary least squares regression modeling strategies for each outcome, while using school dummy codes and cluster robust standard errors. The findings across all of these modeling choices were substantively similar, so we present only the findings from the modeling strategy described in the previous paragraph. All data analyses were conducted in Stata 15.

Results

Table 1 displays descriptive statistics for all the variables used in the analyses. As shown, 39.2% of the sample was black and 88.6% was FRPL eligible. At the school level, there was an average of 41.8% black students per school (range = 22.8 – 92.0%) and 88.0% who were FRPL eligible (range = 58.6 – 100%). About 7.0% of the sample was

Table 1 Descriptive statistics for key variables (N=11,006)

	M	SD	Min	Max
Dependent Variables				
Ever suspended after 9th grade	0.14	0.35	0.00	1.00
Times suspended after 9th grade	0.25	0.85	0.00	14.00
Days per suspension after 9th grade	3.45	1.91	0.33	10.00
Student-Level Predictors				
Suspended in 9th grade	0.07	0.26	0.00	1.00
Race (white=0, black=1)	0.39	0.49	0.00	1.00
FRPL eligible	0.89	0.32	0.00	1.00
School-Level Predictors				
Percent black	39.20	15.71	22.79	92.00
Percent FRPL eligible	86.05	13.47	58.56	100.00

FRPL free or reduced-price lunch

suspended in 9th grade and 14.1% was suspended after 9th grade. On average, students experienced 0.25 suspensions after 9th grade and were suspended an average 0.78 days.

Among students who were suspended, the average number of days suspended after 9th grade was 3.45.

As shown in Table 2 Panel A, two of the models examining the relationship between being suspended in 9th grade and each measure of suspension after 9th grade indicated a significant positive relationship. Specifically, students who were suspended in 9th grade had 482% higher odds of being suspended in subsequent grades (OR = 4.82, $p < 0.001$, 95% CI [4.05, 5.73]). Similarly, students who were suspended in 9th grade had 458% higher incident rates of the total number of suspensions (IRR = 4.58, $p < 0.001$, 95% CI [3.92, 5.34]). However, there was not a statistically significant relationship between being suspended in 9th grade and the number of days per suspension after 9th ($b = -0.11$, $p = 0.337$, 95% CI [-0.33, 0.11]). Additionally, consistent with prior literature on racial disparities in school discipline, black students had higher odds of ever being suspended in 10th–12th grade (OR = 2.06, $p < 0.001$, 95% CI [1.81, 2.34]) and higher incident rates of the number of suspensions (IRR = 2.06, $p < 0.001$, 95% CI [1.81, 2.34]). Student race was not related

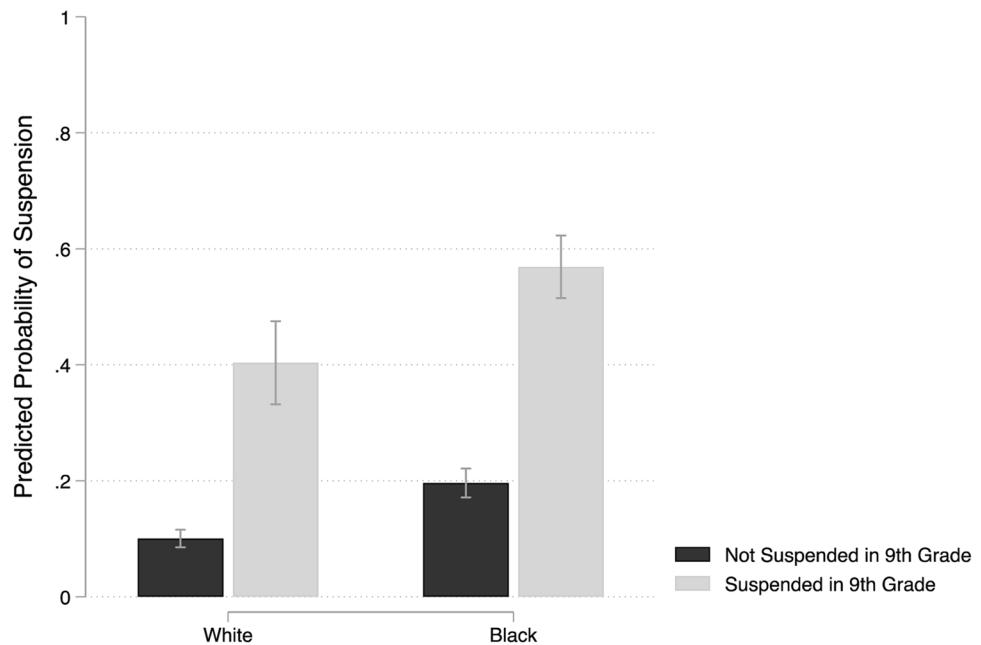
Table 2 Regression models predicting suspension (N=11,006)

	Ever suspended after 9th grade			Times Suspended After 9th grade			Days Per Suspension after 9th Grade		
	OR	p	95% CI	IRR	p	95% CI	b	p	95% CI
Panel A: No moderator									
Suspended in 9th grade	4.82	<.001	[4.05, 5.73]	4.58	<.001	[3.92, 5.34]	-0.11	.337	[-0.33, 0.11]
Black	2.21	<.001	[1.92, 2.55]	2.06	<.001	[1.81, 2.34]	-0.07	.514	[-0.28, 0.14]
FRPL eligible	1.28	0.058	[0.99, 1.65]	1.26	0.051	[1.00, 1.59]	-0.25	.234	[-0.66, 0.16]
Percent black student body	0.30	0.016	[0.11, 0.80]	0.30	0.042	[0.09, 0.96]	0.28	.735	[-1.35, 1.92]
Percent FRPL Eligible student body	91.44	<.001	[20.99, 398.38]	117.96	<.001	[21.35, 651.79]	-3.52	.005	[-5.97, -1.08]
Cohort	0.93	0.071	[0.09, 0.12]	0.94	0.098	[0.87, 1.01]	-0.23	<.001	[-0.34, -0.11]
Panel B: race as moderator									
Suspended in 9th grade	5.19	<.001	[3.86, 7.00]	4.98	<.001	[3.79, 6.54]	0.04	.848	[-0.36, 0.44]
Black	2.25	<.001	[1.93, 2.62]	2.10	<.001	[1.82, 2.42]	-0.02	.847	[-0.26, 0.22]
Suspended in 9th grade X Black	0.89	0.546	[0.62, 1.29]	0.88	0.459	[0.64, 1.23]	-0.21	.395	[-0.68, 0.27]
FRPL eligible	1.28	0.062	[0.99, 1.65]	1.25	0.055	[0.99, 1.58]	-0.26	.222	[-0.67, 0.15]
Percent black student body	0.30	0.016	[0.11, 0.80]	0.30	0.042	[0.09, 0.96]	0.27	.742	[-1.36, 1.91]
Percent FRPL eligible student body	91.25	<.001	[20.90, 398.43]	117.36	<.001	[21.16, 650.86]	-3.53	.005	[-5.98, -1.09]
Cohort	0.93	0.072	[0.85, 1.01]	0.94	0.096	[0.87, 1.01]	-0.22	<.001	[-0.34, -0.11]
Panel C: FRPL as moderator									
Suspended in 9th grade	6.69	<.001	[2.80, 15.98]	4.87	<.001	[2.17, 10.91]	-0.71	.241	[-1.91, 0.48]
FRPL eligible	1.31	0.046	[1.01, 1.70]	1.26	0.052	[1.00, 1.60]	-0.31	.154	[-0.74, 0.12]
Suspended in 9th grade X FRPL	0.71	0.453	[0.29, 1.73]	0.94	0.878	[0.41, 2.13]	0.63	.311	[-0.59, 1.84]
Black	2.21	<.001	[1.92, 2.55]	2.06	<.001	[1.81, 2.34]	-0.07	.503	[-0.29, 0.14]
Percent black student body	0.66	0.401	[0.25, 1.75]	0.62	0.411	[0.19, 1.96]	0.21	.800	[-1.41, 1.82]
Percent FRPL eligible student body	71.61	<.001	[16.31, 314.40]	93.66	<.001	[16.85, 520.62]	-3.27	.009	[-5.72, -0.83]
Cohort	0.93	0.071	[0.85, 1.01]	0.94	0.098	[0.87, 1.01]	-0.23	<.001	[-0.34, -0.11]

Ever suspended model uses logit link function, times suspended relies on log link function, and days suspended model relies on identity link function

OR odds ratio, IRR incident rate ratio, FRPL free or reduced-price lunch

Fig. 1 Predicted probability of suspension after 9th grade by 9th grade suspension and race



to number of days per suspension ($b = -0.11, p = 0.514, 95\% \text{ CI } [-0.28, 0.14]$).

Next, we examined the extent to which student race (black versus white) moderated the relationship between being suspended in 9th grade and each outcome (see Table 2 Panel B). Following best practices for interpreting non-linear interaction effects (Mize, 2019), we calculated and plotted the predicted probabilities and predicted counts for each model at each level of both the predictor (i.e., 9th grade suspension) and moderator (i.e., race) and calculated the statistical significance of the second differences. The first outcome we examined was whether students were ever suspended after 9th grade. Figure 1 displays the predicted marginal probability of being suspended after 9th grade at each level of 9th grade suspension and race. The model predicted that white non-suspended students would have a 0.10 probability of suspension, white suspended students would have a 0.40 probability of suspension, black non-suspended students would have a 0.20 probability of suspension, and black suspended students would have a 0.57 probability of suspension. The influence of being suspended in 9th grade on the predicted probability of being suspended after 9th grade was not significantly different for black students from white students (2nd difference = 0.07, $p = 0.088$).

The next outcome we examined was the number of suspensions following 9th grade. Figure 2 displays the predicted marginal count of times suspended after 9th grade at each level of 9th grade suspension and race. The model predicted that white non-suspended students would have 0.15 suspensions, white suspended students would have 0.93 suspensions, black non-suspended students would have 0.32 suspensions, and black suspended students would have 1.73

suspensions. The influence of being suspended in 9th grade on the predicted count of times suspended after 9th grade was significantly higher for black students than white students (2nd difference = 0.63, $p = 0.001$). Together, the findings from Figs. 1 and 2 suggest that being suspended in 9th grade has no racially disparate relationship with the likelihood of *ever* being suspended in later grades, but it is linked with a larger *number* of suspensions for black students relative to white students.

The final outcome we examined was the number of days per suspension following 9th grade. Figure 3 displays the predicted marginal mean number of days per suspension after 9th grade at each level of 9th grade suspension and race. The model predicted that white non-suspended students would be suspended 3.36 days per suspension, white suspended students would be suspended 3.32 days per suspension, black non-suspended students would be suspended 3.34 days per suspension, and black suspended students would be suspended 3.09 days per suspension. The influence of being suspended in 9th grade on the predicted number of days per suspension after 9th grade was not significantly higher for black students than white students (2nd difference = -0.21, $p = 0.395$).

Next, we examined the extent to which FRPL moderated the relationship between being suspended in 9th grade and each outcome (see Table 2 Panel C). Figure 4 displays the predicted probability of being suspended after 9th grade at each level of 9th grade suspension and FRPL eligibility. The model predicted that non-suspended students with paid lunch would have a 0.12 probability of suspension, suspended students with paid would have a 0.53 probability of suspension, non-suspended students with FRPL would have

Fig. 2 Predicted number of suspensions after 9th grade by 9th grade suspension and race

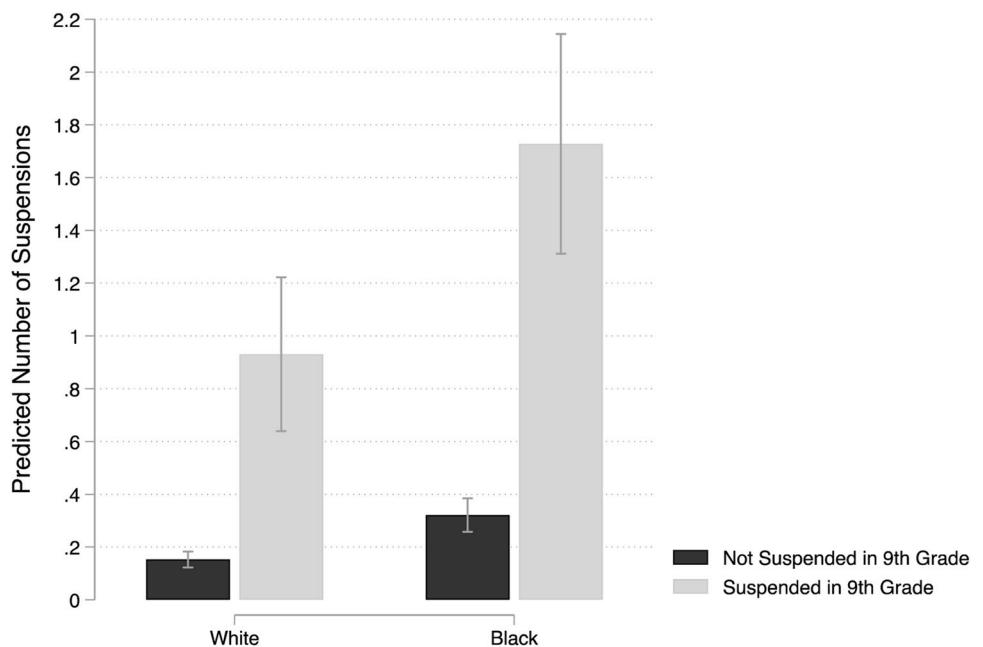
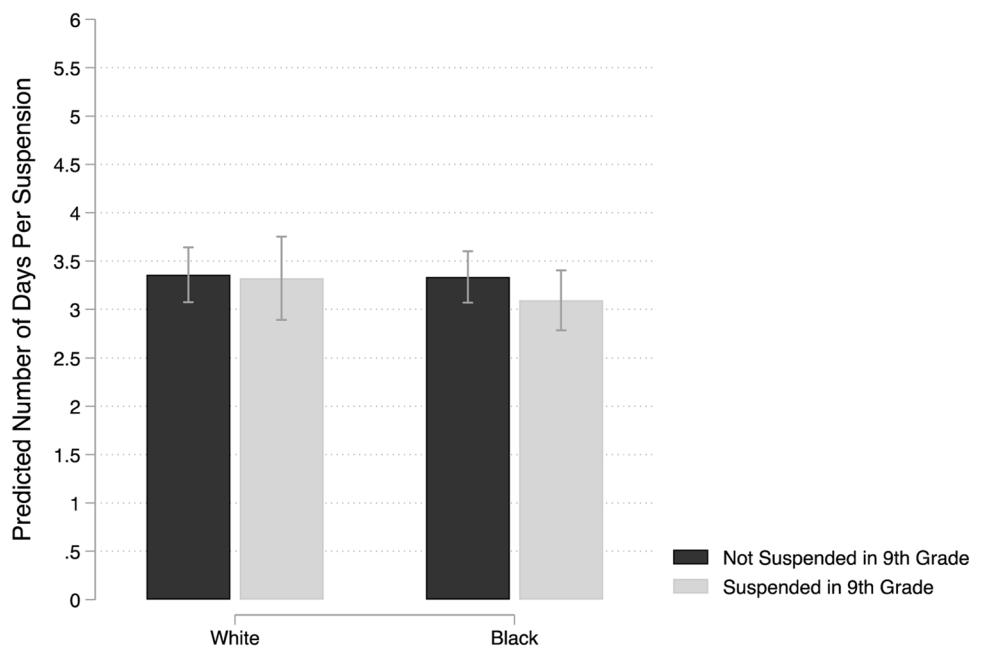


Fig. 3 Predicted number of days per suspension after 9th grade by 9th grade suspension and race



a 0.15 probability of suspension, and suspended students with FRPL would have a 0.52 probability of suspension. The influence of being suspended in 9th grade on the predicted probability of being suspended after 9th grade was not significantly different for FRPL eligible and non-eligible students (2nd difference = -0.05, $p = 0.657$).

Figure 5 displays the predicted count of times suspended after 9th grade at each level of 9th grade suspension and FRPL eligibility. The model predicted that non-suspended students with paid lunch would have 0.18 suspensions, suspended students with paid lunch would have 1.25

suspensions, non-suspended students with FRPL would have 0.23 suspensions, and suspended students with FRPL would have 1.48 suspensions. The influence of being suspended in 9th grade on the predicted count of times suspended after 9th grade was not significantly different for FRPL eligible and non-eligible students (2nd difference = 0.18, $p = 0.720$).

Figure 6 displays the predicted number of days per suspension after 9th grade at each level of 9th grade suspension and FRPL eligibility. The model predicted that non-suspended students with paid lunch would be suspended 3.63 days per suspension, suspended students with paid

Fig. 4 Predicted probability of suspension after 9th grade by 9th grade suspension and FRPL eligibility

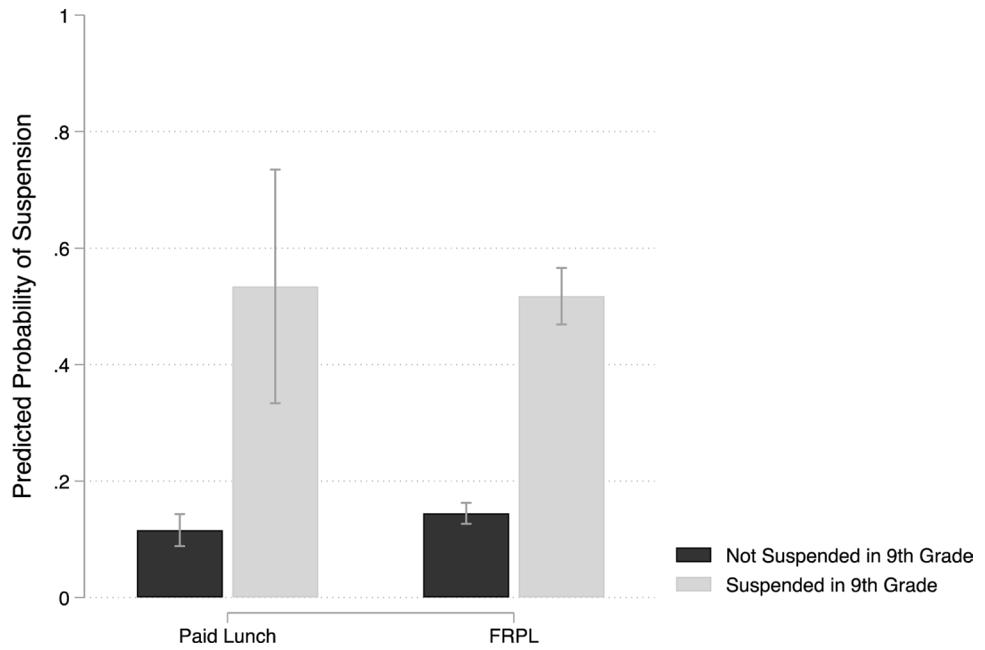
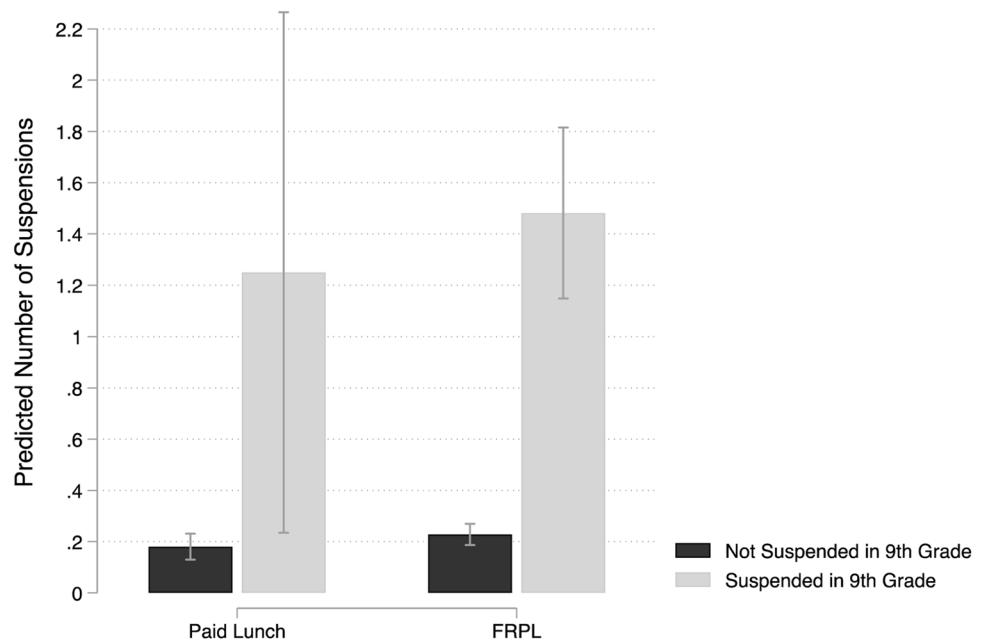


Fig. 5 Predicted number of suspensions after 9th grade by 9th grade suspension and FRPL eligibility

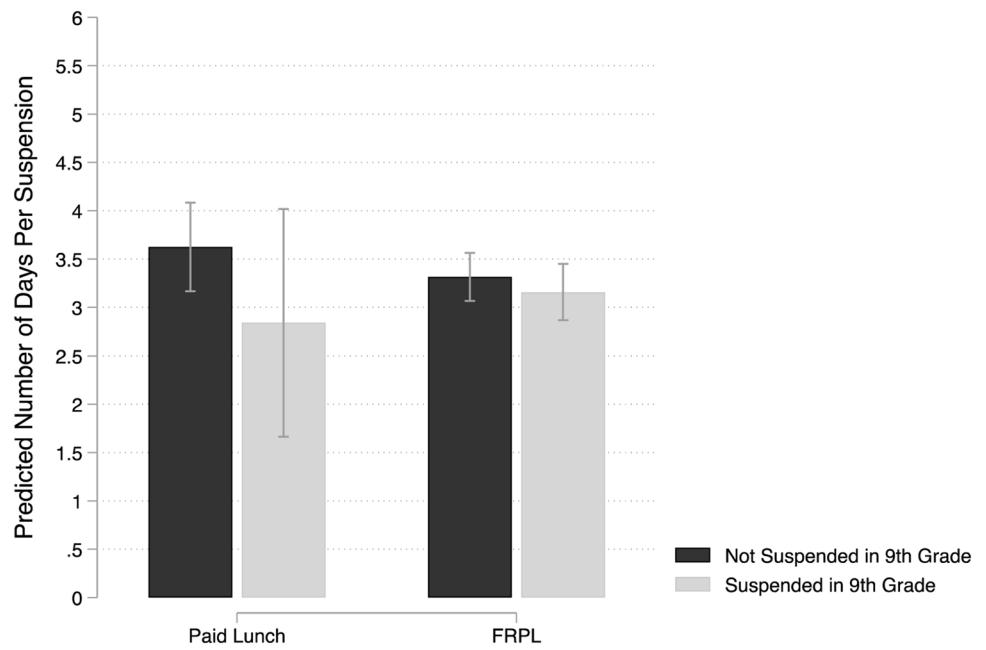


lunch would be suspended 2.84 days per suspension, non-suspended students with FRPL would be suspended 3.32 days per suspension, and suspended students with FRPL would be suspended 3.16 days per suspension. The influence of being suspended in 9th grade on the predicted number of days per suspension after 9th grade was not significantly higher for FRPL eligible than non-eligible students (2nd difference = 0.63, $p = 0.311$).

As noted, one sensitivity analysis included re-estimating the models using school dummy variables and cluster robust

standard errors rather than multilevel models. The overall pattern of results was similar using this modeling strategy. It is noteworthy, though, that the results of each of the six models (three outcomes, each with two moderators) indicated that there was a substantial range of coefficients on the school dummy variables, as well as statistically significant differences among schools. For instance, in the model predicting the number of suspensions after grade 9 using student race as a moderator, the coefficients on the dummy variables ranged from -3.49 to 3.40. Moreover, all but three

Fig. 6 Predicted number of days per suspension after 9th grade by 9th grade suspension and FRPL eligibility



of the schools had coefficients that were significantly different from the reference school, including values that were both higher and lower than the reference school. This indicates that school-level differences explain, at least in part, the pattern of results identified here.

As an additional exploratory analysis, we examined the three-way interaction between being suspended in 9th grade, race, and FRPL eligibility. This three-way interaction was not statistically significant for any of the three dependent variables.

Discussion

The burgeoning body of school discipline literature indicates that punishments lead to a range of unintended consequences, including decreased academic achievement, school dropout, delinquent behavior, and justice system involvement (Arcia, 2006; Hemez et al., 2020; Monahan et al., 2014; Mowen & Brent, 2016; Skiba, Arredondo, et al., 2014; Skiba, Chung, et al., 2014; Wolf & Kupchik, 2017). Another potential consequence of school punishment is that it subjects youth to increased surveillance and future punitive disciplinary responses. The extent to which such a secondary sanctioning process occurs within schools has only recently come under question (see Wiley et al., 2020), but the findings from this study bolster the limited body of research which suggests that students are at greater risk of future disciplinary action once they have been suspended.¹

Specifically, we find that youth who are suspended in 9th grade have increased odds of receiving a suspension in grades 10 through 12. While these students experience more subsequent suspensions, these suspensions are not significantly longer than the suspensions of students who were not suspended in 9th grade.

Research consistently shows that black youth disproportionately experience exclusionary punishment, and our study furthers this line of research by examining whether black youth were more susceptible to subsequent punishments than their white peers. Here, we find a significant difference between black and white youth for one of our outcomes: the number of subsequent suspensions. The fact that interaction effects were not significant for odds of later suspension or average number of days suspended may suggest that a greater number of subsequent suspensions among black students is a function of increased scrutiny rather than behavioral differences. That is, that there was not a similar increase in the number of days per suspension suggests that the offenses for which students were suspended were not severe ones that merited being removed from school for many days, but minor ones that led to relatively short removals. This finding is consistent with ethnographic research that found that subsequent to black boys' initial punishments, teachers increasingly monitored them, often assuming that their misbehavior was intentional, and stereotyped them as "criminal" and even "unsalvageable" in subsequent decisions to punish (Ferguson, 2000). This finding is also consistent with experimental evidence indicating that students' records of prior infractions had different impacts on educators' perceptions of black and white students' behavior. For example, a prior study found that having a record of

¹ But for examination of secondary sanctioning within the criminal justice system, see Liberman et al. (2014), McGlynn-Wright et al. (2020).

two behavioral infractions rather than one led teachers to recommend discipline for black students, but not for white students (Okonofua & Eberhardt, 2015). Although the current study does not measure behavioral infractions, it comes to a similar conclusion that prior engagement with discipline has implications for students' future experiences with discipline. Both the current study and findings by Okonofua and Eberhardt (2015) are consistent with an interpretation that educators' decisions around school discipline are based in part on students' histories, and that these histories matter differently for black and white students.

Our underlying theoretical framework, that school personnel decisions to punish are rooted in stereotype congruence, guided our expectations around how race would moderate the relationship between initial and subsequent punishment. In particular, given the widespread racist stereotypes about black criminality (Muhammad, 2010), we presumed that being suspended in 9th grade would be perceived as stereotype-congruent for black students, but not for white students. These differences, then, would influence schools' decisions to suspend students in the future. Although our findings offer some support for this idea, it is necessary to note that we are limited in the conclusions we can draw around the causes of subsequent discipline. Consistent with stereotype congruence theory, a secondary sanctioning process assumes that subsequent punishment is largely the result of heightened scrutiny of otherwise tolerated behavior. Our data do not include measures of the mechanisms that may account for subsequent suspensions, therefore it is unknown whether additional punishment was based on actual misbehavior. However, prior research has consistently shown that racial disparities in suspension are not fully explained by differences in behavior (Huang, 2020; Skiba, Arredondo, et al., 2014; Skiba, Chung, et al., 2014). Thus, while we cannot empirically rule out the possibility that the racial discipline gaps identified in this study—and the racialized patterns of discipline following an initial suspension in 9th grade—are driven by racial differences in student behavior, prior research indicates that this explanation is likely insufficient. Additionally, the sensitivity tests that incorporated school dummy variables indicates that there are meaningful between-school differences in the pattern of outcomes, suggesting that individual behavior alone does not explain this study's findings.

This study also examined the role that income—as measured by students' eligibility for free or reduced price lunch—played in schools' administration of suspensions following 9th grade suspension. Class is a distinct construct from race, although the two are highly correlated in American society. We examined the extent to which classism may put students from low-income families at increased risk for additional punishment in a similar way that racism may put black students at increased risk for additional punishment.

We found no such result; students eligible for FRPL were statistically indistinguishable from their peers in the relationship between being suspended in 9th grade and all three suspension outcomes. This suggests that racism more than classism is a driving force behind the year-to-year patterns in how schools suspend students.

This study's findings have implications for theory and research related to school discipline and the attendant racial disparities. This study suggests that stereotype congruence theory explains some of the longitudinal patterns in how schools punish white and black students differently. Future research would benefit from better understanding what leads schools to the decision to punish students, with a particular focus on understanding the stereotypes of school personnel. For research, this study addresses a gap in the literature about how being suspended in one grade predicts being suspended in later grades within the same school. Given that prior research has found that being suspended in multiple grades has a compounding effect on students' risk of arrest (Mowen & Brent, 2016), this may be a key mechanism for reducing the negative sequelae of school suspension which are disproportionately borne by black students.

Implications for Policy and Practice

Our study also has important implications for schools' practices, and points to a malleable school policy that has negative consequences for all students, but particularly for black students. If being suspended in 9th grade starts students on a trajectory toward being suspended in future grades, schools might benefit from prioritizing other ways of addressing student behaviors. For example, given that schools administer the vast majority of suspensions for non-criminal, non-violent offenses (Skiba, 2000), educators might take a more positive view of student behaviors and acknowledge that their judgment of what behaviors they consider appropriate might be racist. When student behaviors do cause harm, focusing on healing and restoring all parties involved may be a more productive approach than removing resources from students (Gonzalez, 2012).

However, simply trading suspensions for other disciplinary practices does not address the underlying problem of racial stereotypes. Offering professional development and training for teachers to counter their racial stereotypes has been suggested as one potential strategy to reduce the harmful impact of racial stereotyping, although as Kupchik (2020) notes, unless the underlying school structure is altered to support students who do not fit white, middle-to-upper class standards then training is unlikely to have a lasting impact on racial disparities. As such, a creative reimagining of how to address student behaviors without relying on exclusionary forms of punishment is likely necessary.

Additionally, schools and students are likely to benefit from providing additional resources to students who do experience a suspension. That is, rather than removing resources from students by excluding them from school and expecting them to subsequently conform to the school's behavioral expectations, providing students with extra resources to reintegrate them into the school environment would be useful. For example, intentionally building trusting relationships between suspended students and school personnel is likely to have lasting benefits for students (Crosnoe et al., 2004), especially if these relationships are culturally responsive and appropriate (Stevenson, 2008). Similarly, fostering a school climate that prioritizes a sense of belonging among black students in particular may be useful for interrupting racialized patterns of suspension (Fisher et al., 2020).

Limitations

This study has several limitations that should be considered. First, no information was available about what led to the school's decision to suspend students. We know nothing about school policies, the stereotypes held by school personnel, or the behavior of students. These factors, and others, are likely to matter in schools' decision to suspend students (Huang, 2020; Skiba, Arredondo, et al., 2014; Skiba, Chung, et al., 2014). Future research would benefit from measuring these constructs and including them in a similar analysis. Along similar lines, future research should replicate these findings in a variety of settings to examine the extent to which these findings generalize outside of the school district from which the data came. Second, schools punish students in ways besides suspension, including expulsion, transfers to other schools, detention, and less formal punishments that may not appear in administrative records; none of that information was available here. Finally, our measures of student race and income are used as proxies for the extent to which students are likely to experience the consequences of racism and classism. These measures are imprecise and fail to capture the complexity of how systems of oppression bear upon individuals. Additionally, the measure of income in particular is dependent on self-reporting; if families do not self-identify as low income, they may not be categorized by the school district as FRPL eligible, even if they are in fact eligible. Similarly, we are unable to disentangle the differences between families eligible for free lunch or reduced-price lunch. Moreover, these measures do not account for other parallel systems of oppression. Future qualitative work in the vein of Annamma's (2017) research may be particularly beneficial in this regard.

Conclusion

Although research has long acknowledged the negative consequences of being suspended and the prevalence of racial disparities in suspensions, little research to date has examined longitudinal patterns of suspension. Using four-wave longitudinal data from three cohorts of students in a large urban school district in the USA, this study found that being suspended in 9th grade predicted continued suspension in subsequent years in the same school, including the overall likelihood of being suspended and the number of suspensions, but not the number of days per suspension. Moreover, this relationship was particularly strong for black (relative to white) students in terms of the number of future suspensions, suggesting that the cascading consequences of initial school punishment are more disruptive for black youth. No differences existed by students' household income. These findings can inform ongoing conversations about minimizing the negative effects of school punishment and seeking racial equity in how schools treat students.

Authors' Contributions BWF conceived of the study, conducted the data analysis, and contributed to the writing of the manuscript. SAW assisted with data analysis and contributed to the writing of the manuscript. AMW assisted with data analysis and contributed to the writing of the manuscript.

Funding This project was supported by National Institute of Justice, Office of Justice Programs, U.S. Department of Justice under Award No. 2016-CK-BX-0017. The opinions, findings, and conclusions or recommendations expressed in this presentation are those of the authors and do not necessarily reflect those of the Department of Justice.

Data Availability The data used for this study have been provided by a school district under the agreement that they will remain confidential and not shared with anyone aside from the approved researchers.

Code Availability Stata syntax used in this study is available upon request.

Declarations

Conflict of interest The authors declare no conflicts of interest.

Ethical Approval This study was approved by the University of Louisville IRB under application number 16.0419.

References

- Annamma, S. A. (2017). *The pedagogy of pathologization: Dis/abled girls of color in the school-prison nexus*. Routledge.
- Arcia, E. (2006). Achievement and enrollment status of suspended students: Outcomes in a large, multicultural school district. *Education and Urban Society*, 38(3), 359–369.
- Crosnoe, R., Johnson, M. K., & Elder, G. H., Jr. (2004). Intergenerational bonding in school: The behavioral and contextual correlates

of student-teacher relationships. *Sociology of Education*, 77(1), 60–81.

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17(5), 383–386.

Edwards, L. (2016). Homogeneity and inequality: School discipline inequality and the role of racial composition. *Social Forces*, 95(1), 55–76.

Ferguson, A. A. (2000). *Bad boys: Public schools in the making of black masculinity*. University of Michigan Press.

Fisher, B. W., Dawson-Edwards, C., Higgins, E. M., & Swartz, K. (2020). Who belongs in school? Examining the link between racial disparities in sense of school belonging and suspension. *Journal of Community Psychology*, 48, 1481–1499.

Gansen, H. M. (2020). Disciplining difference(s): Reproducing inequalities through disciplinary interactions in preschool. *Social Problems*. <https://doi.org/10.1093/socpro/spaa011>

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions. *Yale University Child Study Center*, 9(28), Retrieved from https://medicine.yale.edu/childstudy/zigler/publications/Preschool%20Implicit%20Bias%20Policy%20Brief_final_9_26_276766_5379_v1.pdf (accessed 20 May 2021).

Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology*, 94(2), 292.

González, T. (2012). Keeping kids in schools: Restorative justice, punitive discipline, and the school to prison pipeline. *JL & Educ.*, 41, 281.

Hemez, P., Brent, J. J., & Mowen, T. J. (2020). Exploring the school-to-prison pipeline: How school suspensions influence incarceration during young adulthood. *Youth Violence and Juvenile Justice*, 18(3), 235–255.

Hilton, J. L., & Von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47(1), 237–271.

Hirschfield, P. J. (2008). Preparing for prison? The criminalization of school discipline in the USA. *Theoretical Criminology*, 12(1), 79–101.

Holland, P. W. (2008). Causation and race. In T. Zuberi & E. Bonilla-Silva (Eds.), *White logic, white methods: Racism and methodology*. Lanham, MD: Rowman & Littlefield Publishers Inc.

Huang, F. L. (2020). Prior problem behaviors do not account for the racial suspension gap. *Educational Researcher*, 49(7), 493–502.

Jacobsen, W. C., Pace, G. T., & Ramirez, N. G. (2019). Punishment and inequality at an early age: Exclusionary discipline in elementary school. *Social Forces*, 97(3), 973–998.

Kozol, J. (1991). *Children in America's Schools*. New York: Crown.

Kunesh, C. E., & Noltmeyer, A. (2019). Understanding disciplinary disproportionality: Stereotypes shape pre-service teachers' beliefs about black boys' behavior. *Urban Education*, 54(4), 471–498.

Kupchik, A. (2020). Beyond the low-hanging fruit: Reducing racial inequality by rethinking school safety efforts. *Sociological Forum*, 35(3), 813–821.

Kupchik, A. (2010). *Homeroom security: School discipline in an age of fear*. NYU Press.

Lewis, A. E., & Diamond, J. B. (2015). *Despite the best intentions: How racial inequality thrives in good schools*. Oxford University Press.

Liberman, A. M., Kirk, D. S., & Kim, K. (2014). Labeling effects of first juvenile arrests: Secondary deviance and secondary sanctioning. *Criminology*, 52(3), 345–370.

McGlynn-Wright, A., Crutchfield, R. D., Skinner, M. L., & Haggerty, K. P. (2020). The usual, racialized, suspects: The consequence of police contacts with black and white youth on adult arrest. *Social Problems*. <https://doi.org/10.1093/socpro/spaa042>

Mize, T. D. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, 6, 81–117.

Monahan, K. C., VanDerhei, S., Bechtold, J., & Cauffman, E. (2014). From the school yard to the squad car: School discipline, truancy, and arrest. *Journal of Youth and Adolescence*, 43(7), 1110–1122.

Morris, M. (2016). *Pushout The criminalization of black girls in schools*. The New Press.

Mowen, T., & Brent, J. (2016). School discipline as a turning point: The cumulative effect of suspension on arrest. *Journal of Research in Crime and Delinquency*, 53(5), 628–653.

Muhammad, K. G. (2010). *The condemnation of blackness: Race, crime, and the making of modern urban America*. Harvard University Press.

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

Omi, M., & Winant, H. (2014). *Racial formation in the United States*. Routledge.

Payne, A. A., & Welch, K. (2010). Modeling the effects of racial threat on punitive and restorative school discipline practices. *Criminology*, 48(4), 1019–1062.

Pearman, F. A., Curran, F. C., Fisher, B., & Gardella, J. (2019). Are achievement gaps related to discipline gaps? *Evidence from National Data*. *Aera Open*, 5(4), 2332858419875440.

Petras, H., Masyn, K. E., Buckley, J. A., Ialongo, N. S., & Kellam, S. (2011). Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual-and context-level influences. *Journal of Educational Psychology*, 103(1), 223.

Ramey, D. M. (2018). The social construction of child social control via criminalization and medicalization: Why race matters. *Sociological Forum*, 33(1), 139–164.

Rios, V. (2011). *Punished: Policing the lives of Black and Latino boys*. NYU Press.

Rothbart, M., Evans, M., & Fulero, S. (1979). Recall for confirming events: Memory processes and the maintenance of social stereotypes. *Journal of Experimental Social Psychology*, 15(4), 343–355.

Sampson, R. J., & Laub, J. H. (1997). A life-course theory of cumulative disadvantage and the stability of delinquency. *Developmental Theories of Crime and Delinquency*, 7, 133–161.

Shedd, C. (2015). *Unequal city: Race, schools, and perceptions of injustice*. Russell Sage Foundation.

Simon, J. (2007). *Governing through crime: How the war on crime transformed American democracy and created a culture of fear*. Oxford University Press.

Skiba, R. (2000) *Zero Tolerance, Zero Evidence: An Analysis of School Disciplinary Practice*. Bloomington, IN: Education Policy Center Indiana University. Retrieved from <https://files.eric.ed.gov/fulltext/ED469537.pdf> (accessed 20 May 2021).

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

Skiba, R. J., Chung, C. G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal*, 51(4), 640–670.

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., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, 34(4), 317–342.

Stevenson, H. C. (2008). Fluttering around the racial tension of trust: Proximal approaches to suspended black student–teacher relationships. *School Psychology Review*, 37(3), 354–358.

United States Department of Education, Office of Civil Rights. (2019). School Climate and Safety. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/docs/school-climate-and-safety.pdf> (accessed 25 January 2021).

Valentino, N. A., Hutchings, V. L., & White, I. K. (2002). Cues that matter: How political ads prime racial attitudes during campaigns. *American Political Science Review*, 96, 75–90.

Wiley, S. A., Slocum, L. A., O'Neill, J., & Esbensen, F. A. (2020). Beyond the breakfast club: Variability in the effects of suspensions by school context. *Youth & Society*, 52(7), 1259–1284.

Wolf, K. C., & Kupchik, A. (2017). School suspensions and adverse experiences in adulthood. *Justice Quarterly*, 34(3), 407–430.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.