

# The Punishment Gap: School Suspension and Racial Disparities in Achievement

Edward W. Morris<sup>1</sup> and Brea L. Perry<sup>2</sup>

<sup>1</sup>University of Kentucky and <sup>2</sup>Indiana University

## ABSTRACT

While scholars have studied the racial “achievement gap” for several decades, the mechanisms that produce this gap remain unclear. In this article, we propose that school discipline is a crucial, but under-examined, factor in achievement differences by race. Using a large hierarchical and longitudinal data set comprised of student and school records, we examine the impact of student suspension rates on racial differences in reading and math achievement. This analysis—the first of its kind—reveals that school suspensions account for approximately one-fifth of black-white differences in school performance. The findings suggest that exclusionary school punishment hinders academic growth and contributes to racial disparities in achievement. We conclude by discussing the implications for racial inequality in education.

**KEYWORDS:** achievement gap; school discipline; racial disparity; punishment; at-risk students.

Racial disparities in educational achievement are one of the most important sources of American inequality. Racial inequalities in adulthood in areas as diverse as employment, incarceration, and health can be traced to unequal academic outcomes in childhood and adolescence (Belfield and Levin 2007). While the racial “achievement gap” has been consistently documented over several decades, scholars are still working to understand the mechanisms that produce this gap (Jencks and Phillips 1998; Magnuson and Waldfogel 2008). In this article, we propose that school discipline is a crucial, but under-examined, factor in achievement differences by race. Though large racial disparities in discipline exist, this pattern has never been empirically examined as an explanation of racial gaps in school performance. This article presents evidence that exclusionary school punishment may hinder academic growth and contribute to racial inequalities in achievement.

Using detailed data from school records and controlling for a host of school and non-school factors, we confirm that minority students are more likely to be suspended from school. Moreover, using variance decomposition methods that isolate within-student trajectories, we show that suspension is associated with significantly lower achievement growth across time. Finally, we conduct the first comprehensive analysis of suspension as an explanation for the racial gap in achievement. This analysis reveals that school suspensions account for approximately *one-fifth* of black-white differences in school

This research was supported by a grant from the Spencer Foundation. The authors wish to thank Rebecca DiLoretto and the Children’s Law Center for their contributions to this project and for their commitment to equity and justice for all children in public education. Direct correspondence to: Edward W. Morris, 1515 Patterson Office Tower, Department of Sociology, University of Kentucky, 40506. E-mail: edward.morris@uky.edu.

performance, demonstrating that exclusionary discipline may be a key driver of the racial achievement gap. We suggest that the escalation of exclusionary discipline in schools can result in severe academic consequences for minority students.

### THE ACHIEVEMENT GAP

Racial differences in achievement between white and African American children have long been a concern for researchers and policy makers. Today, this issue continues to present a complex and vexing social problem. Data from the National Assessment of Educational Progress (NAEP) reveals that although gaps in reading and mathematics achievement between black students and white students have narrowed in the past 40 years, they remain significantly different (Hedges and Nowell 1999; Jencks and Phillips 1998; Magnuson and Waldfogel 2008). In 2013, for example, African American students, on average, scored 31 points below white students in eighth-grade math and 26 points below in eighth-grade reading (NCES 2014).<sup>1</sup> Historically, black students made steady gains in closing the gap after school desegregation in the 1960s; however, this progress leveled off in 1990. The gap has fluctuated slightly since then, but has ultimately changed little over the past two decades. For twelfth grade students, in fact, the gap in NAEP reading is wider now than it was in 1992 (NCES 2014).

Scholars have offered an array of explanations for these differences in academic performance. Racial gaps in school readiness exist when children enter school, which suggests that inequalities outside of schools play an important role (Downey, von Hippel, and Broh 2004). Studies in this vein focus on family and neighborhood effects ranging from economic inequality (Berends, Lucas, and Penalzoa 2008; Magnuson and Waldfogel 2008) to non-cognitive skills (Grissmer and Eiseman 2008) to parental incarceration (Wildeman 2009). Such effects may be compounded when concentrated in specific schools and neighborhoods (Condrón et al. 2012; see also Coleman et al. 1966). Another proposed outside-of-school factor is student resistance to schooling. The widely debated oppositional culture model asserts that minority students perceive schools as white dominated and this prompts ambivalence toward achievement and disengagement from school.<sup>2</sup>

Other explanations focus within education itself. Dennis Condrón (2009) argued that outside-of-school factors explain learning gaps by socioeconomic status, but not by race per se. Condrón (2009) and others point to de facto school segregation, which decreased through the 1980s before reversing course and increasing beginning in the 1990s (Condrón et al. 2012; Vigdor and Ludwig 2008). Related research asserts that certain characteristics of predominately minority schools depress student achievement, such as per pupil funding (Condrón and Roscigno 2003), teaching experience (Corcoran and Evans 2008), and school-level poverty (Rumberger and Palardy 2005). In addition to across-school differences, research has examined processes within schools, especially ability grouping or tracking (Berends et al. 2008; Tyson 2011). This work suggests that the learning opportunities of minority students are restricted by instructional differentiation, which increases learning gaps over time.

Certainly, these explanations are not mutually exclusive, and racial inequality in achievement arises from a complex interplay of school and non-school factors. However, we argue that this literature has not adequately considered one indispensable piece of the puzzle: *school punishment*. Anne Gregory, Russell Skiba, and Pedro Noguera (2010) have proposed that school discipline could be related to achievement differences, but no empirical work has tested this claim. Yet, school punishment is a logical explanation for achievement differences for several reasons. First, punishment varies widely by race, meaning that it is potentially related to racial variation in achievement. Second, exclusionary

1. The data for fourth-grade and twelfth-grade students show similar patterns. The scale of the NAEP tests ranges from 0 to 500 points. The gap is equivalent to nearly one standard deviation, on average (Condrón et al. 2012; NCES 2014).

2. It is impossible to summarize the extensive debate on oppositional culture in the limited space here. For key works, see Ainsworth-Darnell and Downey (1998), Fordham and Ogbu (1986), and Harris (2011).

forms of school punishment, such as suspension, extract students from the learning environment, which can threaten academic progress. Third, school suspensions increased markedly beginning in the 1990s at the same time that progress on narrowing the achievement gap waned. This indicates that overuse of exclusionary discipline may pose barriers in efforts to reduce racial inequalities in education. The consideration of school punishment adds an important dimension to the argument that school-level processes help reproduce the racial achievement gap.

### RACIALIZED DISCIPLINE AND THE INCREASED USE OF SUSPENSION

Beginning in the 1990s, school discipline approaches became increasingly authoritarian and intrusive. Several scholars have proposed that contemporary regimes of school discipline “criminalize” student misbehavior in ways that mirror the criminal justice system (Hirschfield 2008; Kupchik 2010; Kupchik and Monahan 2006; Wacquant 2001; Welch and Payne 2010). School resource officers (uniformed police officers stationed in schools), security cameras, random searches, and “zero tolerance” policies requiring automatic suspension or expulsion for specified offenses all exemplify this strict, encompassing approach. This shift in disciplinary mentality has resulted in a sharp increase in school suspensions. Suspension rates in U.S. public schools have doubled since the 1970s, and in 2010, almost three million students were suspended from school (Losen and Gillespie 2012).

Zero tolerance policies in particular have markedly impacted school suspensions. Disciplinary reformers modeled these policies after “tough on crime” approaches to policing and sentencing that grew in popularity in the late 1980s (Garland 2001; Simon 2007). According to the logic underpinning these approaches, loose social control will allow deviance to flourish. Thus, even small transgressions left unpunished can evolve into larger transgressions and eventually create a deviant normative context. This logic dictated that early, “tough” punishments were critical to maintaining social order. Hence, criminal sentencing guidelines such as “three strikes” laws emerged in the late 1980s. Zero tolerance policies in schools, which mandated automatic suspension or expulsion for serious or repeated offenses, soon followed suit. Despite evidence that zero tolerance does not actually enhance school climate or safety (see American Psychological Association 2008; Skiba and Peterson 1999), schools across the country continue to be enamored with strict disciplinary policies (Hoffman 2014). Under such policies, exclusionary school punishments such as suspension and expulsion have become widespread, replacing milder repercussions such as detention or loss of privileges.

Although these new punitive policies intend to mete out discipline fairly, they disproportionately impact minority students, especially African Americans. Since the publication of the Children’s Defense Fund’s, *School Suspensions: Are They Helping Children* (1975), research has consistently revealed that African American students are punished at higher rates, including classroom reprimands (Ferguson 2000; Morris 2005), office referrals (Rocque 2010; Skiba et al. 2002), suspensions (Losen 2011; McCarthy and Hoge 1987; Wallace et al. 2008), and expulsions (Kewal Ramani et al. 2007; Wallace et al. 2008). Black students are also more likely to experience severe punishment, such as court action or notification of the police (Welch and Payne 2010). Research suggests that African American students are approximately three times as likely as white students to be suspended (Gregory et al. 2010; Wallace et al. 2008). A recent report found that nationwide, one out of six black students has been suspended at least once (Losen and Gillespie 2012).<sup>3</sup> In addition, predominately minority schools are most likely to rely on punitive forms of discipline such as out-of-school suspension or expulsion (Irwin, Davidson, and Hall-Sanchez 2013; Rocque and Paternoster 2011; Welch and Payne 2010). While the discipline of minority students has long occurred at higher rates

3. For Latinos, the picture is more complex. Some research finds that the punishment of Latino students tends to be less extreme, but still occurs at higher rates than whites (Losen and Gillespie 2012; Peguero and Shekarkhar 2011). Other research (including our own analyses) finds that Latinos are not punished at higher rates after controlling for background factors such as free and reduced lunch eligibility. We focus on black-white gaps in this article, but future research should explore school discipline and Latinos in more depth.

compared to white students, this gap has widened as the prevalence of suspension has increased overall (Verdugo 2002). Using a natural experiment, Stephen Hoffman (2014) found that strict, punitive discipline policies increase the racial gap in suspension and expulsion.

These alarming racial disparities in school discipline have prompted a response from the federal government. In January 2014, the U.S. Department of Education issued a set of guiding principles concerning discipline in public schools. Although the federal government cannot dictate local disciplinary policies, this document encourages schools to rely less on exclusionary forms of discipline and reminds schools that they cannot discriminate in administering discipline (U.S. Department of Education 2014). The missive cautions that punishments such as suspension, which remove students from the learning environment, have been linked to ongoing educational problems. However, despite the growing realization of negative consequences, there is surprisingly little research able to specify the direct impact of suspension on outcomes such as academic achievement.

### IMPACTS OF EXCLUSIONARY PUNISHMENT

Suspension appears to have few behavioral or academic benefits for suspended students. Virginia Contenbader and Samia Markson (1998) found that suspension does little to improve subsequent student behavior, and may even exacerbate students' anger or apathy. Exclusionary discipline can weaken school bonds, which may actually increase the likelihood of further deviant behavior (Hirschi 1969; Jenkins 1997). Academically, school suspension has been correlated with low academic performance (Davis and Jordan 1994) and higher risk of dropout (Ekstrom et al. 1986). A quasi-experimental study by Emily Arcia (2006) followed two groups of similar students over time, the only major difference between the groups was that one had been suspended and the other had not. After two years, the suspended group was nearly five grade levels behind the non-suspended group, which suggests that suspension greatly impedes academic progress. More recently, Brea Perry and Morris (2014) found that high rates of suspension at the school level tend to depress student achievement, even for students who were not personally suspended.

However, research that traces the effects of suspension on achievement longitudinally for a large and diverse group of students remains thin. While prior educational research has connected exclusionary discipline to lower achievement, it is still unclear whether or to what extent suspension reduces achievement. In addition, no empirical research to our knowledge has been able to link suspension disparities by race to achievement disparities by race. In this article, we use detailed longitudinal data from school district records and conservative, unbiased fixed-effects modeling to more accurately specify the impact of suspension on achievement over time. Moreover, we extend this analysis directly to the racial achievement gap to determine the extent to which school discipline disparities explain this gap.

Our unique data and analysis provide the first comprehensive study of the impact of suspension on racial differences in achievement. Using advanced multilevel methods that capitalize on the rich explanatory power of longitudinal and hierarchical data, we focus on the following questions: (1) Are racial and ethnic minorities at disproportionate risk for school suspension? (2) Are racial-ethnic background and school suspension associated with academic achievement in reading and math, controlling for other individual characteristics and all school-level heterogeneity? (3) Do racial differences in the likelihood of suspension explain a significant proportion of the racial achievement gap?

### METHODS

This research uses data from the Kentucky School Discipline Study (KSDS) (Perry and Morris 2014). Data are comprised of existing, deidentified school records and supplementary data collected routinely from parents in a large, urban public school district. All data on school discipline and test scores come directly from school records, eliminating any selection bias and social desirability effects that occur when students or parents report on their own behavior. For each student offense resulting

in any disciplinary action (office referral, detention, suspension, expulsion, etc.), school personnel are required to complete an electronic form containing information about the offense, all students involved, and any response by school officials. This information is stored for the purposes of monitoring school safety and reporting discipline statistics to the state, and is well regulated. Only information on family structure (i.e., single parent family, number of people living in the home) is drawn from the parent survey.

Our sample includes students in grades 6 through 10 (middle and high school) who were enrolled in a district public school over a three-year period beginning in August 2008 and ending in June 2011. The full sample includes 24,347 students. However, 8,089 students (33 percent of the full sample) are dropped due to missing data on end-of-year (spring) Measure of Academic Progress (MAP) test scores. MAP testing by the school district was inconsistent prior to 2009 during the pilot phase. By the 2009-2010 school year, full implementation of the testing was in place. Because the piloting process was random, missing data are unlikely to lead to biases. An additional ten cases were dropped due to missing data on other variables.

The analysis sample includes 16,248 students nested in 17 schools, providing a total of 25,221 observations over three years of data. At baseline, about 65 percent of students are in grades 6 to 8 (ages 11 to 13), and 35 percent in grades 9 to 10 (ages 14 to 16). Approximately 49 percent of students in the sample are girls and 51 percent are boys. The majority of these students are white (59 percent) or black (25 percent). However, 10 percent are Latino, 4 percent are Asian, and 3 percent classify themselves as some other race. Also, 48 percent of students qualify for free or reduced-price meals. These data, which are drawn from one school system, are not nationally representative of all public school children. Most notably, a smaller percentage of the U.S. student population is non-Hispanic black (17 percent) compared to our sample, and a greater percentage is Latino (21 percent; NCES 2014). However, black populations tend to be concentrated in the Southeast where this school district is located. Consequently, these data may be reasonably representative of the Southeastern United States.

With respect to patterns of exclusionary discipline, our sample is on par with national trends (Aud, Fox, and KewalRamani 2010). Specifically, rates of out-of-school suspension in the KSDS and nationally representative National Household Education Surveys (NHES 2007) (U.S. Department of Education 2007) are the same (22 percent had ever been suspended). There are also similar patterns of racial disparities in suspension, which is critical for this analysis in particular. In the KSDS, about 42 percent of black students had ever been suspended, compared to 43 percent in the NHES sample (a non-significant difference). Among Latinos, 26 percent in the KSDS district had ever been suspended compared to 22 percent nationally ( $p < .001$ ). Also, Asians in both data sets were less likely to be suspended, though this difference is larger in Kentucky (4 percent and 11 percent, respectively;  $p < .001$ ). Finally, 18 percent of girls and 26 percent of boys in the KSDS had been suspended compared to 15 percent of girls and 28 percent of boys nationally. This indicates that boys in the general population are slightly more likely to have been suspended than students in the Kentucky district. Overall, these patterns are remarkably similar in magnitude and always in the same direction. These results suggest that exclusionary discipline patterns in the data used for this analysis are representative of national trends, supporting the use of cautious inference to students in other districts.

### Measures

Several static characteristics of individual students are examined as independent variables in multivariate models. Gender is coded as a binary variable (1 = female; 0 = male). Race is measured in five categories and coded as binary indicators: white, African American, Latino, Asian, and other. Family structure is measured by a binary variable indicating whether two parents or guardians were listed on each student's parent information form (1 = two parents; 0 = one parent). Because this measure is available only in the final wave of the study, missing values on 14 percent of observations are replaced via a

logistic regression multiple imputation method. Ten imputations are computed, and Stata's *-mi-* commands are used for imputation and estimation of models that include the family structure variable.

Time is coded using academic year beginning with 0 at baseline in 2008-2009 and ending with 2 in 2010-2011. Time-squared and time-cubed are also calculated to assess the non-linearity of the growth or decline in school suspensions and academic achievement over time. All other time-varying measures are divided into their between-person and within-person information to differentiate the degree to which outcomes are due to average differences between students across waves or differences over time in the characteristics of a student compared to him or herself at other waves (Raudenbush and Bryk 2002). Between-person variance is reflected in the average score for the three waves of the study, and is held constant across observations nested within the same individual. Within-person variance is the average score subtracted from the score for the current wave of the study, and measures how different a person is in a given wave from their own average. For binary variables, the between-person measure is equivalent to the proportion of waves in which each student had the characteristic in question. The within-person score is the difference between the binary indicator for a given wave and the between-person proportion. It ranges from  $-.67$  (having the characteristic in every wave except the current wave) to  $.67$  (having the characteristic only in the current wave), with zero indicating no change across waves.

Socioeconomic status is measured using participation in the free or reduced meal program. For this variable, between-person variance is the average of free/reduced lunch status (coded 1 = yes; 0 = no) across three waves of the study. This is also equal to the proportion of waves in which each student participated in the free/reduced lunch program. The within-person measure is the difference between the binary variable for the current wave and the proportion of waves in which the student participated in the free/reduced lunch program. Receipt of special education services is also measured using binary coding, and is decomposed into between- and within-person variation.

Out-of-school student suspension is measured as a dichotomous variable and is the dependent variable in the first set of regressions. Information on student suspensions is drawn from official school records. Though a small minority of students experienced multiple out-of-school suspensions in a given school year, there are insufficient cases to employ a count variable. In subsequent regressions predicting academic achievement, suspension is an independent variable and is split into between- and within-person variation. The between-person measure of suspension is the proportion of waves in which a student is suspended, while the within-person measure is suspension in the current wave minus the proportion of years with suspensions.

Performance on tests in math and reading are used to assess achievement, and are also drawn from official school records. Between 2008 and 2011 in the targeted school district, academic achievement was measured using MAP testing across the state. This is a computerized adaptive test that is designed to help schools monitor academic growth in reading and math and make informed decisions about placement and needed services. Scores are numeric and normally distributed. The tests are not timed, and are administered multiple times per year. To reduce concerns about reverse causation (i.e., low academic performance leading to suspension), scores from the end-of-year MAP testing are used in this analysis, making it unlikely that any suspensions occurred following testing. In cases where data from the end-of-year academic achievement tests are missing, the average scores from MAP testing occurring earlier in the same school year are imputed. Currently, similar racial differences in the NAEP appear for both math and reading components of the test (NCES 2014). Therefore, it is appropriate for us to use both math and reading outcomes here. MAP scores for reading and math are examined separately to provide a strong overall assessment of achievement.

### Analyses

Analyses focus on identifying the association between race and ethnicity, suspension, and academic achievement. Multivariate effects are modeled with multi-level mixed logistic and linear regression

models using Stata 13 (StataCorp 2013). These adjust for the hierarchical data structure and the interdependence among observations resulting from having multiple observations over time for each student and multiple students in schools. The models have a three-level structure where level-one observations (time points) are nested in level-two individual students, which are nested in level-three schools.

Because these models focus on predicting an individual-level outcome using both time-invariant and time-variant characteristics, the models include a random intercept at level two. To control for unmeasured time- and student-invariant characteristics of schools, these models include level-three fixed effects using dichotomous school indicators (estimates not shown in tables). This means that mechanisms of suspension and achievement for students in a particular school are estimated relative to other students in the same school. Variables such as the neighborhood in which the school is located and other potential confounding school-level effects that are time invariant or which can reasonably be expected to change very little over a three-year period are controlled since all comparisons are between students within the same school. This strategy also eliminates the small  $n$  problem at level three (i.e., 17 schools) because school-level variation is controlled in the fixed-effects model rather than being used for prediction.

The basic mixed-effects model with three levels predicting test scores using two independent variables, for example, takes the following form:

$$y_{ijk} = \beta_0 + \beta_1 x_{1ijk} + \beta_2 x_{2ijk} + \zeta_j + \alpha_k + \varepsilon_{ijk}$$

In this model,  $i$  corresponds to time (level one),  $j$  to student (level two), and  $k$  to school (level three). The symbol  $\zeta_j$  represents the random intercept at the student level and  $\alpha_k$  is a fixed parameter representing all differences between schools that are stable over time. The fixed parameter at the school level is accomplished through binary school indicators, as noted above. Finally,  $\varepsilon_{ijk}$  is the level one residual. Together,  $\zeta_j$  and  $\varepsilon_{ijk}$  represent the random parts of the model, while the other components are fixed.

The first set of models examines the effects of race and ethnicity on the log odds of suspension. A baseline model (1a) includes race and ethnicity as well as time, but does not include dichotomous school indicators. This is the only model estimated without school-level fixed effects, and this is to demonstrate that part of the increased susceptibility of minorities to exclusionary discipline is explained by racial and ethnic segregation into different schools. In addition, a supplemental regression of school-level characteristics is computed to confirm that the partial confounding effect of dichotomous school indicators is due to black students attending schools with higher suspension rates, controlling for school size and socioeconomic status composition (results not shown). The second model (2a) predicting student suspension includes race and ethnicity, time, and dichotomous school indicators. The third model (3a) adds potential confounding factors, including sociodemographics and special education placement. The fourth model (4a) adds a family structure variable and is estimated using multiple imputation procedures due to missing data on that variable.

In the second set of analyses, quadratic growth curve models are estimated to determine how reading and math achievement scores change over time in this school system. Baseline models include time and race and ethnicity (1b and 1c). The second set of models (2b and 2c) add between- and within-person measures of suspension to assess the degree to which group differences in exclusionary discipline experiences explain the racial and ethnic academic achievement gap. Mediation of the relationship between race and ethnicity and academic achievement by suspension is formally tested using the `-sgmediation-` command in Stata. The purpose of this analysis, following Michael E. Sobel (1986) and Reuben M. Baron and David A. Kenny (1986), is to test whether a mediator carries the influence of an independent variable (IV) to a dependent variable (DV). The `-sgmediation-` command tests all four relationships required to meet criteria for mediation: (1) the IV significantly affects the mediator,

(2) the IV significantly affects the DV in the absence of the mediator, (3) the mediator has a significant unique effect on the DV, and (4) the effect of the IV on the DV is reduced when the mediator is added to the model. The indirect effect of race on achievement through suspension is tested using a conservative bootstrapped estimation procedure with case resampling (MacKinnon and Dwyer 1993). This method for testing the statistical significance of an indirect effect (i.e., mediation) has been shown to produce less biased estimates than the Baron and Kenny (1986) and Sobel (1986) methods in simulation studies (MacKinnon, Warsi, and Dwyer 1995).

The third set of models (3b and 3c) predicting test scores add student sociodemographic characteristics that may confound the relationship between race, suspension, and academic achievement (e.g., socioeconomic status). In these models, time invariant characteristics (i.e., gender and race and ethnicity) are measured at level two, while time variant characteristics (i.e., suspension, socioeconomic status, and special education status) are measured at level one. All level one variables are separated into between-student effects (e.g., Why are students different from each other, on average?) and within-student effects (e.g., Why are students different from themselves this year compared to other years?). In addition, the family structure variable is added to the final models (4b and 4c), which are estimated using multiple imputation procedures.

To demonstrate the long-term effects of suspension on academic achievement, we use the above models to generate a graph of predicted values for test scores over time. These depict trajectories of academic achievement based on early and repeated suspensions in the academic career. Between-student effects of suspended and never-suspended students are reflected in intercept differences between groups, while within-student effects of suspension are depicted by changes in the angles of the lines over time. This figure is based on a model containing all student- and school-level control variables.

Though between-school (i.e., time invariant) school-level characteristics are controlled by the fixed-effects approach, we conduct supplemental analyses to assess the sensitivity of the models to time-variant school-level variables that might be correlated with test scores and/or suspension. These variables include within-school variation on percent racial/ethnic minority, percent free/reduced lunch, percent special education, expenditures per student, school size, and total number of offenses in a school in a given year. Estimates of these effects are for the most part unreliable because there is little variation over three years in these indicators, with the exception of number of offenses. However, including time-variant school-level indicators has very little impact on the coefficients for race or suspension in models predicting test scores, and did not change the substantive conclusions of this research. Consequently, these models are not included in tables of results.

A number of student- and school-level variables (e.g., race, socioeconomic status, and likelihood of suspension) are correlated, introducing the possibility of multicollinearity. However, variance inflation factors (VIFs) do not exceed 3.08 for any model. This reduces concerns about the degree to which multicollinearity might lead to biased estimates.

## RESULTS

Descriptive statistics in Table 1 suggest that 12 percent of public school students will receive an out-of-school suspension in any given year. Academic achievement scores in reading ( $m = 220.21$ ;  $s = 17.49$ ) and math ( $m = 231.33$ ;  $s = 19.60$ ) vary substantially across the sample, which includes students in grades 6 through 10. However, scores within schools are less variable, ranging from a standard deviation of 10.80 to 23.28 when accounting for time invariant school-level heterogeneity. Also, the interclass correlations for MAP reading and math scores are .71 and .81, respectively, suggesting substantial correlation in academic achievement across time within each student.

### The Racial and Ethnic Gap in Exclusionary Discipline

Table 2 contains the results from a mixed-effects logistic regression of suspension on race and ethnicity. Findings in Model 1a do not include school-level fixed effects, permitting the relationship



**Table 1. Descriptive Sample Characteristics**

<i>Students</i>	<i>Proportion</i>	<i>Mean</i>	<i>SD</i>	<i>Range</i>
Female	.49			
Race and ethnicity				
White	.59			
Black	.25			
Latino	.10			
Asian	.04			
Other	.03			
Free/reduced lunch	.48			
Special education	.09			
Suspended	.12			
Two-parent family	.63			
MAP reading score		220.21	17.49	141.00-280.00
MAP math score		231.33	19.60	143.00-300.00
<i>N</i> = 16,248				

between race and ethnicity and suspension to reflect group differences in the kinds of schools that minority students are likely to attend. These indicate that black students are estimated to be 7.57 times as likely to be suspended as white students ( $p < .001$ ), and Latinos are over twice as likely as whites ( $OR = 2.39$ ;  $p < .001$ ). Students of other races are predicted to be 2.61 times more likely to be suspended than whites ( $p < .001$ ), while Asians are less likely than whites ( $OR = .20$ ;  $p < .001$ ). Findings in Model 2a add school-level fixed effects, controlling for all observed and unobserved time invariant heterogeneity in characteristics of schools. In other words, all estimates reflect differences between students in the same school. These findings indicate that black students are still estimated to be almost six times as likely to be suspended as white students ( $OR = 5.91$ ;  $p < .001$ ), while Latinos are nearly twice as likely ( $OR = 1.87$ ;  $p < .001$ ). Students of other races are 2.47 times more likely to be suspended than whites ( $p < .001$ ), on average, while Asians are estimated to be suspended at lower rates than whites ( $OR = .23$ ;  $p < .001$ ). In all, racial segregation into different schools explains about 12 percent of the effect of being black on the odds of suspension, and supplemental analyses confirm that schools with larger concentrations of black students have significantly higher rates of out-of-school suspension. Each additional percentage of the student body that is black is estimated to increase the annual number of school suspensions by about ten, controlling for school size and socioeconomic composition ( $b = 10.16$ ;  $p < .01$ ).

As shown in Model 3a of Table 2, the addition of sociodemographic covariates reduces the magnitude of the impact of race and ethnicity on suspension, and this result is attributable almost entirely to racial and ethnic differences in socioeconomic status (i.e., free/reduced lunch). Students who qualify for free/reduced lunch in all three waves of the study are predicted to be over six times as likely to be suspended as those who never qualify ( $OR = 6.36$ ;  $p < .001$ ). Students who receive special education services are also estimated to be more likely to be suspended ( $OR = 3.19$ ;  $p < .001$ ), while girls are less likely to be suspended than boys ( $OR = .36$ ;  $p < .001$ ). However, even after controlling for socioeconomic status, special education services, and gender, black students are predicted to have nearly three times the odds of suspension compared to whites ( $OR = 2.80$ ;  $p < .001$ ), and students of other races are 57 percent more likely than white students to be suspended ( $p < .05$ ). In contrast, the elevated risk of suspension associated with being Latino is entirely explained by this group's lower levels of socioeconomic status.

Results in Model 4a of Table 2 include a variable measuring family structure, and are estimated using multiple imputation procedures. Students with two parents are 56 percent less likely to be suspended, on average, than those with only one parent or guardian ( $p < .001$ ). Family structure

**Table 2. Mixed-Effects Logistic Regression of Suspension on Race and Ethnicity over Time**

	<i>Model 1a</i> <sup>a</sup>	<i>Model 2a</i>	<i>Model 3a</i>	<i>Model 4a</i>
Within-student $\Delta$				
Time (years)	1.05 (.96-1.15)	.99 (.90-1.09)	.96 (.88-1.06)	.97 (.88-1.06)
Free/reduced lunch			.90 (.61-1.31)	.90 (.61-1.31)
Special education			.95 (.45-2.02)	.96 (.45-2.05)
Between-student				
Race and ethnicity <sup>b</sup>				
Black	7.57 (6.36-9.01)***	5.91 (4.98-7.00)***	2.80 (2.38-3.30)***	2.46 (2.09-2.89)***
Latino	2.39 (1.89-2.03)***	1.87 (1.48-2.36)***	.84 (.67-1.07)	.90 (.71-1.13)
Asian	.20 (.11-.35)***	.23 (.13-.40)***	.29 (.17-.50)***	.33 (.19-.57)***
Other	2.61 (1.74-3.91)***	2.47 (1.66-3.68)***	1.57 (1.06-2.33)*	1.39 (.94-2.05)
Female			.36 (.31-.42)***	.35 (.30-.41)***
Free/reduced lunch			6.36 (5.30-7.63)***	4.81 (4.01-5.75)***
Special education			3.19 (2.60-3.92)***	2.92 (2.36-3.56)***
Two-parent family				.44 (.38-.52)***
<i>N</i>	16,284	16,284	16,284	16,284
Obs	25,221	25,221	25,221	25,221
$\rho$	.63	.61	.57	.56
Wald $X^2/F$	555.87***	709.11***	978.35***	38.99***

Notes: Odds ratios are presented, confidence intervals in parentheses. Models 2 through 4 control for dichotomous school indicators.

<sup>a</sup>Model 1 does not control for dichotomous school indicators (i.e., there is no school-level fixed effect).

<sup>b</sup>Omitted category is white.

\*  $p < .05$  \*\*  $p < .01$  \*\*\*  $p < .001$  (two-tailed tests)

explains a small amount of the variation in the effect of being black on suspension, but black students are still estimated to be nearly two and a half times as likely to be suspended as white students in this model (OR = 2.46;  $p < .001$ ). The effect of being some other race or ethnicity becomes non-significant in this model, suggesting that differences in suspension rates for this group are entirely explained by socioeconomic status and family structure. Also, the effect of free/reduced lunch qualification on odds of suspension is partially explained by family structure, but continues to have a large significant effect in this full model (OR = 4.81;  $p < .001$ ).

### Effects of Exclusionary Discipline on Academic Achievement

Table 3 displays the effects of race and ethnicity and suspension on academic achievement in reading. There is evidence of significant curvilinear growth in academic achievement over the study period such that the test scores grow more substantially early in the study period, but that growth begins to taper off over time ( $p < .001$ ). This is consistent with expectations for MAP growth, where gains are

**Table 3. Mixed-Effects Linear Regression of Reading Achievement on Student Suspension Over Time**

	<i>Model 1b</i>	<i>Model 2b</i>	<i>Model 3b</i>	<i>Model 4b</i>
<b>Within-student <math>\Delta</math></b>				
Time (years)	-3.40 (.54)***	-3.43 (.54)***	-2.94 (.52)***	-2.94 (.52)***
Time-squared	2.37 (.20)***	2.33 (.20)***	2.14 (.19)***	2.14 (.19)***
Suspended		-1.01 (.31)***	-1.10 (.31)***	-1.10 (.31)***
Free/reduced lunch			-.55 (.44)	-.55 (.44)
Special education			1.73 (1.07)	1.71 (1.07)
<b>Between-student</b>				
Race and ethnicity <sup>a</sup>				
Black	-10.87 (.30)***	-8.72 (.30)***	-4.60 (.29)***	-4.39 (.29)***
Latino	-12.95 (.44)***	-12.36 (.41)***	-7.97 (.41)***	-8.04 (.41)***
Asian	-2.04 (.65)**	-2.82 (.63)***	-4.15 (.56)***	-4.28 (.56)***
Other	-4.79 (.78)***	-4.02 (.76)***	-1.62 (.68)*	-1.42 (.68)*
Suspended		-15.05 (.45)***	-8.61 (.42)***	-8.37 (.42)***
Female			1.49 (.22)***	1.55 (.22)***
Free/reduced lunch			-9.21 (.27)***	-8.78 (.28)***
Special education			-20.19 (.40)***	-20.07 (.40)***
Two-parent family				1.39 (.28)***
Constant	215.42 (.42)***	217.68 (.62)***	221.93 (.61)***	220.78 (.65)***
N	16,284	16,284	16,284	16,284
Obs	25,221	25,221	25,221	25,221
$\rho$	.78	.76	.71	.71
Wald $X^2/F$	4,055.50***	5,371.83***	10,696.51***	352.93***

Notes: Unstandardized coefficients, standard errors in parentheses; models control for dichotomous school indicators.

<sup>a</sup>Omitted category is white.

\*  $p < .05$  \*\*  $p < .01$  \*\*\*  $p < .001$   
(two-tailed tests)

more substantial in earlier grades relative to later ones. As shown in Model 1b, students who are black ( $b = -10.87$ ;  $p < .001$ ), Latino ( $b = -12.95$ ;  $p < .001$ ), Asian ( $b = -2.04$ ;  $p < .01$ ), and some other race ( $b = -4.79$ ;  $p < .001$ ) are all predicted to have significantly lower scores on achievement in reading compared to white students, controlling for school-level fixed effects.

Model 2b adds between- and within-person variation in suspension over time, and demonstrates that out-of-school suspension is significantly related to academic achievement. The proportion of waves in which a student is suspended (i.e., propensity to be suspended) is associated with decreases in reading such that those who have been suspended each year of the study are predicted to have a MAP reading score that is over 15 points lower than those who have never been suspended ( $b = -15.05$ ;  $p < .001$ ). This is nearly a one-standard deviation decrease in academic achievement. In other words, being suspended is a strong predictor of a student's academic performance relative to other students in the same school. Also, having a suspension in a given wave is associated with significantly lower performance on reading evaluations ( $b = -1.01$ ;  $p < .001$ ) at the end of that academic year relative to other years, comparing each student to him or herself.

As seen in Models 3b and 4b of Table 3, girls tend to score higher than boys in reading achievement ( $b = 1.49$ ;  $p < .001$ ), on average. Between-person variation in proportion of waves spent in free/reduced lunch status is associated with significant differences in reading achievement ( $b = -9.21$ ;  $p < .001$ ), as is between-student variation in special education placement ( $b = -20.19$ ;  $p < .001$ ). However, within-person changes in these statuses over time do not significantly affect reading or math achievement. Model 4b includes a measure of family structure, suggesting that students in two-parent families perform better in reading than those with one parent ( $b = 1.39$ ,  $p < .001$ ). Most importantly, the addition of these potential confounding factors only partially explains differential academic achievement by race and ethnicity and by suspension.

Findings in Table 4 reflect the effects of race and ethnicity and suspension on math achievement. Again, there is evidence of curvilinear growth in math scores over the study period ( $p < .001$ ), as anticipated. As shown in Model 1c, students who are black ( $b = -13.34$ ;  $p < .001$ ), Latino ( $b = -12.57$ ;  $p < .001$ ), and some other race ( $b = -6.97$ ;  $p < .001$ ) are all predicted to have significantly lower scores on achievement in math compared to white students, controlling for school-level fixed effects. In contrast, Asian students are estimated to perform better in math than whites, on average ( $b = 9.40$ ;  $p < .001$ ).

The effects of suspension on math achievement are included in Model 2c of Table 4. The proportion of waves in which a student is suspended is associated with decreases in math performance such that those who have been suspended each year of the study are predicted to have a MAP math score that is 16.21 points lower than those who have never been suspended ( $p < .001$ ; nearly a one standard deviation reduction). Also, having a suspension in a given wave is associated with significantly lower math performance ( $b = -.56$ ;  $p < .05$ ) at the end of that academic year relative to other years, comparing each student to him or herself.

Effects of control variables on math achievement mirror those for reading achievement. Gender differences are the exception (see Models 3c and 4c of Table 4), as girls score lower than boys, on average, in math achievement ( $b = -2.09$ ;  $p < .001$ ). Also, between-person variation in free/reduced lunch ( $b = -11.14$ ;  $p < .001$ ) and special education status ( $b = -24.21$ ;  $p < .001$ ) are associated with lower math performance. Model 4c shows that students with two parents are estimated to score higher in math achievement than those with one parent ( $b = 1.68$ ;  $p < .001$ ). As with reading achievement, the addition of these potential confounding factors only partially explains differential math achievement by race and ethnicity and suspension.

Figure 1 depicts results from Model 3c in Table 4. Differences in math achievement between suspended and never-suspended students (i.e., between-student effects) are reflected in baseline predicted values of math MAP performance (year 0). Within-student effects of suspension are depicted by changes in predicted values over time. A student who is never suspended has a linear growth in math performance that is reflected in a six-point increase across the three measures, as would be expected for students making normal academic progress. Suspended students have lower baseline scores than never-suspended students, on average, possibly reflecting other unmeasured mechanisms of student success that are correlated with suspension. However, suspension does have meaningful and lasting adverse effects over time independent of early disparities between ever- and never-suspended students. Though students experiencing one early suspension begin with only a three-

**Table 4. Mixed-Effects Linear Regression of Math Achievement on Student Suspension Over Time**

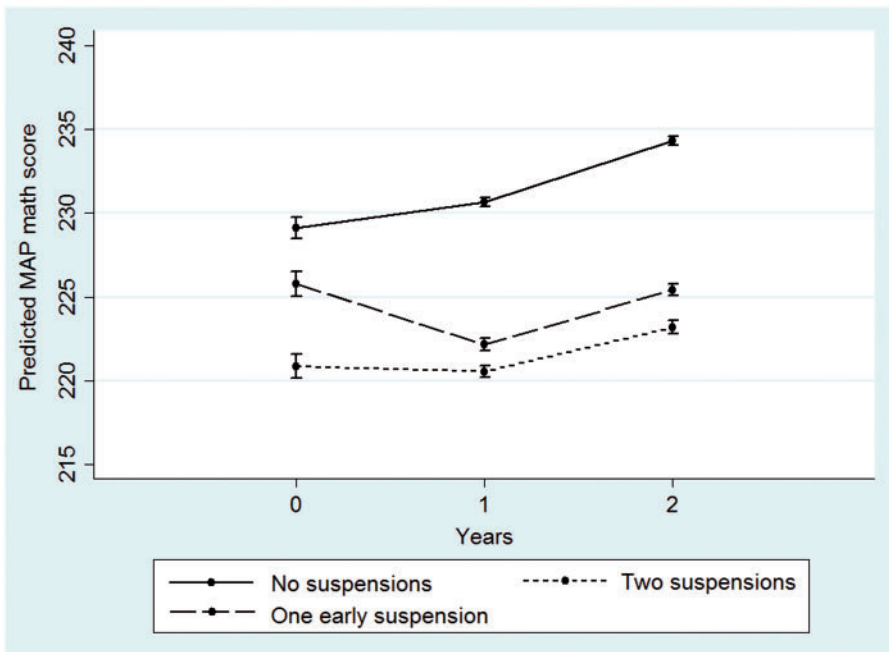
	<i>Model 1c</i>	<i>Model 2c</i>	<i>Model 3c</i>	<i>Model 4c</i>
<b>Within-student <math>\Delta</math></b>				
Time	-1.91	-1.98	-1.61	-1.61
(years)	(.50)***	(.50)***	(.49)***	(.49)***
Time-squared	1.99	1.97	1.82	1.81
	(.18)***	(.18)***	(.18)***	(.18)***
Suspended		-.56	-.60	-.60
		(.28)*	(.27)*	(.27)*
Free/reduced lunch			-.16	-.16
			(.39)	(.39)
Special education			1.14	1.11
			(.98)	(.98)
<b>Between-student</b>				
Race and ethnicity <sup>a</sup>				
Black	-13.34	-10.99	-5.76	-5.51
	(.34)***	(.34)***	(.32)***	(.33)***
Latino	-12.57	-11.90	-6.50	-6.59
	(.50)***	(.49)***	(.46)***	(.46)***
Asian	9.40	8.51	6.89	6.73
	(.73)***	(.71)***	(.63)***	(.63)***
Other	-6.97	-6.13	-2.99	-2.76
	(.89)***	(.86)***	(.76)***	(.76)***
Suspended		-16.21	-9.40	-9.11
		(.51)***	(.47)***	(.47)***
Female			-2.09	-2.02
			(.25)***	(.25)***
Free/reduced lunch			-11.14	-10.62
			(.30)***	(.31)***
Special education			-24.21	-24.06
			(.45)***	(.45)***
Two-parent family				1.68
				(.31)***
Constant	228.01	230.21	237.05	235.65
	(.62)***	(.62)***	(.61)***	(.66)***
N	16,284	16,284	16,284	16,284
obs	25,221	25,221	25,221	25,221
$\rho$	.86	.85	.81	.81
Wald $X^2$	4,460.13***	5,640.90***	11,539.23***	379.40***

Notes: Unstandardized coefficients, standard errors in parentheses; models control for dichotomous school indicators.

<sup>a</sup>Omitted category is white.

\*  $p < .05$  \*\*  $p < .01$  \*\*\*  $p < .001$   
(two-tailed tests)

point deficit relative to those without a suspension, that deficit grows to nine points at the end of the two-year study period. Students with an early suspension experience no significant growth in math achievement. Students with two years of suspension do demonstrate modest growth (three points), but they begin with a much larger eight-point deficit relative to never-suspended students. By the end



**Figure 1.** Predicted Values of MAP Math Scores Over Time as a Function of Suspensions

Note: Based on Model 3c in Table 4.

of the study period, that deficit has grown to 11 points. Importantly, *this figure suggests that when students who were initially at risk for low performance are suspended, this event places them at further risk of academic decline.*

### Reproduction of Racial Inequality through Exclusionary Discipline

The first set of analyses demonstrates that racial and ethnic minorities are disproportionately susceptible to suspension. This effect is particularly pronounced for black students, and this effect is only partially explained by socioeconomic status, family structure, and other variables. The suspension disparity operates at both the school and individual levels such that black students are more likely than white students to attend schools that employ higher levels of exclusionary discipline, and black students are also more likely to be suspended than their white peers within the same schools. In turn, racial and ethnic minorities underperform on reading and math achievement tests relative to white students in this school system. As shown in Model 2 of Tables 3 and 4, adding between- and within-person measures of suspension to the regression of academic achievement on race and ethnicity reduces the effect of minority status. To assess the extent to which group differences in exclusionary discipline experiences explain the racial and ethnic academic achievement gap, mediation analyses with a bootstrapped estimation of the indirect effect are conducted. These findings suggest that 20 percent of the effect of being black on reading achievement ( $b = -2.07$ ;  $p < .001$ ) and 17 percent on math achievement ( $b = -2.24$ ;  $p < .001$ ) works indirectly through inequalities in exclusionary discipline experiences. In other words, *the racial achievement gap for black students is reproduced in part through disproportionate exposure to exclusionary discipline in public schools.*

## DISCUSSION

Our analysis provides evidence that school suspension contributes to racial inequalities in achievement. According to our results, African Americans and Latinos are disproportionately susceptible to suspension. Because there are fixed-effects parameters at the school level, this result cannot be

explained by racial and ethnic segregation into different kinds of schools. In other words, African Americans and Latinos are more likely to be suspended than whites and Asians within the same school. For African Americans, this finding persists even after controlling for socioeconomic status and other relevant individual-level variables.

Results indicate that suspension has important linkages to student academic achievement. Students who have been suspended score substantially lower on end-of-year academic progress tests than those who have not, and even students with a propensity to be suspended perform worse in years where they are suspended relative to years when they are not. We find that the effects of suspension are long lasting, setting into motion a trajectory of poor performance that continues in subsequent years, even if a student is not suspended again. Indeed, our results show that academic growth drops precipitously after one early suspension (see Figure 1). In all, our analysis provides strong evidence that suspension is harmful to academic achievement.

As hypothesized, the most striking finding from this research is the important association between suspension and patterns of achievement disparity. Our study is the first to our knowledge to directly examine the implications of racial differences in punishment for racial differences in achievement. The results support the proposition that school discipline is a major source of the racial achievement gap and educational reproduction of inequality (Gregory et al. 2010). Particularly for African American students in our data, the unequal suspension rate is one of the most important factors hindering academic progress and maintaining the racial gap in achievement. Consistent with previous research, we find that family economic background and family structure explain much, but certainly not all, of the achievement gap (Hedges and Nowell 1999) and the discipline gap (Skiba et al. 2002).

Our findings add a critical new dimension to the long-standing discussion of academic disparities by race. Recent perspectives on the achievement gap emphasize a complex interplay of between-school, within-school, and non-school factors, instead of an either-or view (Berends et al. 2008; Condrón et al. 2012). We agree with this multifaceted approach, and our findings on school discipline align with each set of factors. The discipline disparities we observe emanate at least partially from the types of schools black students attend (Condrón 2009). In addition, home-based inequalities are undoubtedly an important part of why suspension reduces achievement (Downey et al. 2004), as schools send suspended students home, often with little academic guidance or oversight. However, our findings on school punishment most directly add to the notion that practices *within* schools contribute to the achievement gap.

Because we find that school discipline is related to racial differences in achievement, we cast our findings as a possible example of hidden inequality embedded within routine educational practices. Scholars of race assert that subtle, covert forms of discrimination are major drivers of racial inequality in the post-civil rights “color-blind” era (Bonilla-Silva 2006; Pager and Shepherd 2008; Quillian 2006). Such inequality occurs indirectly, through the routine enactment of everyday institutional policies and procedures (Pager and Shepherd 2008). Similarly, education scholars from Pierre Bourdieu and Jean-Claude Passeron (1977) to Karolyn Tyson (2011) have argued that seemingly neutral processes in schools conceal certain biases and reproduce inequalities. Indeed, purportedly neutral discipline policies that increase the overall use of suspension in schools (e.g., zero tolerance) have been shown to exacerbate the racial gap in suspension (Hoffman 2014; Verdugo 2002).

Although we lack the data to test racial bias in discipline directly, we do show an alarming racial gap in punishment even after controlling for a host of background variables. This racial difference indicates that the enactment of discipline, while likely holding no discriminatory intent, nevertheless generates *de facto* racial inequalities. While it is possible that black students simply misbehave more than white students, previous studies have found racial discrepancies in how punishment is administered, even for similar offenses (Ferguson 2000; Morris 2005; Skiba et al. 2002). Thus, our results align with evidence of racial inequality (even if subtle and unrecognized) in school punishment. Our analysis advances this research by linking such punishment to disparate academic outcomes. Future

research could complement our study by fleshing out the micro-level processes of discipline and academic progress in greater detail.

### Limitations and Future Directions

While our research reveals a strong relationship between school suspension and achievement, it also has limitations. The primary limitation is that our data, while longitudinal, cannot prove a causal link between suspension and achievement. In particular, unmeasured endogenous factors could be driving the association between exclusionary discipline and achievement.

One of the most likely intervening factors is that black students could demonstrate worse behavior on average, which would lead to more suspensions for black students. This same behavior could also interfere with the learning process, resulting in lower achievement. We do not possess the data to directly examine differences between student behavior and the discipline they receive. However, we can draw from previous studies, which have noted that minority students are disciplined more harshly than white students for similar misbehavior. For example, in an analysis that controls for teacher-reported behavior, [Michael Rocque and Raymond Paternoster \(2011\)](#) found that “(racial) disproportionality in discipline is not explained by differential behavior” (p. 662). Moreover, qualitative studies by [Ann Arnett Ferguson \(2000\)](#) and [Edward Morris \(2005\)](#) have shown that black and Latino students are more closely monitored and more often punished than white students for similar types of infractions. According to [Ferguson \(2000:68\)](#), school officials tend to interpret behavior through a “racialized key” that accentuates transgressions of minority students. Thus, while we cannot assert it definitively based on our data, we can look to previous research to suggest that disciplinary policies and interpretations within schools contribute to at least part of the racial disparity in discipline.

However, we think that the association between discipline and behavior is ultimately complex, and begs further study. Student behavior, discipline, and achievement interact as students progress through schooling. The challenge for future research is to plumb this relationship further to gain a deeper picture of the mechanisms producing differences in punishment. It would be especially fruitful for studies to examine students’ progress over time, as they transition across various levels of schooling, and to examine the types of academic resources students have access to after a suspension. It would also be useful to compare the effects of different types of discipline to ascertain whether any act of punishment is associated with diminished achievement, or whether it is exclusionary discipline per se. Likewise, an important next step is assessing whether missed instruction is a mechanism of our findings. For instance, future research should compare the influence of missed instruction due to suspension and other causes (e.g., illness, truancy, etc.) to determine whether it is punishment per se or lost classroom time in general that underlies the link between exclusionary discipline and achievement.

Another significant limitation is that we do not possess data on student perceptions of discipline and relationships with school officials. Even when strict, if students perceive discipline as *fair*, this may foster a positive relationship with school and result in higher achievement ([Arum 2003](#); [Kupchik and Ellis 2008](#)). For minority students in particular, developing supportive bonds with institutional officials appears critical for academic success ([Conchas 2006](#); [Stanton-Salazar 1997](#)). These bonds may be enhanced by minority teachers, who tend to assess the behavior of minority students more positively ([Downey and Pribesh 2004](#); [McLoughlin and Noltemeyer 2010](#); [Quiocho and Rios 2000](#); [Rocha and Hawes 2009](#)), but such teacher-student dynamics are complex ([McGrady and Reynolds 2013](#)). Future research should examine student perceptions of discipline and relationships with school officials as potentially important factors in school punishment disparities.

### CONCLUSION

This study adds a critical new piece to the puzzle over racial disparities in achievement. In particular, it demonstrates how exclusionary forms of punishment such as suspension have important, racialized academic consequences. Our study presents evidence that disparate suspension lowers school



performance and contributes to racial gaps in achievement. Discipline is a necessary condition for student learning. However, unequal exclusionary discipline severely restricts opportunities for students to learn and grow. For genuine progress to be made in closing the racial achievement gap, we must also make progress in closing the racial punishment gap.

## REFERENCES

- Ainsworth-Darnell, James W. and Douglas B. Downey. 1998. "Assessing the Oppositional Culture Explanation for Racial/Ethnic Differences in School Performance." *American Sociological Review* 63:536–53.
- American Psychological Association. 2008. "Are Zero Tolerance Policies Effective in the Schools? An Evidentiary Review and Recommendations." *American Psychologist* 63:852–62.
- Arcia, E. 2006. "Achievement and Enrollment Status of Suspended Students: Outcomes in a Large, Multicultural School District." *Education and Urban Society* 38:359–69.
- Arum, Richard. 2003. *Judging School Discipline: The Crisis of Moral Authority*. Cambridge, MA: Harvard University Press.
- Aud, Susan, Mary Ann Fox, and Angelina KewalRamani. 2010. *Status and Trends in the Education of Racial and Ethnic Groups (NCES 2010–015)*. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Baron, R. M. and D. A. Kenny. 1986. "Moderator-Mediator Variables Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51:1173–82.
- Belfield, Clive R. and Henry M. Levin. 2007. *The Price We Pay: Economic and Social Consequences of Inadequate Education*. Washington, DC: Brookings Institution Press.
- Berends, Mark, Samuel R. Lucas, and Roberto V. Penaloza. 2008. "How Changes in Families and Schools are Related to Trends in Black-White Test Scores." *Sociology of Education* 81:313–44.
- Bonilla-Silva, Eduardo. 2006. *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in the United States*. New York: Roman and Littlefield.
- Bourdieu, Pierre and Jean-Claude Passeron. 1977. *Reproduction in Education, Society, and Culture*. London, UK: Sage Publications.
- Children's Defense Fund. 1975. *School Suspensions: Are They Helping Children?* Cambridge, MA: Washington Research Project.
- Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederick D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.
- Conchas, Gilberto Q. 2006. *The Color of Success: Race and High Achieving Urban Youth*. New York: Teachers College Press.
- Condron, Dennis J. 2009. "Social Class, School and Non-School Environments, and Black/White Inequalities in Children's Learning." *American Sociological Review* 74:683–708.
- Condron, Dennis J., Daniel Tope, Christina R. Steidl, and Kendralin J. Freeman. 2012. "Racial Segregation and the Black-White Achievement Gap, 1992–2009." *The Sociological Quarterly* 54:130–57.
- Condron, Dennis J. and Vincent J. Roscigno. 2003. "Disparities Within: Unequal Spending and Achievement in an Urban School District." *Sociology of Education* 76:18–36.
- Contenbader, Virginia and Samia Markson. 1998. "School Suspension: A Study with Secondary School Students." *Journal of School Psychology* 36:59–82.
- Corcoran, Sean P. and William N. Evans. 2008. "The Role of Inequality in Teacher Quality." Pp. 212–49 in *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Davis, J. E. and W. J. Jordan. 1994. "The Effects of School Context, Structure, and Experiences on African American Males in Middle and High Schools." *Journal of Negro Education* 63:570–87.
- Downey, Douglas B., Paul T. von Hippel, and Beckett A. Broh. 2004. "Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year." *American Sociological Review* 69:613–35.
- Downey, Douglas B. and Shana Pribesh. 2004. "When Race Matters: Teachers' Evaluations of Students' Classroom Behaviors." *Sociology of Education* 77:267–82.
- Ekstrom, R. B., M. E. Goertz, J. M. Pollack, and D. A. Rock. 1986. "Who Drops out of High School and Why? Findings from a National Study." *Teachers College Record* 87:356–73.
- Ferguson, Ann Arnett. 2000. *Bad Boys: Public Schools in the Making of Black Masculinity*. Ann Arbor: University of Michigan Press.

- Fordham, Signithia and John U. Ogbu. 1986. "Black Students' School Success: Coping with the Burden of 'Acting White'." *Urban Review* 18:176–206.
- Garland, David. 2001. *The Culture of Control: Crime and Social Order in Contemporary Society*. Chicago: University of Chicago Press.
- Gregory, Anne, Russell J. Skiba, and Pedro Noguera. 2010. "The Achievement Gap and the Discipline Gap: Two Sides of the Same Coin?" *Educational Researcher* 39:59–68.
- Grissmer, David and Elizabeth Eiseman. 2008. "Can Gaps in the Quality of Early Environments and Noncognitive Skills Help Explain Persisting Black-White Achievement Gaps?" Pp. 139–80 in *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Harris, Angel L. 2011. *Kids Don't Want to Fail: Oppositional Culture and the Black-White Achievement Gap*. Cambridge, MA: Harvard University Press.
- Hedges, Larry V. and Amy Nowell. 1999. "Changes in the Black-White Gap in Achievement Test Scores." *Sociology of Education* 72:111–35.
- Hirschfield, Paul. 2008. "Preparing for Prison? The Criminalization of School Discipline in the USA." *Theoretical Criminology* 12:79–101.
- Hirschi, Travis. 1969. *The Causes of Delinquency*. Berkeley: University of California Press.
- Hoffman, Stephen. 2014. "Zero Benefit: Estimating the Effect of Zero Tolerance Discipline Policies on Racial Disparities in School Discipline." *Educational Policy* 28:69–95.
- Irwin, Katherine, Janet Davidson, and Amanda Hall-Sanchez. 2013. "The Race to Punish in American Schools: Class and Race Predictors of Punitive School-Crime Control." *Critical Criminology* 21:47–71.
- Jencks, Christopher and Meredith Phillips, eds. 1998. *The Black-White Test Score Gap*. Washington, DC: Brookings Institution Press.
- Jenkins, Patricia H. 1997. "School Delinquency and the School Social Bond." *Journal of Research in Crime and Delinquency* 34:337–67.
- Kewal Ramani, Angelina, Lauren Gilbertson, Mary Ann Fox, and Stephen Provasnik. 2007. *Status and Trends in the Education of Racial and Ethnic Minorities*. Washington, DC: National Center for Educational Statistics. Retrieved April 11, 2011 (<http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007039>).
- Kupchik, Aaron. 2010. *Homeroom Security: School Discipline in an Age of Fear*. New York: New York University Press.
- Kupchik, Aaron and Nicholas Ellis. 2008. "School Discipline and Security: Fair for All Students?" *Youth & Society* 39:549–74.
- Kupchik, Aaron and Torin Monahan. 2006. "The New American School: Preparation for Post-Industrial Discipline." *British Journal of Sociology of Education* 27:617–31.
- Losen, Daniel J. 2011. *Discipline Policies, Successful Schools, and Racial Justice*. Boulder, CO: National Education Policy Center.
- Losen, Daniel J. and Jonathan Gillespie. 2012. *Opportunities Suspended: The Disparate Impact of Disciplinary Exclusion from School*. Los Angeles, CA: The Civil Rights Project.
- MacKinnon, D. P. and J. H. Dwyer. 1993. "Estimating Mediated Effects in Prevention Studies." *Evaluation Review* 17:144–58.
- MacKinnon, D. P., G. Warsi, and J. H. Dwyer. 1995. "A Simulation Study of Mediated Effect Measures." *Multivariate Behavioral Research* 30:41–62.
- Magnuson, Katherine and Jane Waldfogel, eds. 2008. *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*. New York: Russell Sage Foundation.
- McCarthy, John D. and Dean R. Hoge. 1987. "The Social Construction of School Discipline: Racial Disadvantage out of Universalistic Process." *Social Forces* 65:1101–20.
- McGrady, Patrick B. and John R. Reynolds. 2013. "Racial Mismatch in the Classroom: Beyond Black-White Differences." *Sociology of Education* 86:3–17.
- McLoughlin, Cavin S. and Amity L. Noltemeyer. 2010. "Research into Factors Contributing to Discipline Use and Disproportionality in Major Urban Schools." *Current Issues in Education* 13(2). Retrieved December 8, 2014 (<http://cie.asu.edu/>).
- Morris, Edward W. 2005. "'Tuck in that Shirt!': Race, Class, Gender, and Discipline in an Urban School." *Sociological Perspectives* 48:25–48.
- National Center for Education Statistics (NCES). 2014. *2013 Mathematics and Reading Assessments*. Retrieved June 12, 2014 (<http://nces.ed.gov/nationsreportcard/>).
- Pager, Devah and Hana Shepherd. 2008. "The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets." *Annual Review of Sociology* 34:181–209.

- Peguro, Anthony A. and Zahra Shekarkhar. 2011. "Latino/a Student Misbehavior and School Punishment." *Hispanic Journal of Behavioral Sciences* 33:54–70.
- Perry, Brea L. and Edward W. Morris. 2014. "Suspending Progress: Collateral Consequences of Exclusionary Punishment in Public Schools." *American Sociological Review* 79:1067–87.
- Quillian, Lincoln. 2006. "New Approaches to Understanding Racial Prejudice and Discrimination." *Annual Review of Sociology* 32:299–328.
- Quirocho, Alice and Francisco Rios. 2000. "The Power of Their Presence: Minority Group Teachers and Schooling." *Review of Educational Research* 70:485–528.
- Raudenbush, Stephan W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models*. 2d ed. Thousand Oaks, CA: Sage Publications.
- Rocha, Rene R. and Daniel P. Hawes. 2009. "Racial Diversity, Representative Bureaucracy, and Equity in Multiracial School Districts." *Social Science Quarterly* 90:326–44.
- Rocque, Michael. 2010. "Office Discipline and Student Behavior: Does Race Matter?" *American Journal of Education* 116:557–81.
- Rocque, Michael and Raymond Pasternoster. 2011. "Understanding the Antecedents of the School-to-Jail Link: The Relationship between Race and School Discipline." *The Journal of Criminal Law and Criminology* 101:633–65.
- Rumberger, R. and G. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *Teachers College Record* 107:1999–2045.
- Simon, Jonathan. 2007. *Governing through Crime: How the War on Crime Transformed American Democracy and Created a Culture of Fear*. New York: Oxford University Press.
- Skiba, Russell J. and Reece L. Peterson. 1999. "The Dark Side of Zero Tolerance: Can Punishment Lead to Safe Schools?" *Phi Delta Kappan* 80:372–82.
- Skiba, Russell J., Robert S. Michael, Abra Carroll Nardo, and Reece L. Peterson. 2002. "The Color of Discipline: Sources of Racial and Gender Disproportionality in School Punishment." *The Urban Review* 34:317–42.
- Sobel, M. E. 1986. "Some New Results on Indirect Effects and their Standard Errors in Covariance Structure Models." *Sociological Methodology* 16:159–86
- Stanton-Salazar, Ricardo. 1997. "A Social Capital Framework for Understanding the Socialization of Racial Minority Children and Youths." *Harvard Educational Review* 67:1–41.
- StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Tyson, Karolyn. 2011. *Integration Interrupted: Tracking, Black Students, and Acting White after Brown*. New York: Oxford University Press.
- U.S. Department of Education. 2007. *Household Education Surveys Program*. National Center for Education Statistics. Retrieved September 2, 2014 (<http://nces.ed.gov/nhes/index.asp>)
- . 2014. *Guiding Principles: A Resource Guide for Improving School Climate and Discipline*. Retrieved June 6, 2014 ([www2.ed.gov/policy/gen/guid/school-discipline/guiding-principles.pdf](http://www2.ed.gov/policy/gen/guid/school-discipline/guiding-principles.pdf).)
- Verdugo, Richard R. 2002. "Race-Ethnicity, Social Class, and Zero-Tolerance Policies: The Cultural and Structural Wars." *Education and Urban Society* 35:50–75.
- Vigdor, Jacob L. and Jens Ludwig. 2008. "Segregation and the Test Score Gap." Pp. 181–211 in *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Wacquant, Loic. 2001. "Deadly Symbiosis: When Ghetto and Prison Meet and Mesh." *Punishment and Society* 3:95–134.
- Wallace, John M., Sara Goodkind, Cynthia M. Wallace, and Jerald G. Bachman. 2008. "Racial, Ethnic, and Gender Differences in School Discipline among U.S. High School Students: 1991-2005." *The Negro Educational Review* 59:47–62.
- Welch, Kelly and Allison Ann Payne. 2010. "Racial Threat and Punitive School Discipline." *Social Problems* 57:25–48.
- Wildeman, Christopher. 2009. "Parental Imprisonment, the Prison Boom, and the Concentration of Childhood Disadvantage." *Demography* 46:265–80.