An Empirical Examination of the Effects of Suspension and Suspension Severity on Behavioral and Academic Outcomes

Christina LiCalsi, David Osher, Paul Bailey

American Institutes for Research

AUGUST 2021

AIR®

Advancing Evidence.
Improving Lives.

An Empirical Examination of the Effects of Suspension and Suspension Severity on Behavioral and Academic Outcomes

AUGUST 2021

Authors:
Christina LiCalsi, David Osher, Paul Bailey

American Institutes for Research®
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

Contents

Figures

Tables

Abstract

The use of exclusionary discipline practices, such as out-of-school suspension (OSS) and in-school suspension (ISS), is prevalent in the United States. Of the 50.6 million students enrolled in K–12 public schools in 2015–16, 2.7 million students received one or more OSS. In addition, substantial disparities persisted for certain subgroups of students, such as Black students and students with disabilities (SWD) (U.S. Department of Education, Office for Civil Rights, 2018).

A large body of research has found that exclusionary discipline practices are associated with negative educational outcomes for individual students and their peers (Noltemeyer et al., 2015). More recent quasi-experimental evidence suggests that these associations might be causal (Hinze-Pifer & Sartain, 2018; Hwang, 2018; Hwang & Domina, 2020; Lacoe & Steinberg, 2019; Steinberg & Lacoe, 2018). Yet little is known about how the type and length of suspension are related to academic and nonacademic outcomes for disciplined students and their peers. In addition, to our knowledge, no prior study has compared later academic and nonacademic outcomes for students with the same disciplinary incidents but different disciplinary responses, or for the peers of such students.

In this paper, we build on the evidence by examining, using a quasi-experimental propensity score–weighting method, the effects of the type and length of exclusionary disciplinary responses on (a) middle and high school students' educational outcomes, (b) their same-school same-grade peers' educational outcomes, and (c) school climate. To accomplish this, we used linked disciplinary, demographic, and academic administrative data from the New York City Department of Education (NYCDOE). We found that more severe exclusionary discipline does not serve as a deterrent to students' future reported behavior, and for younger students it may instead exacerbate it. In addition, more severe exclusionary discipline has a consistent negative effect on many other long-term educational outcomes for students. Receiving a more severe exclusionary disciplinary response to an incident increases the number of days students miss due to absence during subsequent school years, increases the number of days they miss due to suspension in subsequent school years, decreases their likelihood of earning both English language arts (ELA) and math credits throughout their high school career, and decreases their likelihood of graduating. The severity of exclusionary disciplinary response has no effect on the behavior, academic outcomes, or attendance of peers in the same grade within the disciplined student's school, nor does it have effects on students' or teachers' perceptions of school climate.

Introduction

Exclusionary discipline—including expulsion, transfers to alternative schooling environments, out-of-school suspension (OSS), in-school suspension (ISS), and classroom removal—has deep roots in U.S. history, and its use has increased dramatically during the last 3 decades of the 20th century (Wald & Losen, 2003). It disproportionately affects students of color, students with disabilities (SWD), and students from other historically disadvantaged groups (Losen & Whitaker, 2018; Noltemeyer et al., 2015) and has been linked to a host of negative outcomes, including poor grades, disengagement, chronic absenteeism, grade repetition, dropout, lower graduation rates, adult mental illness, and incarceration (Hwang, 2018; Mendez-Raffaele & Knoff, 2003; Monahan et al., 2014; Morris & Perry, 2016; Noltemeyer et al., 2015; Wolf & Kupchik, 2017).

Although there has been some reduction in suspension and other forms of exclusionary discipline use over the past decade, its use remains widespread, and disparities in its application persist (Gregory et al., 2021). Of the 50.6 million students enrolled in K–12 public schools in 2015–16, 2.7 million students received one or more OSS (U.S. Department of Education, Office for Civil Rights, 2018). Nationally, school children lost more than 11 million days of instruction as a result of OSS in 1 school year alone (11,360,004 days during 2015–16), and days lost were even greater for students from historically disadvantaged groups. During the same year, Black students lost 66 days of instruction to OSS compared to 14 days for White students; that is, Black students lost nearly 5 times the amount of instruction as White students. Additionally, SWD lost more than double the number of days as their nondisabled peers (44 days compared to 20 days) (Losen & Whitaker, 2018). Research syntheses suggest that these disparities in suspension are inextricably linked to achievement gaps and the ability to engage students in deeper learning (Gregory et al., 2010; Osher et al., 2020). One analysis found that differences in suspension rates account for one fifth of the Black-White test score gap (Gopalan, 2019; Morris & Perry, 2016).

Concerns with reliance on exclusionary discipline, disparities in its application, and its consequences on students have led to systemic efforts to incorporate less punitive disciplinary approaches, such as positive behavioral interventions and supports (PBIS), social-emotional learning (SEL) curriculum, trauma-sensitive approaches, and restorative practices, into schools (Gregory et al., 2021). Although an overwhelming number of correlational, qualitative, and longitudinal studies exist demonstrating the negative relationship between suspension and other exclusionary practices and social, emotional, and academic outcomes for students, causal evidence about the effects of exclusionary discipline on excluded students and their peers remains relatively scant.

Current Study

In this paper, we build on the evidence by examining, using quasi-experimental propensity score–weighting methods, the effects of the type and length of exclusionary disciplinary responses on (a) middle and high school students' educational outcomes, (b) their same-school same-grade peers' educational outcomes, and (c) teachers' and students' reports of school climate. We also examine these questions for student subgroups of interest. To accomplish this, we used linked disciplinary, demographic, and academic administrative data from the New York City Department of Education (NYCDOE).

As is always the case, from a causal standpoint, the ideal way to estimate the unbiased effect of one disciplinary response as compared to another is to randomly assign disciplinary responses to every behavioral incident that occurs and compare later outcomes between students who were reported for the same incident but received different responses. This is not, however, ethical, given what is already known about the harmful effects of suspension. Thus, the research base is mired with omitted-variable bias, and the causal evidence from randomized controlled trials (RCTs) is limited to school-level interventions that aim to reduce the use of exclusionary discipline, often for specific types of infractions. These experiments do not allow for an examination of the effect of changes in severity, nor do they allow for a comparison of responses to a wide variety of incident types.

We used propensity score weighting to balance both detailed data on behavioral incidents as well as student and school covariates among groups of students that received different severity levels of punishment. We created a quasi-experimental framework in which propensity score weighting created optimally equivalent groups of students, thus allowing us to come closer than previous studies to controlling for omitted-variable bias. Because of the breadth of our data, both in terms of the number of incidents as well as the available covariates, we used extreme gradient boosting (XGB), tuned with cross-validation (a type of machine learning method), to automatically select variables and interaction terms.

Using these models, past events can be described and future outcomes predicted based on (a) characteristics of the students; (b) characteristics of their school, including prior year measures of school climate; (c) measures of the outcomes of interest preceding the incident (e.g., students' attendance, grades, and behavior in prior years); (d) information on the incident itself (e.g., the exact type of incident, the number of students involved, the date of the incident); and (e) information on the outcomes of interest. We were then able to use this information to create groups of students that are balanced on everything we know about them—their demographic, academic, and behavioral characteristics; their school environment; and their specific behavioral incidents—essentially everything that would predict both the severity of discipline and their outcomes, except that they received a different disciplinary

response. Then, predictions can be compared across decisions. For example, simplifying to a situation in which there are only two possible choices related to discipline for a student, each outcome can be thought of as having two possible future values depending on which disciplinary decision is made in response to the student's transgression: one if a more severe exclusionary response is assigned and the other if a less severe response is assigned. The goal is to estimate how each outcome would change if one response or another was assigned for a particular incident.

Machine learning can support these analyses. Several machine learning methods offer the ability to automatically select features (covariates) of a student's record that are associated with an outcome and develop a model for predicting that outcome in what amounts to a matching framework in which students with similar histories—minus the disciplinary response to the particular incident—are compared. Specifically, we employed propensity score weighting based on a gradient-boosting model to balance covariates of interest (Chen & Guestrin, 2016; McCaffrey et al., 2013). We further used doubly robust estimation in which we regression-adjusted for covariates that were slightly imbalanced (McCaffrey et al., 2013). When balance is achieved, these propensity score weights can be used to estimate the causal effect of increasing the severity of the exclusionary discipline response on those who were, in fact, excluded more severely.

The gradient-boosting model approximates the probability of receiving the various penalties (including in a multinomial model) with a successive series of interaction terms. At each step, it minimizes the residual error (conditional on all previous steps) by adding a single binary tree of a tuned depth that classifies the incidents by covariates and then assigns an additive term to the predicted probability of receiving each outcome based on the final cell. The results from these trees are not directly added in but attenuated, in order to slow learning and avoid overfitting. Hyperparameters, fit with cross-validation, determine the tree depth (number of branches per tree, which is also the maximum number of interaction terms) and how much the learning in each tree is slowed. For a full description of XGB, see Chen and Guestrin (2016).

An advantage of adding binary trees is that it adds every variable nonparametrically and with as many interactions as possible. How variables, cut points, and interactions are chosen is in the model and not the result of analyst interpretation/interference. This allowed us to control for the large amount of information about behavioral incidents and the large number of covariates without concern for overfitting or having to select interaction terms. The overfitting is handled explicitly—though cross-validation—rather than implicitly by limiting the number of variables included in the model. Using this method allowed us to take advantage of the incredible detail contained in the NYCDOE data and include more than 80 variables in our propensity weighting models.

We found that more severe types and lengths of exclusionary discipline have no positive effect on students' future involvement in behavioral incidents, suggesting that more severe exclusionary punishment fails to reduce students' undesirable behavior either through changing their internal thought processes (i.e., the student learns that the behavior was wrong and changes in order to be a better person) or even by simply serving as an external deterrent (i.e., the student changes their behavior, not due to a change in desire to be a better person but in a desire not to experience the punishment). In addition, for younger students there is evidence that more severe punishment can result in increased reporting of behavioral infractions in the future, an indication that exclusionary discipline either results in an increase in negative behavior in school for younger students, increased perception and reporting of student misbehavior by teachers and administrators, or a combination of these responses.

Receiving a more severe type and length of exclusionary discipline in response to a behavioral incident also has substantial long-term negative effects on students, including increases in the number of days students miss due to absence, increases in the number of days they miss due to suspension, decreases in their likelihood of earning both English language arts (ELA) and math credits, and decreases in their likelihood of graduating. The severity of exclusionary disciplinary response has no effect on same-school same-grade peers' academic outcomes or attendance, nor does it have an effect on students' or teachers' perceptions of school climate. Effects are consistent across student racial and ethnic subgroups and for SWD and students from economically disadvantaged backgrounds.

Background

Exclusionary practices, such as suspension and expulsion, have a long history in U.S. education, and these practices have been disproportionately applied to students of color, students with a high incidence disabilities, and other minoritized groups (Osher, 2016). During the last decades of the 20th century, politicized and racialized fears concerning violence in neighborhoods and schools led to a dramatic increase in the implementation of zero-tolerance school discipline policies, disproportionately affecting students of color and students with emotional and behavioral challenges (Skiba, 2014). National estimates suggest that the proportion of students suspended or expelled doubled between 1970 and the early 2000s (Losen, 2011; Wald & Losen, 2003). These zero-tolerance policies are grounded in deterrence theory, presuming that harshly punishing a student will have a deterrent effect on future misbehavior for the offending student and the student's peers (Ewing, 2000) and that the removal of certain students will yield a more productive learning climate for those students who remain (Public Agenda, 2004). Many districts and schools adopted further zero-tolerance policies for nonviolent and minor

incidents, such as a school uniform violation or refusing to turn off a cell phone (Watanabe, 2013). In fact, nonviolent behaviors have accounted for an increase in the share of all suspensions for 5+ days, transfers to specialized schools, and expulsions (Steinberg & Lacoe, 2017). By 2001, 90% of school systems had implemented some form of zero-tolerance or three-strikes discipline policy (American Civil Liberties Union, 2017).

Critics argue that the widespread use of exclusionary discipline and the disparities in its application are particularly concerning in light of a large body of research finding that exclusionary discipline is not only ineffective at producing positive behavioral change but also linked with a host of short- and long-term negative outcomes in schools, in the community, and intergenerationally (Anderson et al., 2019; Dong & Krohn, 2020; Hemphill et al., 2013; Monahan et al., 2014; Rosenbaum, 2020). Students who are suspended once are at greater risk of being suspended again in the future (Mendez-Raffaele & Knoff, 2003), receive poorer grades, and are at greater risk of dropping out (Hwang, 2018; Morris & Perry, 2016; Noltemeyer et al., 2015). In addition, exclusionary discipline puts students at greater risk of becoming involved with the juvenile justice system (Monahan et al., 2014) and increases the likelihood that they will experience criminal victimization and incarceration as adults (Hughes et al., 2020; Osher et al., 2002; U.S. Department of Education & Department of Justice, 2014; Wolf & Kupchik, 2017). These findings are meaningful to matters of intervention and systems change, as empirically demonstrated alternatives exist (Osher et al., 2010; Valdebenito et al., 2019). In addition, similar students attending demographically similar schools experience variant outcomes (Anderson & Ritter, 2017; Fabelo et al., 2011; Sartain et al., 2015). For example, Welsh and Little (2018) found that disparities in disciplinary outcomes may be better explained by the behavior of teachers and principals in schools rather than student characteristics such as misbehavior, poverty, or race—an encouraging finding, as the behavior of teachers and principals may be more readily changed than complex underlying economic, political, and social structures.

Over the past 2 decades, calls for curtailing suspension and developing alternative approaches to managing students' behavior have come from students, families, judges, civil rights organizations, educational associations, and school safety researchers (AASA, n.d.; Academy Council on School Health, 2013; American Federation of Teachers, 2020; American Psychological Association, 2008; Interdisciplinary Group on Preventing School and Community Violence, 2013). Research has also accumulated on the effectiveness of alternative approaches grounded in mental health research, prevention science, applied behavioral analysis, and SEL, as well as with regard to connectedness, cultural competence, motivation, engagement, and classroom management (Bear, 2014; Bradshaw et al., 2012; Osher et al., 2008, 2016; Osher et al., 2010; Resnick et al., 1997). This research fueled a public-private effort that focused on supportive school discipline and called for minimizing the use of suspension and providing

multitiered school- and community-focused alternatives (Morgan et al., 2014; Skiba et al., 2016). Consequently, many states and districts moved toward an approach of limiting exclusions and keeping students in the classroom (Steinberg & Lacoe, 2017). Between 2012 and 2016 alone, 27 states revised or released guidance to change disciplinary practices that rely on exclusionary consequences, and more than 50 of America's largest districts reformed their discipline policies to include nonpunitive strategies, limit the use of suspensions, or both (Kamenetz, 2017).

And yet, punitive and exclusionary approaches to school discipline persist. In the 2017–18 school year, middle schools and high schools suspended an average of 7.4% of their students. Although this was a reduction from 9.6 percent in 2011–12, 2.5 million students still faced suspension in 2017–18, with Black students and SWD more than twice as likely to be suspended as White students and students without disabilities. Further, the reductions in suspension were not universal: Mississippi and South Carolina had suspension rates that were twice the national average in 2017–18, both having increased since 2011–12 (Ryberg et al., 2021). In a nationally representative survey of K–12 teachers, the proportion of teachers who agreed that ISS, OSS, and expulsion were somewhat or very effective discipline strategies were 47%, 39%, and 39%, respectively (Educators for Excellence, 2018).

Consequences of Suspension for Suspended Students

Student behavior has individual and contextual and roots (Osher et al., 2004; Osher et al., 2020). Troubling behavior, including what is viewed as antisocial behavior, may indicate deeper problems that can be examined and addressed at a student, teacher, school, and system level (Benbenishty & Astor, 2019; Dwyer et al., 2000; Osher et al., 2004; Osher et al., 2012; Osher et al., 2020). Research and intervention at the student level target social-emotional competencies, health, and cognitive needs (Dwyer et al., 2000; Gresham & Lane, 2019; Mayer & Salinger, 2019). School targets include school organizational capacity (Benbenishty et al., 2005; Diamond & Lewis, 2016; Osher et al., 2004; Sugai & Horner, 2002), educator stress, teachers' social-emotional skills, and bias (Artiles et al., 2010; Jennings et al., 2019; Osher et al., 2004). Often, these sets of needs dynamically interact (Nese et al., 2019; Osher et al., 2019) in a manner that leads to a cascade of problems, including harsh and exclusionary discipline and treatment disparities, which in turn exacerbate student alienation, disengagement, and academic and social problems (Walker et al., 2004). Exclusionary discipline and other trauma-insensitive processes may be part of this cascade (Dishion & Patterson, 2006; Guarino & Caverly, 2019; Kellam et al., 2008; Okonofua et al., 2016).

From a social control deterrence perspective, exclusionary discipline, like other forms of punitive punishment, is meant to signal the severity of the student's misbehavior, be unpleasant enough to the misbehaving student that it discourages future misbehavior, and

simultaneously impart norms and values of proper behavior to the other students in the school, thus deterring the misbehaving student's peers from misbehavior as well (Perry & Morris, 2014). Although suspension has been challenged on theoretical grounds from the perspectives of reinforcement theory (suspension may be experienced positively) and differential association theory (suspended students may be more likely to associate with antisocial as opposed to prosocial peers) and on empirical grounds based on studies of both the educational and juvenile justice systems, it remains intuitively attractive to many practitioners and politicians (Hemphill et al., 2013; Lipsey, 2009; Mayer, 1995; Mayer et al., 1983; Noltemeyer et al., 2015).

Simply removing students from the school environment does nothing to deal with students' and schools' deeper issues and may lead to further disengagement from school, anger, and erosion of trust. These negative psychosocial outcomes have been shown to exacerbate recidivism (Costenbader & Markson, 1998). Exclusionary discipline may also lead to students who feel disconnected from the school, reinforcing one another's antisocial behavior (Dodge et al., 2006). In addition, students who are removed from the instructional environment due to suspension may also experience additional negative outcomes associated with the loss of instructional time. For example, an analysis conducted in the Cleveland Metropolitan School District (Linick & D'Amico, 2014) found that for 9th- and 10th-grade students, missing 10 or more days of school was associated with a 40.9% drop in the probability of being on track to graduate.

The empirical literature, whether ethnographic, qualitative, or correlational, largely provides evidence demonstrating a negative relationship between experiencing exclusionary discipline and students' academic achievement (Balfanz et al., 2014; Balfanz et al., 2015; Carpenter & Ramirez, 2007; Chu & Ready, 2018; Fabelo et al., 2011; Suh & Suh, 2007). Noltemeyer et al. (2015) examined 24 studies on the effect of suspension on student outcomes and found a significant inverse relationship between school suspensions and academic achievement as well as a significant positive relationship between school suspensions and dropping out of school. Researchers have recently sought to use quasi-experimental methods to examine the effects of suspensions on students' outcomes and have also largely found negative consequences, although of smaller magnitude than those of correlational studies. For example, Steinberg and Lacoe (2018) used a difference-in-difference approach to examine the effect of a policy change in which schools were pushed to no longer suspend students for classroom disruptions. The authors found that the change did result in a decrease in suspension of students for classroom disruption, while also resulting in improved attendance outcomes for students suspended prior to the policy change. It should be noted that the authors did not find that behavior deemed to be disruptive decreased—only suspensions for such behavior. Unlike prior correlational research (Noltemeyer et al., 2015), the authors did not find a significant change in math or ELA performance for

students suspended prior to the policy change. This more rigorous analytic approach provides helpful insight into the effect of a specific discipline policy reform on achievement.

Using panel data, the same researchers (Lacoe & Steinberg, 2019) employed student fixed effects and instrumental variable estimates to study the effect of suspensions on math and ELA performance. They found that suspensions had a negative effect on both math and ELA performance and that students who had been suspended once were more likely to be suspended again the following school year. Lacoe and Steinberg's use of student-level fixed effects and instrumental variable analysis provides a more rigorous analysis of the relationship between being suspended and the outcomes of suspension than previous correlational studies.

A second study employing student fixed effects also found negative associations between suspension and academic achievement, although it was not consistent across types of suspension and was of a smaller magnitude than when comparing suspended students to nonsuspended students (Hwang, 2018). The author examined linked panel data sets of student suspensions and academic achievement, measured quarterly in both math and ELA. The study first compared suspended students to nonsuspended students, controlling for observable characteristics of the student and school, as well as teacher, year, and school quarter fixed effects, finding negative associations between ISS and OSS on both ELA and math achievement. The author then ran a student fixed-effects model comparing students with a suspension in a particular quarter to themselves in other quarters. Using this model, OSS was associated with a decrease in ELA achievement (of approximately half the magnitude of the prior model) but no difference in math achievement, and ISS was not associated with a difference in math or ELA achievement. The reduction in effect magnitude between models suggests that the latter controlled for omitted-variable bias that the former did not.

Sorensen et al. (2021) exploited principal turnover to examine the impact of principal-driven disciplinary decisions on the outcomes of excluded middle school students and their peers. They found that principals who were more likely to exclude students saw reductions in reported rates of minor student misconduct, suggesting a deterrence effect for these types of infractions. However, these principals also saw a reduction in high school graduation rates for all students in their schools and an increase in the rate of juvenile justice complaints. Additionally, the level of racial bias a principal exhibited—defined as the difference in propensity to remove a Black student for the same offense and same offense history as a White student—predicts substantial academic losses for Black and Hispanic students enrolled at that school.

To our knowledge, two studies have examined the effects of suspension on academic outcomes of suspended students by comparing them to other students who committed a similar behavioral infraction (Anderson et al., 2019; Swanson et al., 2017). Swanson et al. examined

grade retention between 9th and 10th grade for students who experienced an exclusionary disciplinary event in 8th grade compared to students with similar behavioral histories who did not receive exclusionary discipline. They found that students who received 1+ days of exclusionary discipline were more likely to be retained than nonexcluded students. Swanson et al. also examined whether the length of the exclusionary discipline moderated the relationship, finding that longer exclusionary discipline events had a greater negative impact than shorter events. Anderson et al. examined test scores and grade retention for K–12 students with a behavioral infraction who received more and less severe exclusionary discipline and found that more exclusionary discipline was associated with more negative academic outcomes during the year of the infraction.

Spillover Effects on Peers and Teachers

Student misconduct and the resulting consequences do not only affect the students who misbehave. One of the theoretical foundations of punitive discipline is that if nonoffending students witness misbehavior and the resulting consequences are severe, the nonoffending students should be further incentivized not to engage in such misconduct. Specifically, the substantial rise of zero-tolerance policies in schools, even for relatively minor infractions, during the 80s and 90s was based in part on the broken-windows theory that was simultaneously changing criminal policing. Wilson and Kelling (1982) argued that strict enforcement of low-level crimes would put the "real" criminals on notice that they were being watched and that their actions would have consequences, thus reducing serious crime. Applied to the educational setting, it was thus theorized that strict consequences such as suspension for relatively minor misbehavior such as talking back or being disruptive would reduce the occurrence of more major violent misbehavior.

Additionally, the primary critique of efforts to reduce exclusionary discipline is that removing misbehaving students is necessary to maintain order in the classroom (Eden, 2017, 2019). Indeed, a number of studies examining exogenous shocks to peer composition and configuration of schools have found negative effects of disruptive students on peers' learning and behavior (Carrell & Hoekstra, 2010; Deming, 2011; Figlio, 2007; Fletcher, 2010; Gottfried, 2010; Imberman et al., 2012; Kristoffersen et al., 2015). Studies leveraging the change in student composition in classrooms associated with the inclusion of students with emotional and behavioral disorders in general education classrooms found evidence of negative spillover effects on peers' behavior problems (Gottfried, 2014) as well as on their reading and math skills (Fletcher, 2010). In addition, when students with discipline problems were forced to relocate to new school districts, there were negative effects on their new peers' behavior and attendance (Imberman et al., 2012). The ability to extrapolate from these studies to the effects of exclusionary discipline more broadly are unclear, however. For example, exogenous shocks can

destabilize classes and schools, which are dynamic systems; students with emotional and behavioral disorders do not represent all suspended students; and forced student mobility has additional effects on the incoming student and receiving students.

Alternatively, more punitive discipline practices may have unintended adverse consequences on nonsuspended students. Criminological research has found that mass incarceration has had large negative effects that extend past the incarcerated individual and onto their families and communities. Mauer and Chesney-Lind (2002) termed these spillover effects of highly authoritarian approaches to social control "collateral consequences." Perry and Morris (2014) extended this perspective to the U.S. education system, conceptualizing exclusionary discipline practices as the manifestation of increased social control in schools. The authors cited work by Arum (2003), which found that school discipline is most effective when it is moderately strict, consistently applied, and perceived by students as fair. In addition, they argue that, conversely, overly punitive environments erode students' views of their schools' moral authority, resulting in alienation and resistance, because such contexts promote "legal cynicism," the perception by a group that law enforcement is illegitimate (Kirk & Papachristos, 2011), even among group members who follow the rules themselves (Kirk & Matsuda, 2011). For example, the authors drew on ethnographic evidence by Nolan (2011) and Kupchik (2010), who describe a ubiquitous sense of anxiety for all students in punitive educational settings, with the goal of social control overshadowing the goal of education.

It is also plausible that the effect of disciplinary decisions on school climate depends on the type of incident that occurred and the context of the school. Suspensions for trivial, nondisruptive, nonviolent infractions may undermine school culture, whereas suspensions for infractions that disrupt learning and jeopardize the safety of students and teachers may be viewed as justified. Social psychology and studies of adolescent development suggest that one of the consequences of suspension, even if objectively warranted, is compromising trust and reinforcing a negative climate of racial and adult control that can undermine identity safety and affirmation, connection to school, and the building of an academic community (Astor et al., 2020; Eccles et al., 1993; Ferguson, 2000; Hammond, 2020; Nasir, 2020; Rios, 2011).

Empirical studies have found mixed spillover effects of exclusionary discipline on peers' academic and behavioral outcomes and on school climate. A number of correlational studies have found that higher school-level suspension rates are associated with teachers and students feeling less safe than their peers in schools serving similar students with lower school-level suspension rates (Bradshaw et al., 2009; Lacoe, 2015; Steinberg et al., 2011). Recently, researchers have built upon this work by creatively employing fixed effects to more rigorously explore the potential spillover effects of exclusionary discipline on school climate and peer outcomes, with mixed findings.

Perry and Morris (2014) estimated the effects of suspension on peers, controlling for time-invariant heterogeneity in students and schools, finding that nonsuspended students had lower math and reading scores when school-level suspension rates rose. Hinze-Pifer and Sartain (2018) examined whether suspensions were associated with peer outcomes, using student and school fixed effects and taking advantage of policy-induced changes in school suspensions. The researchers found moderate improvements in academic outcomes and attendance as well as improvements in school climate when schools reduced their reliance on exclusionary discipline. These findings are consistent with research suggesting that excessive use of suspensions and other exclusionary discipline policies can damage classroom learning and school climate. Hinze-Pifer and Sartain found that attendance and school climate improvements were driven primarily by schools that had the highest baseline suspension rates, suggesting that reductions in suspension usage might be most important in schools that have relied particularly heavily on them in the past.

Conversely, Hwang and Domina (2020) examined the link between suspensions and nonsuspended students' learning, employing student and classroom fixed effects, and found that student achievement in math increased when classmates receive suspensions—particularly suspensions attributed to disruptive behavior—whereas no effects were found on ELA achievement. The authors note, however, that their results came from schools in which suspensions were relatively rare and therefore may not generalize to other settings. Of particular interest, the authors note that although students of color and low-income students are more likely to receive a suspension, they are also more likely to be in classes with disruptive students and thus disproportionally subjected to any negative spillover effects.

Other studies have built on the early correlational literature by exploiting changes in suspensions due to district policy changes limiting the use of suspension for more minor infractions such as insubordination (e.g., Augustine et al., 2018; Hinze-Pifer & Sartain, 2018; Steinberg & Lacoe, 2018). These studies have also found inconsistent effects. Hinze-Pifer and Sartain (2018) found improvements in academic outcomes, attendance, and school climate. Steinberg and Lacoe found a reduction in the use of suspension, particularly for classroom disruption (the intended policy target), and a small improvement in attendance among students who had been previously suspended. However, they found negative impacts on academic performance and increased truancy among peers who had not previously been suspended, which was driven by schools that did not fully implement the policy. In addition, one RCT (Augustine et al., 2018) found that policy shifts resulting in a reduced reliance on suspension improved teacher reports of school climate but harmed student achievement.

Gaps in the Empirical Research

Although numerous creative strategies have been used over the past few years to provide stronger causal evidence of the possible effect of suspension on suspended students and their peers, these studies have several limitations. As detailed in the previous section, a number of studies have used student fixed effects to control for time-invariant characteristics of students who are and are not suspended (Hinze-Pifer & Sartain, 2018; Hwang, 2018; Hwang & Domina, 2020; Lacoe & Steinberg, 2019; Perry & Morris, 2014). By using student fixed effects, students serve as their own comparison; that is, students' academic outcomes during one time period when they experienced a suspension can be compared to their academic outcomes during another time period when they did not experience a suspension. By using students as their own comparison, omitted-variable bias resulting from underlying differences between suspended and nonsuspended students is accounted for because they are the same students. These studies largely found a negative relationship between suspension and academic outcomes, although of smaller magnitude than in correlational studies.

However, this method does not control for time-variant omitted-variable bias. For example, imagine that a student who had been in stable housing has recently become homeless. It is not hard to imagine that the student might begin acting out, or that a teacher might misread and punish such depression- or stress-related behavior, responding in a manner that triggers suspension, while at the same time the student's grades may slip. Did the suspension cause the fall in grades, or did the upheaval in his home life cause him to both act out and simultaneously have difficulty concentrating in school and lack a quiet place to focus on homework? We cannot know. In addition, by using students as their own comparison, it is impossible to examine long-term effects of suspension. It is only possible to examine differences in outcomes as they occur contemporaneously with changes in suspension.

Other studies exploit changes in policy to identify whether such changes are accompanied by simultaneous changes in outcomes. These studies have three important limitations that must be considered. The first is that they only examine differences in outcomes that are the result of a decrease in suspension for particular types of behavioral infraction—generally more minor infractions, such as classroom disruption. Thus, the results are not necessarily generalizable to other types of incidents. The second is that they only examine changes in outcomes that occur in a particular context, again limiting generalizability. Exclusionary discipline is a strategy that teachers and administrators use in an attempt to create a learning environment in which students are able to learn and are held accountable for their actions. If it is removed as a possible response and teachers are not provided with the resources and training they need to respond in more positive ways to attain these same goals, we might anticipate negative effects of the policy on misbehaving students and their peers. However, if the policy shift is

accompanied by increased funding and training for PBIS, SEL, and restorative practices, the effects might change direction. For example, negative effects of academics such as those found by Augustine et al. (2018) may be the result of a lack of adequate planning and resources devoted to ensuring that implementation of restorative practices does not come at the expense of time spent on academic tasks. The third limitation, again, is that none of these studies examine long-term academic effects such as graduation, nor are they longitudinal in nature. This is particularly important because system change that involves practice change has been consistently found to take 5–8 years to be able to demonstrate the quality of intended impacts (Aladjem et al., 2010; Borman et al., 2003). Effective implementation of what is a historically wicked problem of policy and practice that challenges many educators' deeply held beliefs regarding discipline and requires changes in their practice is difficult to make sense of over the short term.

Another limitation of the current literature is that few studies have included a comparison group of students who also committed a behavioral infraction. Specifically, we identified two studies that examined the effects of suspension by comparing suspended students to other students who committed a behavioral infraction (Anderson et al., 2019; Swanson et al., 2017). However, both studies used a small number of broad categories, whereas we included 45 specific incident codes as well as additional information, such as the date of the incident and the number of people involved. Also, again, neither of these studies examined long-term academic outcomes.

Finally, to our knowledge, only three studies specifically examine the effects of the length of suspension (Lacoe & Steinberg, 2019; Mader et al., 2016; Swanson et al., 2017). Mader et al. and Swanson et al. found that longer OSS was associated with increasingly negative academic outcomes. Lacoe and Steinberg found that an OSS of 2 days has a negative effect on test scores, with additional days producing increasingly negative effects, yet the effects of receiving an ISS are concentrated in the first day. However, although all three studies found that longer suspensions were associated with more negative student outcomes, none examined spillover effects on other students or long-term academic effects for suspended students.

Data

To answer our research questions, we employed data shared with us by the NYCDOE, including data on student behavior, discipline decisions, course grades, attendance, credits, test scores, demographics, and graduation status over a 10-year period.

The cohort for this study comprises middle and high school students involved in a behavioral incident between the 2008–09 school year and the 2017–18 school year. Because outcomes are only available until 2017–18, results are always from the subset of years when data are available. For example, on-time graduation outcomes are only available for students who started high school in 2013–14 or earlier because a student who started 9th grade in 2015–16 or later has not had enough time to provide a graduation outcome by 2017–18. Similarly, an outcome such as attendance 2 years after the incident is only available for students who were involved in an incident during the 2015–16 school year or earlier. Table 1 details the types of data sources and the years and grade ranges for which data were available.

Table 1. Data Availability, by Source and Year

Data source type	Availability	Grades
Incidents	2007–08 to 2017–18	K–12
Discipline	2007–08 to 2017–18	K–12
Demographics	2008–09 to 2017–18	K–12
Attendance	2008–09 to 2017–18	K–12
Grades	2013–14 to 2017–18	1–12
Graduation	2008–09 to 2017–18	--
State assessment	2008–09 to 2017–18	3–8
ELA/math credit accumulation	2008–09 to 2017–18	6–12
School surveys, student	2008–09 to 2017–18	K–12
School survey, teacher	2015–16 to 2017–18	K–12

Note. ELA = English language arts.

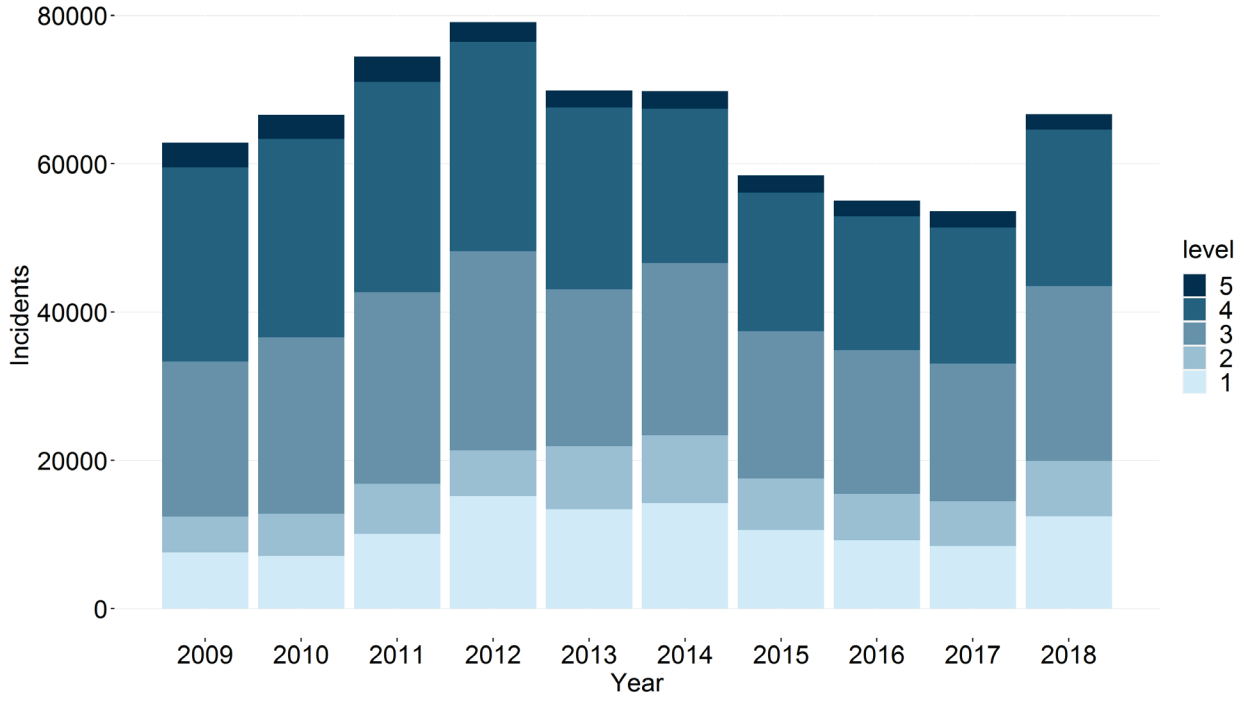
Incident and Discipline Data

The NYCDOE behavior records are broken up into two files: an incident file and a discipline file. Any student involved in an incident has a record on the incident file, but only students who received an exclusionary disciplinary response have a record in the discipline file. In the incident file, each record contains linking information for the student and school, a date, a single

incident type that indicates the most severe incident type associated with the incident,[1] the level of that incident type, a flag for if the New York Police Department (NYPD) was contacted, and a flag for if an emergency medical service (EMS) was contacted. Incidents may include several students; for example, if one or more students involved in an incident has a weapon, this would be reflected in all students' data.

Each incident type is categorized by severity into one of five levels of increasing severity.[2] The distribution of incidents by level by year for middle and high school students is shown in Figures 1 and 2, respectively (see Appendix A, Tables A.1 and A.2, for full descriptive data). Total incidents peaked in 2012 and have decreased in the following years, primarily driven by changes in the number of Level 1 incidents.

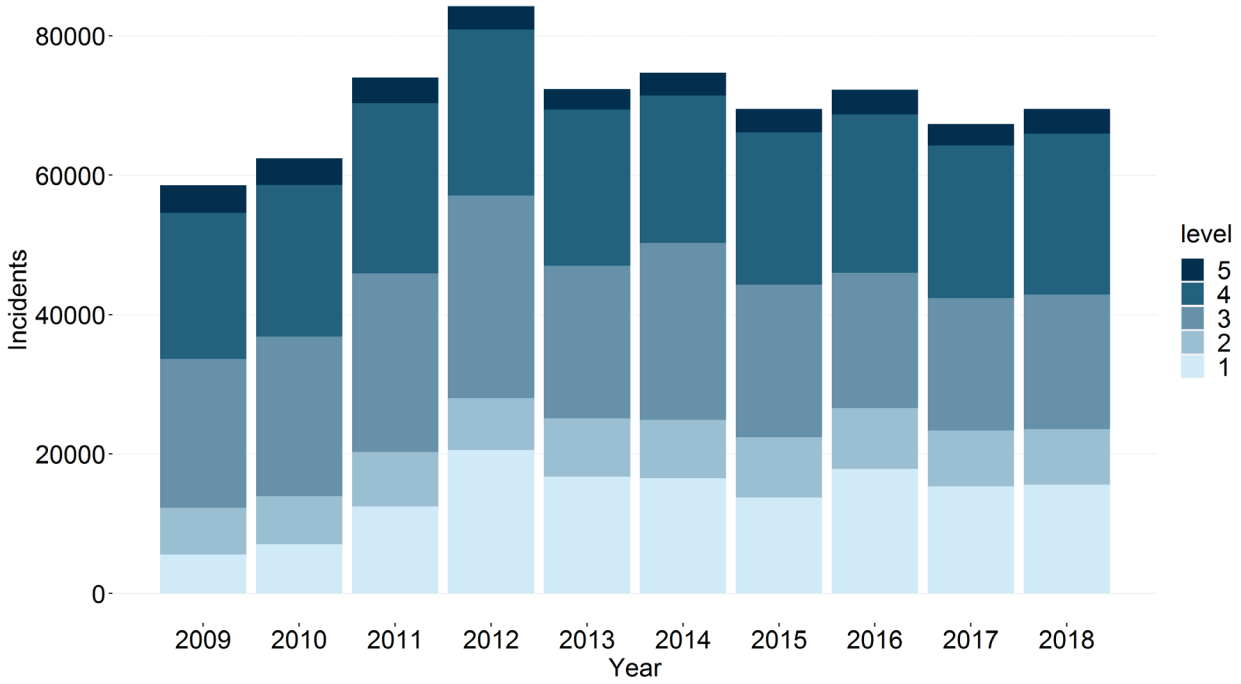
Figure 1. Incident Level by Year for Middle School Students



[1]Eighty-six incident codes are present in the data. Not all 86 are used in the analyses, however, because we limited to codes with 400 or more cases.

[2]The most common incident codes associated with each level are as follows: Level 1 are "cutting class" and "disrupting the educational process"; Level 2 are "profane, obscene, vulgar language or gestures" and "leaving class or school premises without permission"; Level 3 are "insubordination" and "minor altercation"; Level 4 are "coercion or threats" and "altercation and/or physically aggressive behavior"; and Level 5 are "group violence" and "category I weapons possession" in which category I weapons include, for example, knives and guns.

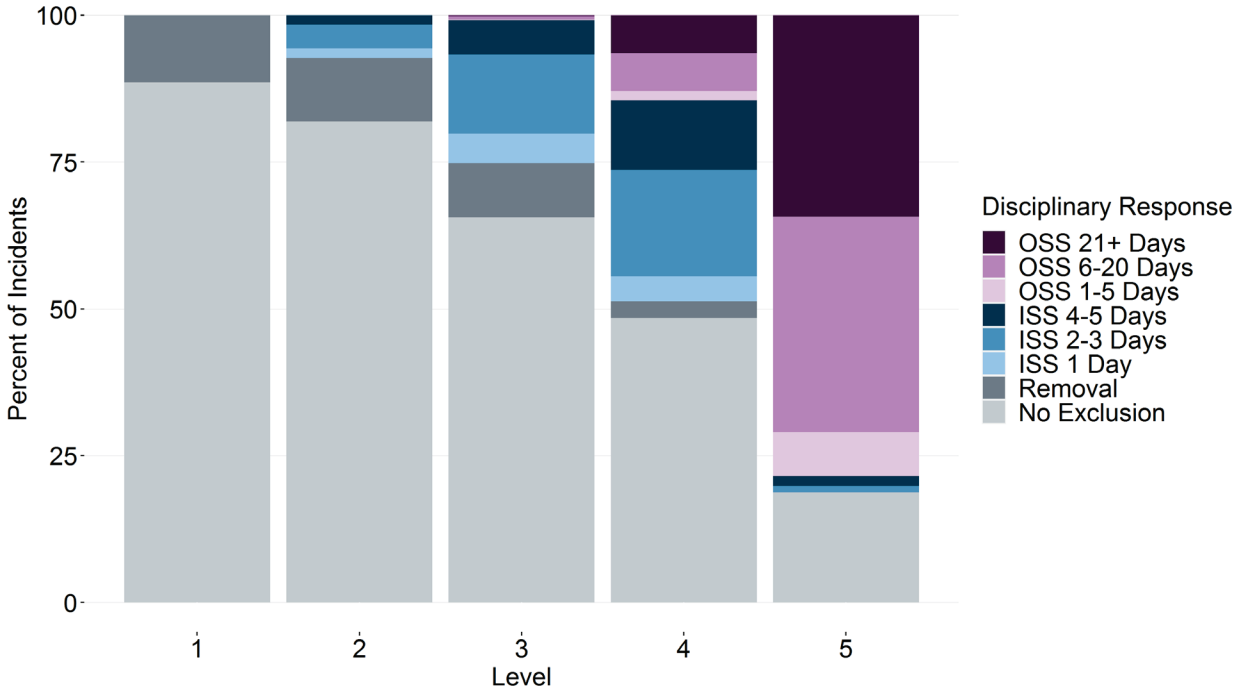
Figure 2. Incident Level by Year for High School Students



The NYCDOE discipline data file contains the type of disciplinary response and the length in days. The types of exclusionary discipline responses in the data are teacher removal (which can be for just a specific classroom period or a portion of a single day, commonly referred to as a *classroom removal*), *principal suspension* (the term NYCDOE uses for ISS), and *superintendent suspension* (the term NYCDOE uses for OSS). Not surprisingly, disciplinary responses become, on average, harsher as the level of the incident increases. Figures 3 and 4 show the distribution of disciplinary responses by level of the incident (see Appendix A, Tables A.3 and A.4, for full descriptive data). Approximately 90% of Level 1 incidents resulted in no exclusionary disciplinary action, with the remainder resulting in classroom removal. For the most severe incidents, more than three fourths of the incidents resulted in OSS, with more than one fourth of OSS being 21+ days.

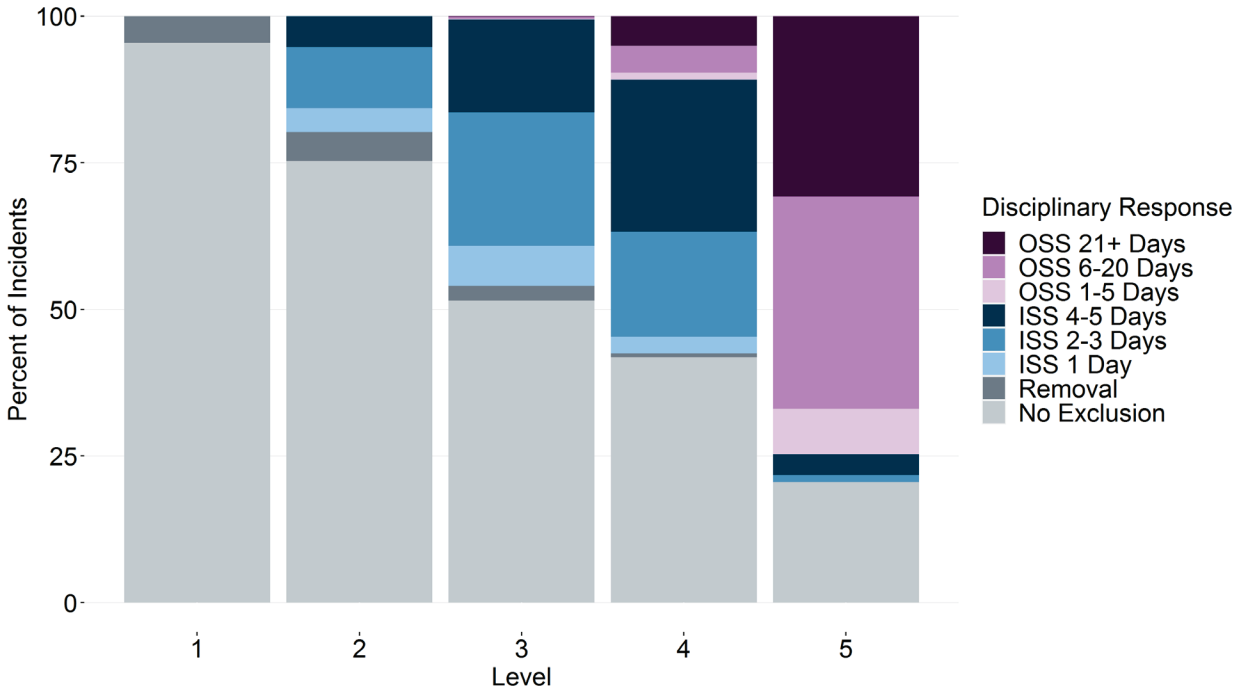
For propensity score–weighting models, the length of discipline is included as length in days. For outcomes, disciplinary response lengths were separated into bins based on frequency and district policy. For example, the common lengths of OSS were 5, 6, 10, and 30 days. In addition, NYCDOE has enacted policy around limiting the use of suspensions that are 20+ days long, making this a policy-relevant category.

Figure 3. Disciplinary Response by Level of Infraction for Middle School Students



Note. ISS = in-school suspension; OSS = out-of-school suspension.

Figure 4. Disciplinary Response by Level of Infraction for High School Students



Note. ISS = in-school suspension; OSS = out-of-school suspension.

Demographic Data

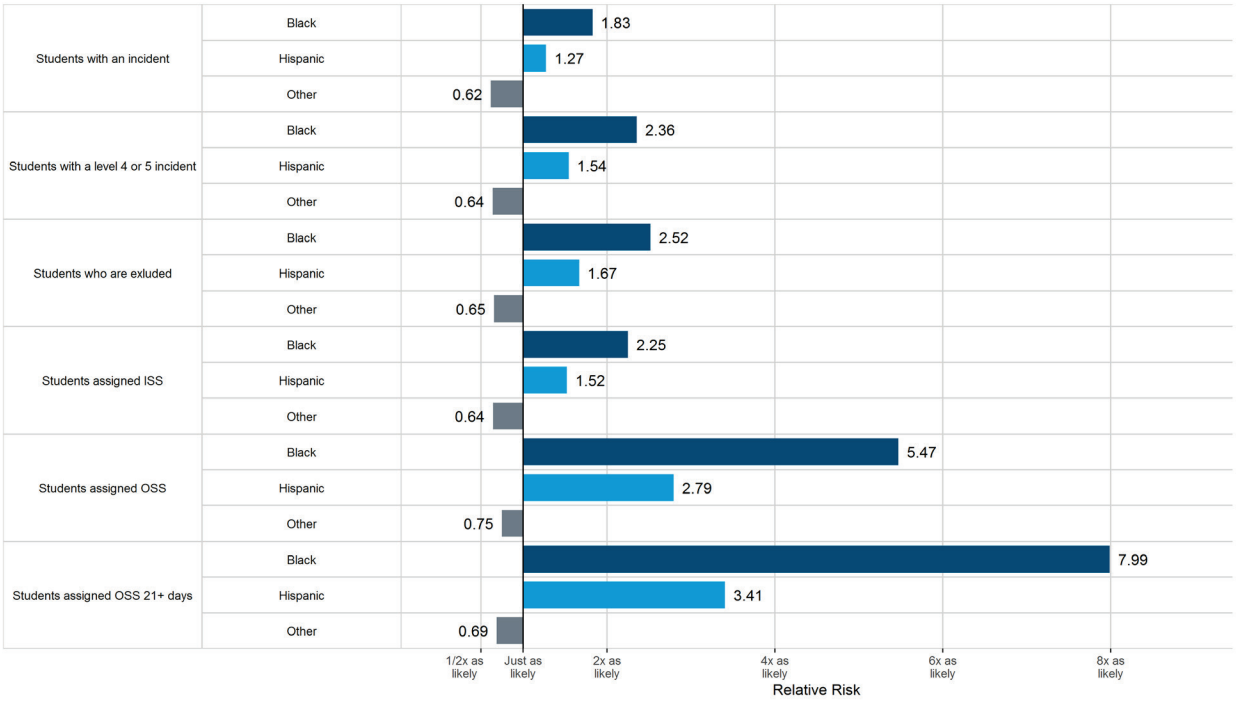
All demographic data were taken from the June demographics file each year. These data include sex, ethnicity, birth month and year, English language learner (ELL) status, SWD status, economic disadvantage, and grade level. Demographic data are used in propensity weighting models and to create subgroups of interest.

Relative risk ratios

Consistent with national data, the risk of being reported as involved in a behavioral incident is not equal across racial and ethnic groups in the NYCDOE data. For example, Black students are overrepresented among students with reported incidents relative to all students; although they comprise approximately 30% of the student body, they make up half of students with reported incidents (see Appendix A, Table A.5, for full descriptive data). Additionally, as suspension type becomes more severe, the proportion of Black students grows, as Black students comprise more than 60% of OSS of 21+ days—a proportion that is twice their overall proportion in the general population.

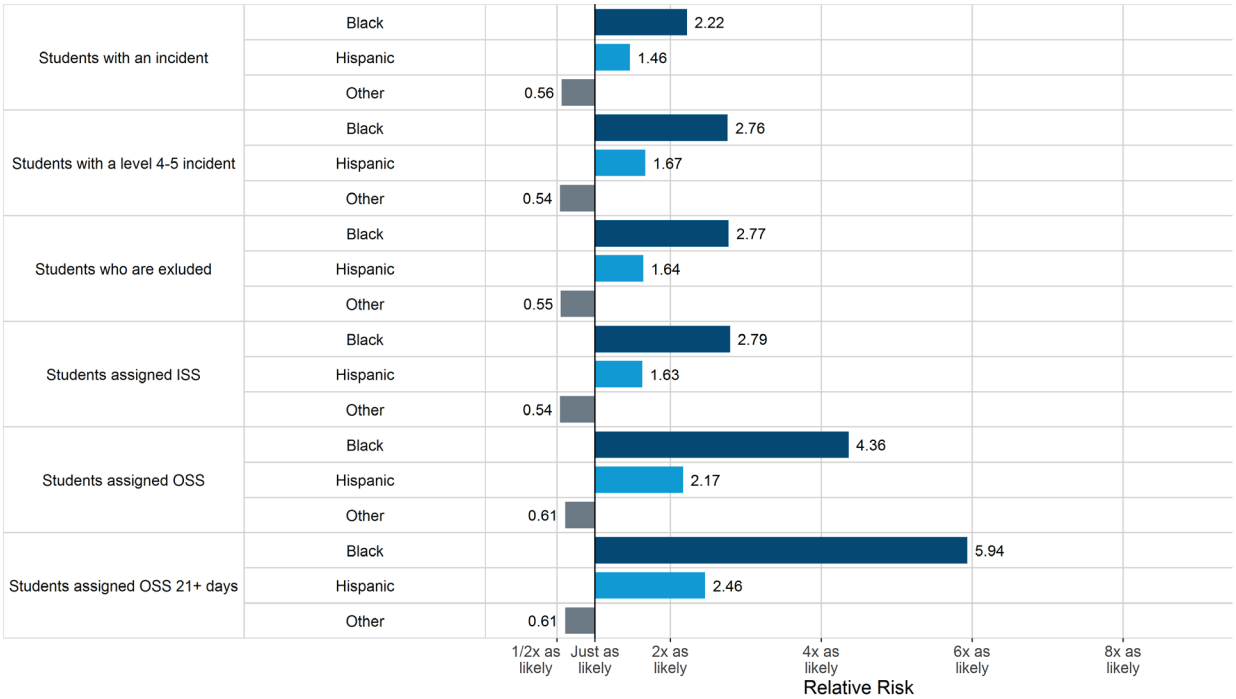
We address this by reporting relative risk ratios (RRRs), an approach that was recommended in the U.S. Department of Education's guidelines for addressing the root causes of disciplinary disparities in education (Osher et al., 2015). Figures 5 and 6 present the RRRs for Black, Hispanic, and other ethnicity middle and high school students, respectively, as compared to White students, for the following: (a) being reported for a behavioral incident; (b) being reported for a severe (Level 4 or 5) behavioral incident; (c) being given any exclusionary discipline at all (i.e., teacher's removal, ISS, or OSS); (d) being given an ISS; (e) being given an OSS; and (f) being given an OSS of 21+ days. As these graphs illustrate, the risk of being reported for an infraction and the risk of being disciplined through an exclusionary mechanism is substantially greater for Black students as compared to White students. Black middle school students have a risk of being reported for a behavioral incident nearly 2 times that of White middle school students (1.83 RRR), and for Black high school students the RRR is 2.22. The greatest differences in risk occur when comparing the risk of receiving a suspension of 21+ days. Black middle school students have a risk nearly 8 times that of White middle school students of receiving a suspension of 4+ weeks (7.99 RRR), and Black high school students have a risk that is 6 times that of White high school students (5.94 RRR).

Figure 5. Relative Risk Ratios of Discipline Incidents and Responses by Race: Middle School



Note. ISS = in-school suspension; OSS = out-of-school suspension.

Figure 6. Relative Risk Ratios of Discipline Incidents and Responses by Race: High School



Note. ISS = in-school suspension; OSS = out-of-school suspension.

Attendance and Academic Data

Annual attendance files include records of students from all district schools. We summed enrollment and absences across schools per student and then calculated the absence rate as a proportion of enrolled days. All results annualized these by multiplying them by a long-run average of 180 days per school year. NYCDOE attendance data count an absence at the middle and high school levels as being absent for at least one period during the school day. Individual student attendance data are used both in the propensity score–weighting models and as an outcome. School average attendance is calculated from individual student records and included in the propensity score–weighting model.

Assessment data were taken from the contemporary New York State assessment files for Grades 3–8 and include the ELA and math scale score. Because scores are not equated through time, we used a within-year, -subject, and -grade *z* score. Assessment data are also used both in the propensity score–weighting models and as an outcome. School average assessment scores are calculated from individual student records and included in the propensity score–weighting model.

We used the course grade file to record the grade point average (GPA) of students in Grades 9–12 from 2013–14 to 2018, as well as Grades 6–8 from 2015 to 2018, the only years the course grade files were available. GPAs were converted to grade points by ignoring plus/minus grades and weighting by credits. For each school level, only grades from that level were used as a predictor, so elementary school GPAs were not used in middle school, nor were middle school GPAs used in high school.

Because year-end assessments in ELA and math are not given to all grades in high school and because not all students take the New York State Regents exams, we instead used ELA and math credits as the academic outcomes for years when a student is in high school. Outcomes were simply measured as a 1 if the student earned at least one credit in that subject and year, and a 0 otherwise.

Graduation data reflected New York State Education Department (NYSED) rules for graduation. Students appear in this file when U.S. Department of Education graduation rate formulas would have them attributed to a school in New York. We used an on-time (4-year) graduation definition for graduation outcomes.

School Climate Data

School climate data came from the NYC School Survey, a yearly survey taken by students (Grades 6–12), teachers, parents, and select support staff and aligned to the DOE's Framework

for Great Schools.[3] Data were only available aggregated to the school level in a tabular format showing the number of respondents at every response level. These were averaged to a single score per item per year per school.

For teacher data, a fixed set of items was identified by the study team a priori as relevant to the study. These included items asking about teachers' perceptions of school safety, discipline, culture, social and emotional learning (SEL) and support, classroom environment, and greater school environment. Specific items that addressed these constructs varied by year. Two constructs were identified through a factor analysis, but there were only sufficient data available to support one. These items were largely about civility of the school and classroom environment.

For student data, again, a fixed set of items was identified by the study team a priori as relevant to the study, also including items on students' perceptions of their school's safety, discipline, culture, social and emotional learning support, classroom environment, and greater school environment. Specific items for the student survey also changed over time. An exploratory factor analysis was conducted within data collection year. This analysis consistently identified two factors in every year from among the relevant items. The two factors were persistent from 2008–09 to 2013–14, as evidenced by serial correlation. In the 2014–15 school year, however, the student survey items were updated, and although there were still two constructs, the constructs from 2008–09 to 2003–14 were not correlated with the constructs from 2014–15 to 2017–18 throughout time (within school). Because of this, we reported the two climate outcomes for the 2014–15 to 2017–18 school years only. By observing the items that loaded onto the constructs, we named them *classroom environment* and *school environment*.

Scores from the items that loaded onto each construct were averaged to determine an overall score for the construct for each school during each year. These school-level average scores were used for the propensity score model (year prior to the incident for years 2008–09 through 2017–18) and as an outcome (year of the incident for 2014–15 through 2017–18). For a full list of items included in all constructs, see Appendix A, Table A.6.

Methods

To estimate the effect of discipline type and duration on students' outcomes, same-school same-grade peers' outcomes, and school climate, we conducted a set of quasi-experimental

[3]See https://infohub.nyced.org/reports/school-quality/nyc-school-survey for more information on survey questions, psychometric properties, and response rates.

propensity score analyses. These analyses leveraged the fact that many students are included in the NYCDOE data for being involved in a particular type of behavioral incident yet received different severity of disciplinary responses. We estimated the effect of receiving a more exclusionary disciplinary response on the students who received the more exclusionary discipline (the average treatment effects on the treated [ATT]) by comparing the average outcome to what would be expected if the student had received the less exclusionary discipline in a variety of contrasts. Specifically, we compared (a) students who received any form of exclusionary discipline to those who did not; (b) students who received a 2- to 3-day ISS to those who received a 1-day ISS; (c) students who received a 4- to 5-day ISS to those who received a 2- to 3-day ISS; (d) students who received an OSS to those who received an ISS; (e) students who received a 6- to 20-day OSS to those who received a 1- to 5-day OSS; and (f) students who received an OSS of 21+ days to those who received a 6- to 20-day OSS.

To do this, we implemented a doubly robust estimator. First, we generated propensity scores using generalized gradient boosting. Generalized gradient boosting is a machine learning method that allows for automatic variable selection and can handle intricate relationships between a large number of covariates to generate a propensity score that represents the individual's propensity for receiving the "treatment"—in this case, the more exclusionary discipline response. We then used those scores to generate inverse probability of treatment weighting (IPTW), which allowed us to create two balanced groups of students: (a) those who received a harsher discipline and (b) those who received a less harsh discipline in response to being involved in an incident but who, on average, resembled the students who received the harsher discipline. Second, we used regression adjustment for selected covariates. This method is doubly robust because if either the propensity score model or the regression model is correct, the results will be unbiased.

In this section, we first explain our use of generalized gradient boosting to generate propensity scores. Next, we describe our propensity weighting models in greater detail. Finally, we detail our data analysis steps.

Generating Propensity Scores Using Generalized Gradient Boosting

Propensity scores represent the probability that an individual is assigned to the treatment group based on a set of observed covariates that are both predictive of their treatment assignment and associate with the outcome(s) of interest. In the case of this study, the treatment assignment is not a "treatment" as we usually think about it in causal research. Instead, it refers to the severity of the discipline a student received in response to a reported behavioral incident. For each analysis, the "treatment" group is the group of students who were given the more exclusionary discipline in response to their behavioral incident. The "comparison" group is the group of students who were given the less exclusionary discipline.

As a result of the rich data provided to us by NYCDOE and through the use of a gradient-boosting model, we were able to include 80 variables in our propensity score models, including 45 specific incident codes. This level of specificity is uncommon; typical propensity score models in education include a handful of characteristics of students and schools provided through administrative data. This is an important difference in that the more predictive variables included in a propensity score model, the better it can predict treatment assignment, leading to less biased effects. As such, all variables that were available and that would have theoretical basis for being associated with both the severity of the punishment a student received as well as their later outcomes were included. (A full list of variables included in the propensity score model is available in Appendix A, Table A.7.) These covariates fall into seven broad categories: (a) individual student demographic characteristics, such as race/ethnicity, gender, and economic disadvantage; (b) individual student achievement in years prior to the year of the disciplinary incident being examined, such as math and ELA standardized test scores and credit accumulation; (c) individual student attendance; (d) individual student behavioral incidents and discipline in months or years prior to the incident being examined, including the type and level of behavioral incidents, whether the student was suspended, the type of suspension, and the days of suspension; (e) school means of all student demographic, achievement, attendance, and behavioral incidents during the year prior to the incident; (f) school climate measures during the year prior to the incident; and (g) characteristics of the incident the student was involved in, including the specific incident code or codes (a student may be written up for two codes at once, e.g., smoking and unexcused absence), date of the incident, and number of other students involved.

Of all covariates in the propensity score models, the level, or severity, of the incident is the most predictive of the severity of the response. Year is also highly predictive, with the severity of responses decreasing in more recent years as NYCDOE has encouraged less punitive disciplinary responses. Also interesting is that student demographic characteristics, including race, disability status, gender, and socioeconomic background, although highly predictive of who is identified as being involved in a behavioral incident as evidenced by the RRRs presented in the prior section, were not highly predictive of receiving a more severe disciplinary response conditional on all the other variables in the model.[4] A table of all variables included in the models (Table A.7) and a table of the most predictive variables by model (Table A.8) are provided in Appendix A.

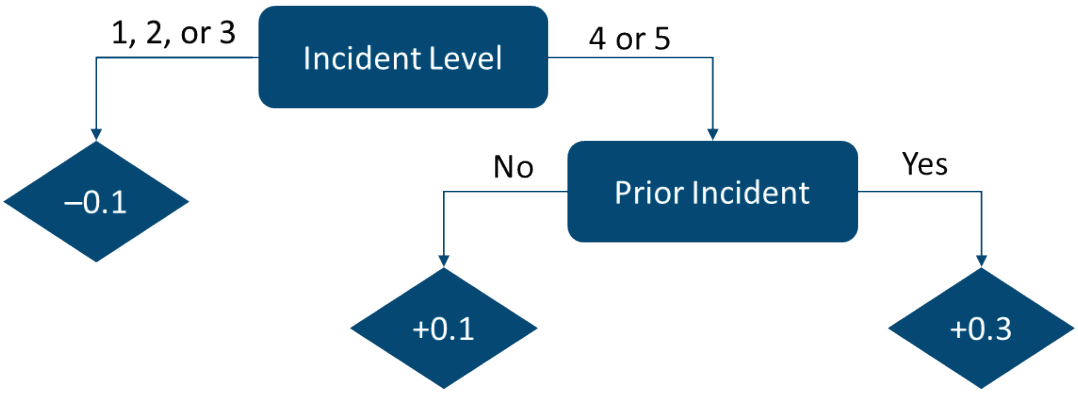
Following the approach outlined by McCaffrey et al. (2013), we implemented XGB (Chen & Guestrin, 2016), a machine learning algorithm, to estimate our propensity scores. XGB models

[4]This finding stands at odds to numerous prior findings in juvenile justice and other school contexts that suggest race affects the severity of sanctions (e.g., Osher et al., 2020).

have been shown to be useful for the estimation of propensity scores with a single treatment group (Lee et al., 2010; McCaffrey et al., 2004; Ramchand et al., 2011).

The core concept in XGB is the use of an individual classification and regression tree (CART), which fits the data by sequentially picking a variable to split the data on that maximizes the subsequent fit and then generates updates to the estimates in each node of the split. For example, if we focus on the dose of ISS, Figure 7 represents a simplified binary single CART example. In this example, the first node splits the data into incidents categorized as Levels 1–3 (left) and incidents categorized as Levels 4 and 5 (right). It then updates the probability of being given a longer or shorter ISS in each node. The left node (Level 1, 2, or 3 incidents) is a leaf, and students in this leaf have .1 subtracted from the overall probability of receiving the more exclusionary discipline response. In the right node (Level 4 or 5 incidents), the tree further splits depending on whether the student did or did not have a prior incident, with each result leading to another leaf node. The figure shows an update in propensity for those who are reported for a Level 4 or 5 incident but who had no prior incidents of a .1 increase, whereas those who did have a prior incident have their propensity for receiving the more exclusionary response increased by .3.

Figure 7. Simplified Example of a Single Classification and Regression Tree Classifying Students' Incidents, First by Level and Then by Prior Incident, for Level 4 and 5 Incidents



This CART represents a single tree. However, the algorithm is a *slow learner* in that it builds a series of trees, each of which increases the quality of the fit, conditional on the previous trees, without attempting to completely fit the data.[5] It does this by attempting to increase the

[5]Several additional protections against overfitting the data are added. First, the likelihood is not directly fit, but penalty terms are added for (a) the size of the change that is added at each leaf node, (b) the number of nodes, (c) the product of the gradient of the likelihood and the update, and (d) the product of the Hessian and the square of the update. Second, the updates are not applied directly, but instead only a portion, termed *eta,* is added from each tree. So, in Figure 1, students in Level 1, 2, or 3 would not have 1 subtracted, but eta times 1 would be subtracted from their fit.

likelihood using the summed predictions from all previous trees. Then the final prediction uses the input from each tree to form a final propensity score.[6]

In addition, to slow the speed of learning and not overfit the data, each individual tree does not contribute its full estimate but only a portion (η) to the model. The final propensity score is the sum of the outcomes from each tree [discussed above, annotated as $f_j(\cdot)$ in the next equation], conditional on covariates (X_i), mapped to the propensity score space with a function $[g(\cdot)]$.[7]

$$\Pr(T = 1|X_i) = g\left[\sum_{j=1}^{K} \eta f_j(X_i)\right]$$

The advantage of the McCaffrey method is that it works for a multinomial response, such as the differing length of suspension, which we treat as polytomous decisions, with three possible levels. Propensity score weighting also allows ready variance estimation using the jackknife, described in more depth in the following section.

Calculating IPTW

Once propensity scores are calculated, we employ IPTW to attain a group of students who received the less exclusionary response but who closely resemble the group of students who received the more exclusionary response. IPTW achieves balance between the "treatment" and "comparison" groups by upweighting individuals in the comparison group with a high probability of receiving treatment and downweighting individuals with a low probability of receiving treatment. For example, if we want to estimate the effect of receiving an OSS instead of an ISS, we need to generate propensity weights for the "comparison" students (those who received an ISS), which, when applied, make this group of students equivalent to the "treatment" students (those who received an OSS). If, for example, students who are involved in a fight with a weapon are very likely to receive an OSS, the individuals who were involved in a fight with a weapon but who received an ISS would be upweighted in the comparison group. However, students who received an ISS for a minor altercation, an incident that rarely results in an OSS, would be downweighted in the comparison group. Because we are calculating the ATT, all students in the "treatment" group (those who received the more severe discipline) are consistently weighted at 1. Table 2 provides an example of the weights associated with different probabilities of treatment for students in the comparison and treatment groups.

[6]This explanation is an oversimplified model that does not deal directly with probabilities but instead uses a latent variable. In addition, although this figure simplifies the outcome to be binary, all the decisions described in this paper are multinomial.
[7]This model described has several parameters to choose, such as η. We use 5-fold cross-validation to set these hyperparameters: η, the maximum depth of a tree, and number of trees to use in the model. In every case, the holdout data were fit nearly as well, regardless of the hyperparameters, once enough trees were used.

Table 2. Propensity Weights, by Treatment Probability

Probability of treatment	Comparison group weight	Treatment group weight
.10	0.1111	1
.20	0.25	1
.30	0.4286	1
.40	0.6667	1
.50	1	1
.60	1.5	1
.70	2.333	1
.80	4	1
.90	9	1

Using the weights, it is possible to calculate the balance of covariates across treatment and comparison groups on all of the covariates in the model. When we did this, any covariates with mean standardized difference between treatment groups greater than 0.05 SD were included in the regression model as a predictor. This doubly robust propensity score weighting and modeling approach should mitigate any residual bias due to covariate imbalance (McCaffrey et al., 2013).

To examine the effect of any exclusionary response as opposed to no exclusionary response, the data set is the full set of students with a recorded incident. To compare ISS to OSS, the analytic sample is limited to students who received one of the two types of suspension. For the question of dosage within ISS and OSS, students were limited to those who received an ISS or an OSS, respectively. What this means is that when we are comparing groups of students who, for example, received an ISS and an OSS, we are comparing groups of students who are balanced on all the variables included in the propensity score model. The groups not only resemble each other in terms of the most common covariates, such as demographics and academics, they also are comparable in terms of their prior behavior and the discipline they received, they are in schools that are comparable in terms of the types of behavior that occur and reports of school climate, and they were involved in comparable incidents leading to the punishment—not just based on a few large categories of offenses, but instead on the specific type of offense and where and when it happened.

Benefits of gradient-boosting models in the estimation of propensity scores

There are two key benefits of using a gradient-boosting model in the estimation of propensity scores in comparison to logistic regression approaches. First, gradient-boosting models allow for automatic variable selection and can handle intricate relationships between the covariates and treatment, including nonlinear relationships. Standard logistic regression would require the

analyst to determine which variables are important for the propensity score model and whether any interactions among variables or polynomial terms would improve covariate balance. With many independent variables (and possible combinations), this process becomes quickly tenuous when handled manually. Gradient-boosting model algorithms handle these decisions automatically. Instead of creating a comparison group that is balanced with the treatment group on a handful of demographic characteristics, such as race/ethnicity, economic disadvantage, and disability, we are able to create a comparison group comprising individuals who received the next less severe disciplinary response but that is balanced with the group who received the more severe disciplinary response on more than 100 variables, including, most significantly, the specifics of the behavioral incidents in which students were involved.

The second key benefit of using a gradient-boosting model in the estimation of propensity scores in comparison to logistic regression approaches is that logistic regression can often produce propensity scores that are very near to 0 or 1, which can result in large IPTW. This yields instability in impact estimates due to a small number of students potentially contributing a large majority of information to the impact estimate relative to the vast majority of the analytic sample. Calculating propensity scores via machine learning methods such as the gradient-boosting approach can help mitigate this problem.

Data Analysis Methods

This section provides details about the analytic approaches to estimating the propensity scores, checking for balance and overlap, and estimating treatment effects.

Step 1: Calculate propensity scores via generalized boosted models.

Using the XGBoost package in *R* (Chen & Guestrin, 2016), we estimated propensity scores via a boosted model while implementing *k*-fold cross-validation to facilitate model evaluation and avoid overfitting the model to the data. Treatment group membership was predicted using pretreatment covariates at the student and school levels.

Following the procedures outlined by McCaffrey et al. (2013), we estimated propensity scores for binary comparisons of treatment groups. These propensity scores were then used to calculate the appropriate weights for the ATTs for each treatment. ATTs provide an estimate of the treatment effect among the group receiving treatment had they received the comparison condition in question (e.g., the impact of receiving an OSS instead of an ISS for the students who received an OSS).

- For the effect of any exclusion, we estimated the propensity of being suspended or removed in any form (vs. not suspended or removed) among students who were involved in a behavioral incident. We calculated similar effects for OSS versus ISS.

- For the effect of suspension length (dosage), we estimated all pairwise propensity scores for the different suspension lengths considered; that is, for each suspension dosage (number of days or range of days), we calculated the propensity to receive that length of suspension for students receiving each alternative treatment (the other dosage levels).

Step 2: Select a propensity score model based on diagnostics.

Gradient-boosting models are implemented with some criteria for determining a sufficient number of regression trees to include in the final model (to avoid overfitting the data). Boosting models work in a forward stepwise process, adding weak learning models to existing weak learning models until some stopping criteria are met. We used *k*-fold cross-validation to identify the ideal number of maximum tree depth as well as the number of trees to fit. We did this by minimizing the negative log likelihood (binary models) in the test data or the exact matching error (multinomial models). After the model was fit, weights were checked. No weights were larger than 100, an indication of a model with overly large variance.

For a binary effect, the absolute standardized mean difference (ASMD) for an ATT estimate is calculated using the Cox index.

When the value was continuous, we calculated ASMD as follows:

$$ASMD_{ATTk} = |\bar{X}_{kT} - \bar{X}_{kC}|/\hat{\sigma}_{kT}$$

where $\bar{X}_{kT}$ is the weighted mean of the covariate *k* for the treatment group and $\bar{X}_{kC}$ is the corresponding weighted mean for the comparison group. $\hat{\sigma}_k$ is the SD of the covariate in the pooled sample—although for *z*-scored variables, such as assessment scores, $\hat{\sigma}_{kT}$ was always set to the population value of 1.

Step 3: Assess balance and overlap.

After a model was selected in Step 2, we assessed covariate balance and treatment group overlap. The balance statistics outlined previously (Cox index, ASMD) were used. When assessing balance after model selection, we examined these balance statistics for each measured pretreatment covariate. Standardized mean differences less than 0.05 are considered ideal. Covariates with greater than 0.05 difference (and less than 0.25 SD difference) are included in the regression models described in Step 4. When a covariate had an imbalance greater than 0.25 SD, one of a variety of changes was applied. If the predictor was discrete and rare (e.g., NYPD contacted), it was either noted as a contrast or, if very large, all cases in the rare group were excluded from the analysis (only one instance). If the covariate was imbalanced (>0.25 SD) and not rare, the results are not reported; this happened only for contrasts between removal and ISS.

Table A.9 in Appendix A presents the covariates that had large imbalances for each model and the decision that was made to resolve the imbalance.

Propensity score weighting substantially improved covariate balance. Tables A.10–A.15 in Appendix A present the balance before and after, in standardized mean differences, for the covariates. The final column of each table shows the unweighted combined mean of the variable. For a binary variable, this is the proportion in that category.

Step 4: Fit doubly robust mixed-effects regression models.

To calculate our treatment effects for binary treatments (e.g., OSS vs. ISS) and our pairwise treatment effects (e.g., 5 days suspended vs. 3–4 days), we fit a series of doubly robust regressions. Using the propensity score models developed in Steps 1–3 for each comparison, we created IPTWs. To make the regression doubly robust, the model included covariates with residual imbalance after weighting (i.e., covariate imbalance >0.05 SD).

The regression model used to estimate the impact of disciplinary actions on the outcomes can be represented as follows:

$$Y_{ij} = \beta_0 + \beta_1 Trt_{ij} + \sum_{p=2}^{P+2} \beta_p X_{pij} + \varepsilon_{ij}$$

where Y_{ij} is the outcome of interest, Trt_{ij} is a binary treatment indicator (either treatment vs. comparison or Treatment A vs. Treatment B), X_{pij} are p covariates with residual imbalance, and ε_{ij} is the residual error term. Binary outcomes were fit with this same model.

All models were fit with robust standard errors using a jackknife procedure that dropped clusters of schools in 62 groups one at a time. The results were subtracted from the full sample estimates and the square deviations summed according to Wolter (2007).

$$s_{\beta_1} = \frac{61}{62} \sum_{i}^{62} \left(\beta_{1,0} - \beta_{1,i}\right)^2$$

where s_{β_1} is the standard error of the treatment effect, $\beta_{1,0}$ is the full sample estimate of the treatment effect, and $\beta_{1,i}$ is the estimate that sets all weights in cluster *i*'s schools to 0. This yields a school-level cluster-robust standard error.

Subgroups

Steps 2–4 were redone separately for five subgroups of students: (a) Black students, (b) Hispanic students, (c) White students, (d) students identified with the NYC "poverty" flag, and (e) students identified as "students with disabilities" (SWD) by the district. The

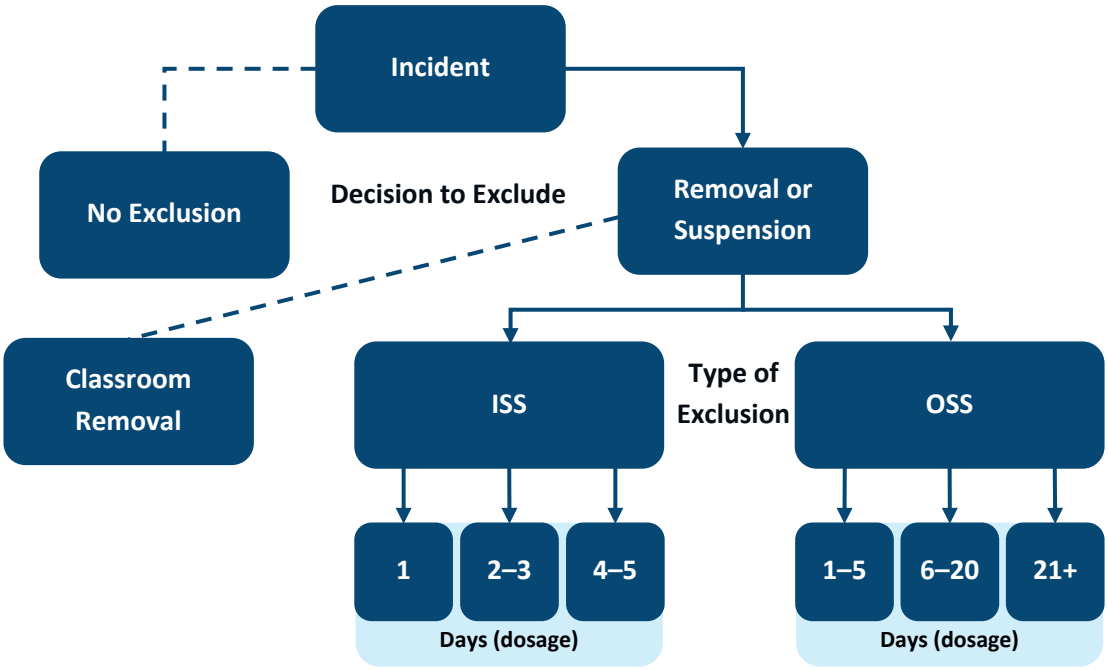
hyperparameter tuning from the full data set was used, but the mating model was refit for each subgroup. Following the recommendations of McCaffrey et al. (2013), when some subgroups had weights more than 100, they were trimmed to 100 to avoid high variance estimators. Then the balance was checked and estimates for the subgroup were generated.

The model was not reported for several subgroups in high school and middle school because it was not possible to achieve balance within the subgroup of students (in part due to the much smaller number of individuals in each subgroup). For this reason, no subgroup effects are reported for Asian students, students of other ethnic groups, or for intersectional categories. In addition, in the model for the dose of OSS, one covariate, percent ELL, was not balanced for the model on Black students. However, this category made up only a few percent of the Black students in the sample, so the results were reported.

Results

To answer the research questions and provide evidence of the effect of the severity of the type and length of exclusionary discipline responses on students and their peers, we attempted to examine a total of seven comparisons (see Figure 8): (a) any removal or suspension versus no exclusion; (b) OSS versus ISS; (c) ISS versus classroom removal; (d) ISS of 2 or 3 days versus ISS of 1 day; (e) ISS of 4 or 5 days versus ISS of 2 or 3 days; (f) OSS of 6–20 days versus OSS of 1–5 days; and (g) ISS of 21+ days versus ISS of 6–20 days. There was insufficient overlap to allow comparisons between students who received a classroom removal and those who received an ISS in response to an incident. As such, we are not able to report the associated effects with this decision. Additionally, for the comparison between a middle school OSS of 6–20 days versus 21+ days, the balance between incident level was just shy of the 0.25 SD cutoff (.27). We decided to report these results but urge the reader to interpret these comparisons with more caution.

Figure 8. Discipline Decision Comparisons



Note. ISS = in-school suspension; OSS = out-of-school suspension.

In this section, we provide results for the effect of each increase in discipline severity—first on the educational outcomes of disciplined students themselves, then on their same-school same-grade peers' educational outcomes, and finally, on teachers' and students' measures of school climate. We ran all analyses separately for the full group of middle and high school students as well as for middle school Black students, Hispanic students, White students, SWD, and students from economically disadvantaged backgrounds and for high school students of the same subgroups. Effects were fairly consistent across populations. As such, figures in the main body of the report are for the full group of students.

Effects on Students' Educational Outcomes

The effects of the type and length of discipline on students' educational outcomes were examined for 1–3 years after the students were involved in a behavioral incident, as present in the data, for high school students, and 1–4 years after the students were involved in a behavioral incident for middle school students. For example, if a student was involved in an incident in 9th grade and progressed on time through grades without dropping out, the student's behavior, attendance, and credit accumulation would be available for 3 years (if the incident happened 2014–15 or before, as the data are only available through 2017–18). However, a student who was involved in an incident in 11th grade would only be able to be

examined for 1 year (their 12th-grade year) if they graduated on time, as would a 10th grader who drops out of school after 11th grade.

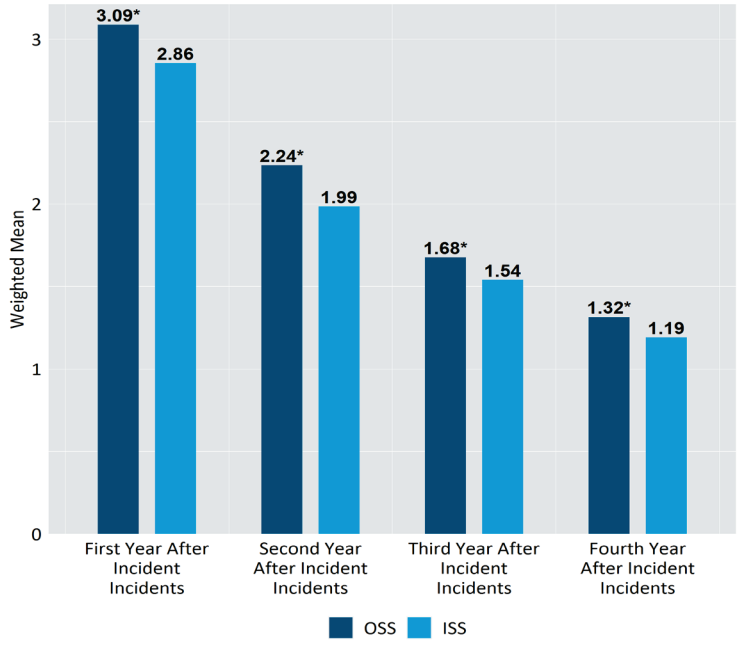
Behavior and Discipline

In the following sections, we present findings of the effects of more severe exclusionary discipline on later behavioral incidents reported and number and length of suspensions. Results are presented first for middle school students and then for high school students.

Middle School Students' Behavioral Incident Reports.

We found no evidence that more exclusionary disciplinary responses reduced future reported behavioral incidents for disciplined students. Rather, receiving an OSS instead of an ISS, in particular, actually had a negative effect on middle school students' future behavioral incidents, both overall and for more severe Level 4 and 5 incidents. These results suggest that longer and more severe forms of exclusionary discipline do not serve as a deterrent and may result in more behavioral incidents, either because students' behavior is negatively impacted by the experience or because their future behavior is viewed more negatively by teachers and school administrators (see Appendix A, Table A.14, for full results). For example, middle school students who received an OSS instead of an ISS were reported for more behavioral incidents in school during each of the 4 years after their incident, as shown in Figure 9.

Figure 9. Effect of Receiving an OSS Instead of an ISS on Number of Reported Behavioral Incidents 1–4 Years Later: Middle School

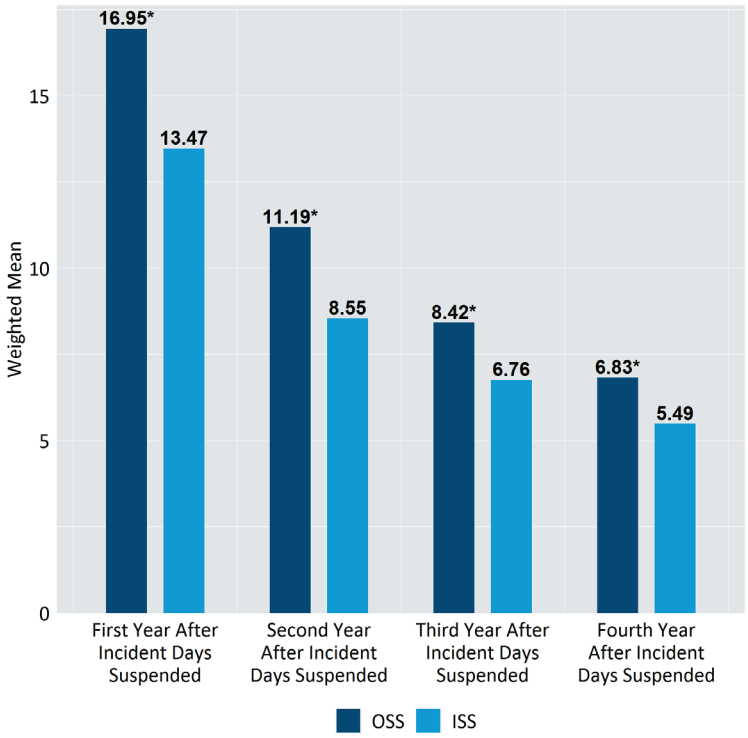


Note. ISS = in-school suspension; OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

Middle School Students' Suspensions.

Middle school students who received a more severe exclusionary discipline response were also more likely to be suspended in the future and to miss more days due to suspension in the future, particularly if they received an OSS instead of an ISS or if they received an OSS of 21 days or more. For example, as shown in Figure 10, students who were given an OSS instead of an ISS in response to their behavioral incident missed 3.5 more days due to suspension the year after the incident, and even 4 years later were suspended for more than a day, on average, than they would have been had they received an ISS for the incident (see Appendix A, Table A.16, for full results).

Figure 10. Effect of Receiving an OSS Instead of an ISS on Days Suspended 1–4 Years Later: Middle School



Note. ISS = in-school suspension; OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

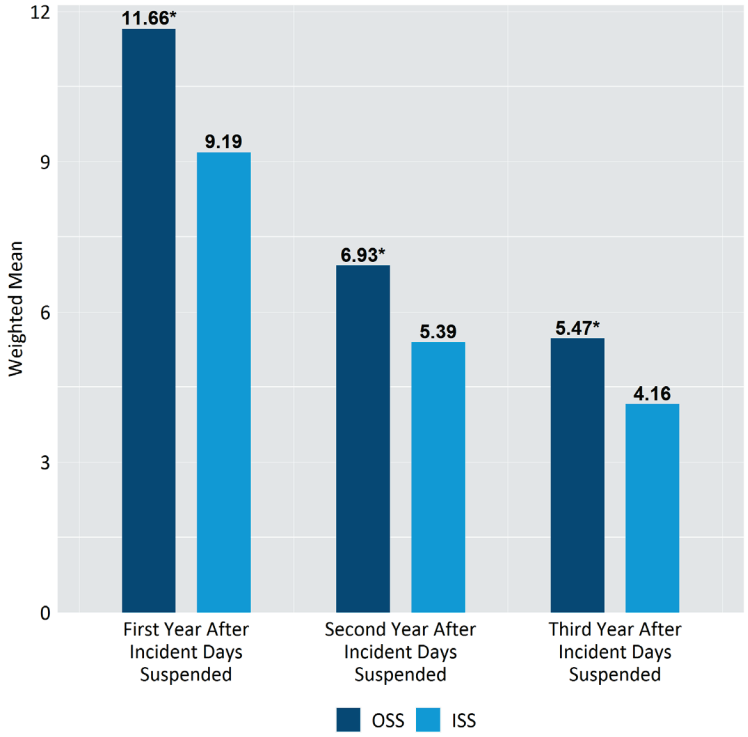
High School Students' Behavioral Incident Reports.

There is no effect of discipline severity on high school students' later reported behavioral incidents or reported severe (Level 4 or 5) behavioral incidents (see Appendix A, Table A.15, for full results). However, it also does not serve as a deterrent for future misbehavior. There is no consistent effect of discipline severity on the number of incidents resulting in a suspension.

High School Students' Suspensions.

There is evidence that students who received a more severe exclusionary punishment in response to an incident, despite not being involved in any more incidents in the future or in any more severe incidents in the future, did spend significantly more days suspended in the following 3 years (see Appendix A, Table A.17, for full results). For example, students who received an OSS rather than an ISS in response to a behavioral incident spent an additional 2.5, 1.5, and 1.3 days suspended 1, 2, and 3 years after the incident, as shown in Figure 11. Again, this suggests that rather than serving as a lesson or deterrent, more severely punishing a student does nothing to reduce the number of behavioral incidents they are involved in later and instead results in continued severity in future punishment.

Figure 11. Effect of Receiving an OSS Instead of an ISS on Days Suspended 1–3 Years Later: High School



Note. ISS = in-school suspension; OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

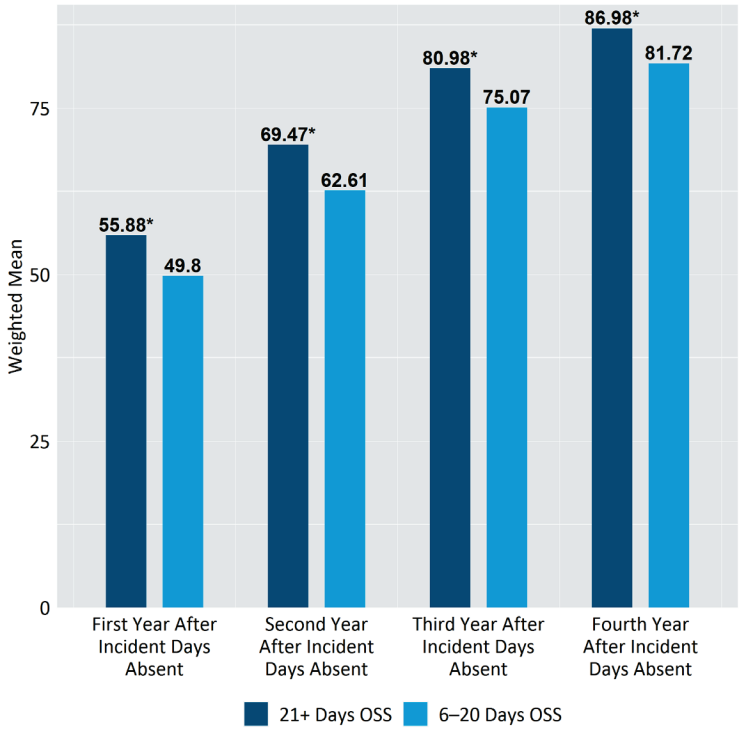
Attendance

In this section, we present findings on the effect of exclusionary discipline on students' (not suspension related) attendance. NYCDOE data consider a day absent if the student missed one or more classes throughout the day. Results are presented for middle school students, followed by high school students.

Middle School Students.

More severe exclusionary discipline has a consistent negative effect on middle school students' attendance 1–4 years after students are involved in a behavioral incident (see Appendix A, Table A.16, for full results). The largest effects were for students who received an OSS instead of an ISS (between approximately 5 and 6 additional days of absence 1–4 years later) and for students who received a suspension of 21+ days instead of a 6- to 20-day suspension (between approximately 5 and 7 additional days of absence 1–4 years later), as shown in Figure 12. However, even small changes in punishment severity, such as receiving a 2- or 3-day ISS rather than a 1-day ISS, negatively affected students' attendance for each of the following 4 years.

Figure 12. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on Days Absent 1–4 Years Later: Middle School



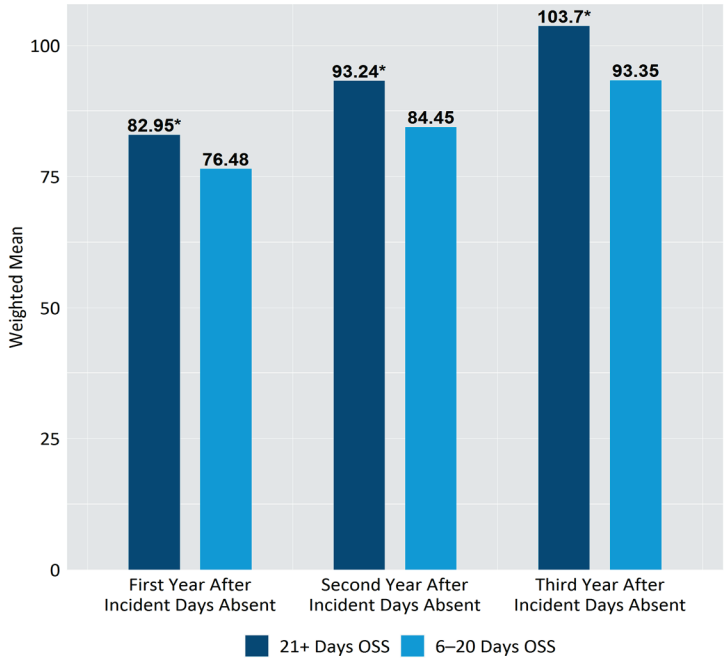
Note. Days absent comprise all days with an absence recorded for one or more classes. OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

High School Students.

More severe exclusionary discipline also has a consistent negative effect on high school students' attendance 1, 2, and 3 years after students are involved in a behavioral incident (see Appendix A, Table A.21, for full results). For all comparisons, with the exception of a 6- to 20-day OSS versus a 1- to 5-day OSS, students receiving the more severe exclusionary discipline missed significantly more school due to absence 1, 2, and 3 years after the incident. For

example, students who received an OSS in response to an incident missed at least one class between 6.1 and 8.4 more days of the school year each year over the following 3 years than they would have had they been given an ISS. To put this in perspective, a 9th grader who receives an OSS rather than an ISS in response to a behavioral incident would be absent from one or more classes, on average, over 4 more weeks of school (21.4 days) than if they had received an ISS. Furthermore, a suspension of 21+ days resulted in between 1 and 2 additional weeks of absences each of the following 3 years compared to what would be expected if those students had received a 6- to 20-day suspension, for a total of 5 additional weeks of missing at least one class, as shown in Figure 13.

Figure 13. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on Days Absent 1–3 Years Later: High School



Note. OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

Academic Achievement

In this section, we present the effect of exclusionary discipline severity on standardized test scores in middle school and credit accumulation in high school. Standardized test scores are presented for the years when a student was still in middle school (1- and 2-year effects for a 6th grader with a behavioral incident and 1-year effects for a 7th grader with a behavioral incident). High school credit accumulation is presented for the years when a student was in high school (3- and 4-year effects for a 6th grader with a behavioral incident; 2-, 3-, and 4-year effects for a 7th grader with a behavioral incident; and all years for students in Grades 8–12 with a behavioral incident).

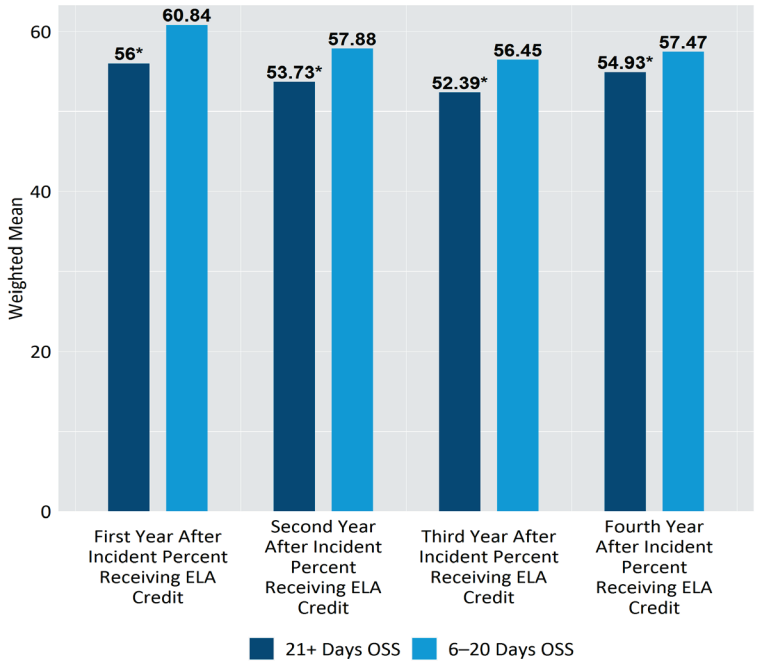
Middle School Students' Standardized Test Scores.

There is evidence that more severe exclusionary discipline may have a small negative effect on students' ELA and math standardized test scores in middle school, although effects were not statistically significant for most comparisons (see Appendix A, Table A.22, for full results). For students who were in 7th or 8th grade 1 and 2 years after a behavioral incident, we found that more severe exclusionary discipline had a small negative effect on standardized test scores (measured in *z* scores) for OSS compared to ISS.

Middle School Students' High School Credit Accumulation.

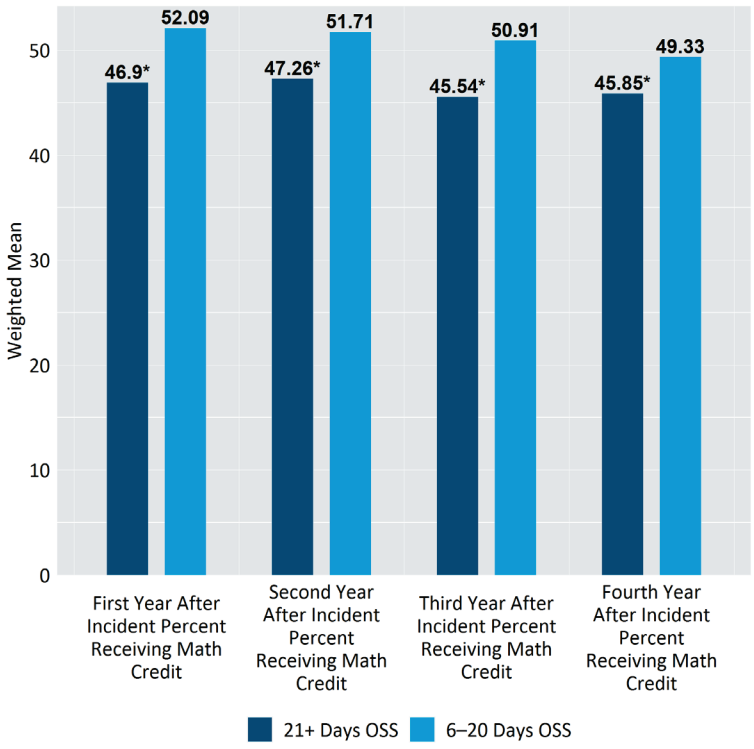
Consistent negative effects of more severe exclusionary discipline were found for high school credit accumulation in ELA and math (e.g., if a student was in 7th grade during an incident, their 1-year effect would be standardized test scores because they were in 8th grade, but their 2- through 4-year effects would be credit accumulation in Grades 9–11). Again, OSS instead of ISS and OSS of 21+ days instead of 6–20 days have the largest and most consistent negative effects on students' credit accumulation. For example, students who received a suspension of 21+ days earned both ELA and math credits in high school at between approximately 2.5 and 5.5 percentage point lower rates 1–4 years later than they would have had they been given a suspension that was 6–20 days in length, as shown in Figures 14 and 15 (see Appendix A, Table A.23, for full results).

Figure 14. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on the Probability of Earning an ELA Credit 1–4 Years Later: Middle School



Note. ELA = English language arts; OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

Figure 15. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on the Probability of Earning a Math Credit 1–4 Years Later: Middle School

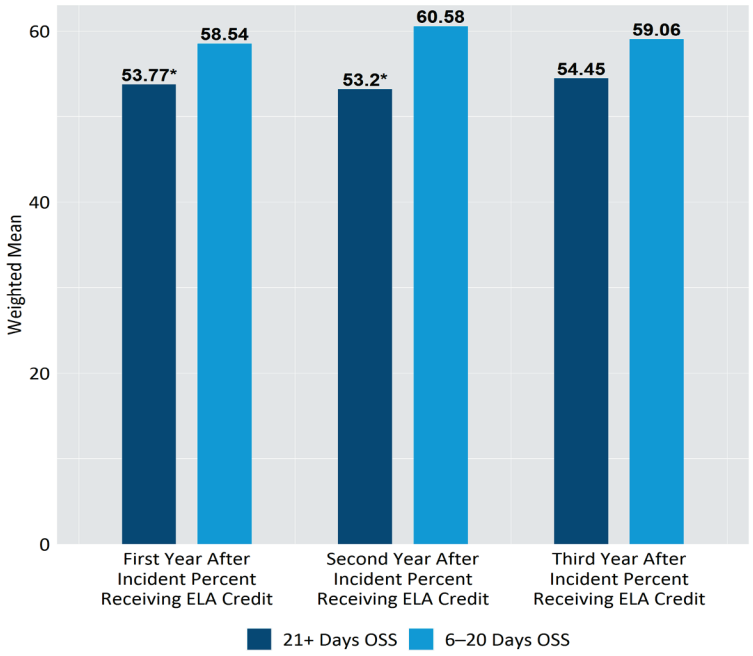


Note. OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

High School Students' Credit Accumulation.

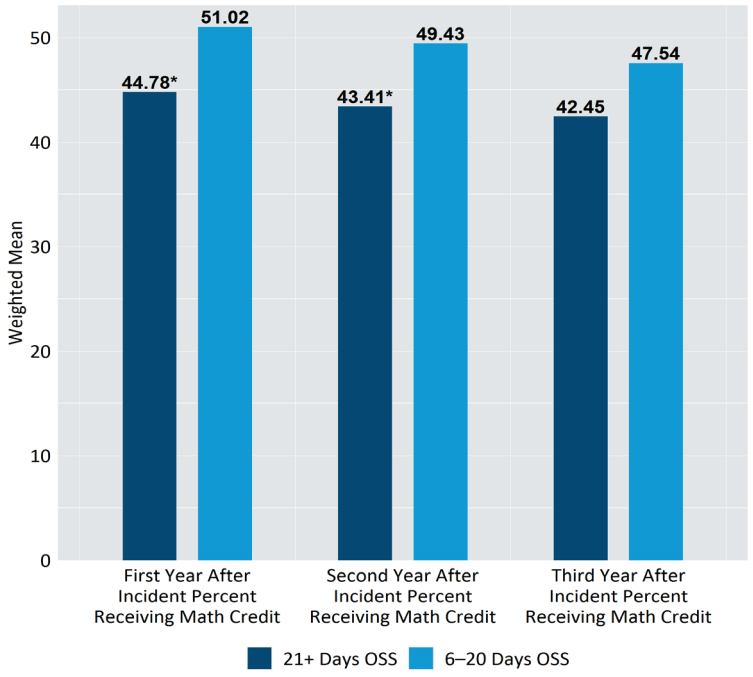
More severe exclusionary discipline had a consistent negative effect on high school students' ELA and math credit accumulation as well (see Appendix A, Table A.24, for full results). The largest effects were seen for students who were given an OSS of 21+ days. These students were between approximately 4.5 and 7 percentage points less likely to earn a credit in ELA and math each of the 3 following years, as presented in Figures 16 and 17, than if they had received an OSS of 6–20 days. However, even small changes in the length of a student's ISS also negatively affected high school credit accumulation in the 2 years after the incident, although to a smaller degree.

Figure 16. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on the Probability of Earning an ELA Credit 1–3 Years Later: High School



Note. ELA = English language arts; OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

Figure 17. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on the Probability of Earning a Math Credit 1–3 Years Later: High School



Note. OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

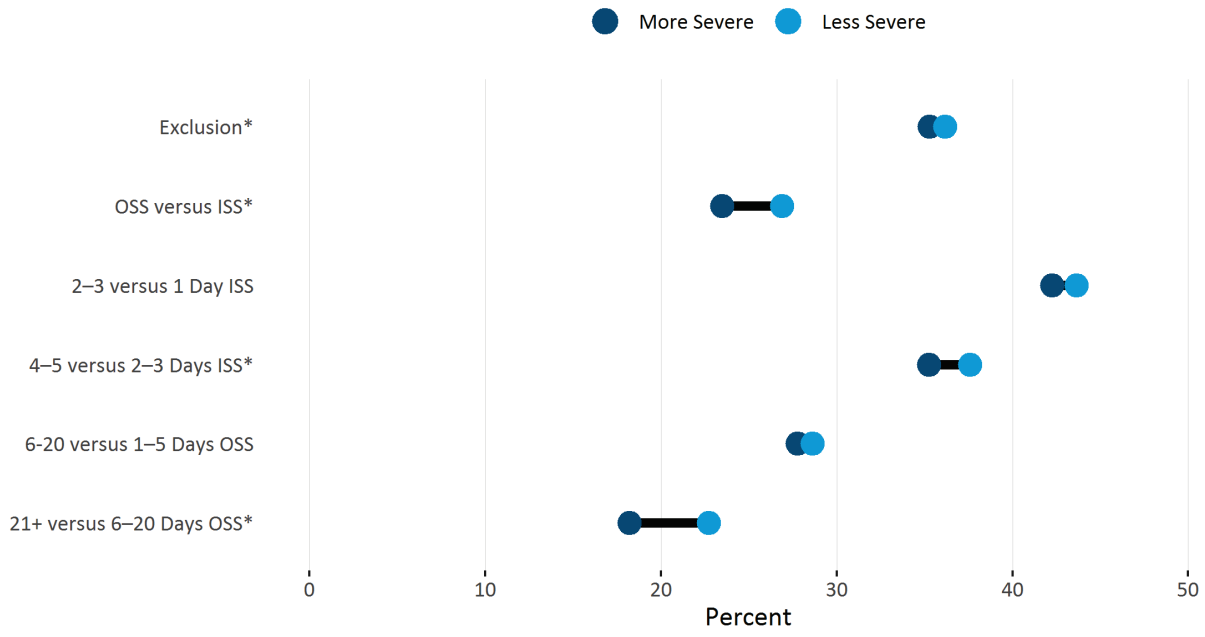
On-Time Graduation

In this section, we present the effects of exclusionary discipline severity on the likelihood that a student graduates on time (4 years after beginning 9th grade) from high school. First, we present results of behavioral incidents for students in middle school and then for students in high school.

Middle School Students.

More severe exclusionary discipline, even as early as in middle school, has a consistent and substantive negative effect on the likelihood that a student will graduate on time from high school (see Appendix A, Table A.25, for full results). For four of the six severity comparisons, students who received the more severe punishment rather than the less severe were statistically significantly less likely to graduate from high school on time (the other two comparisons were also negative in direction but not statistically significantly different from 0). For example, middle school students who received an OSS instead of an ISS in response to a behavioral incident were 3.4 percentage points less likely to graduate on time from high school, and middle school students who received the most severe punishment (OSS of 21+ days) were 4.5 percentage points less likely to graduate on time from high school, as shown in Figure 18.

Figure 18. Effect of Receiving Subsequently More Severe Exclusionary Discipline on the Percentage of Students Graduating on Time: Middle School

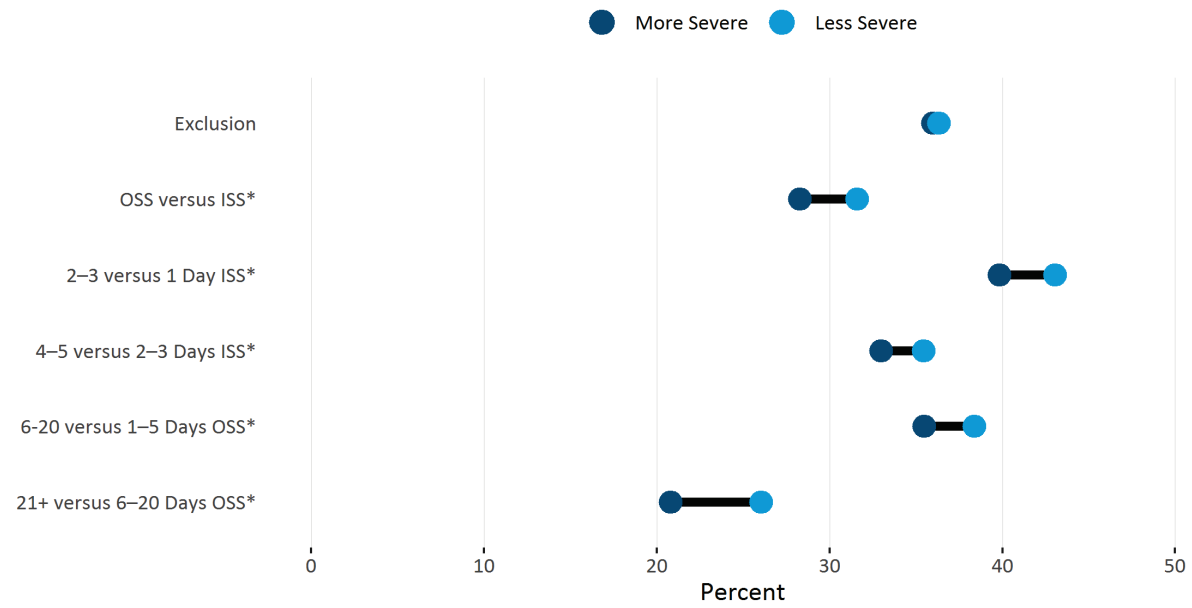


Note. ISS = in-school suspension; OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

High School Students.

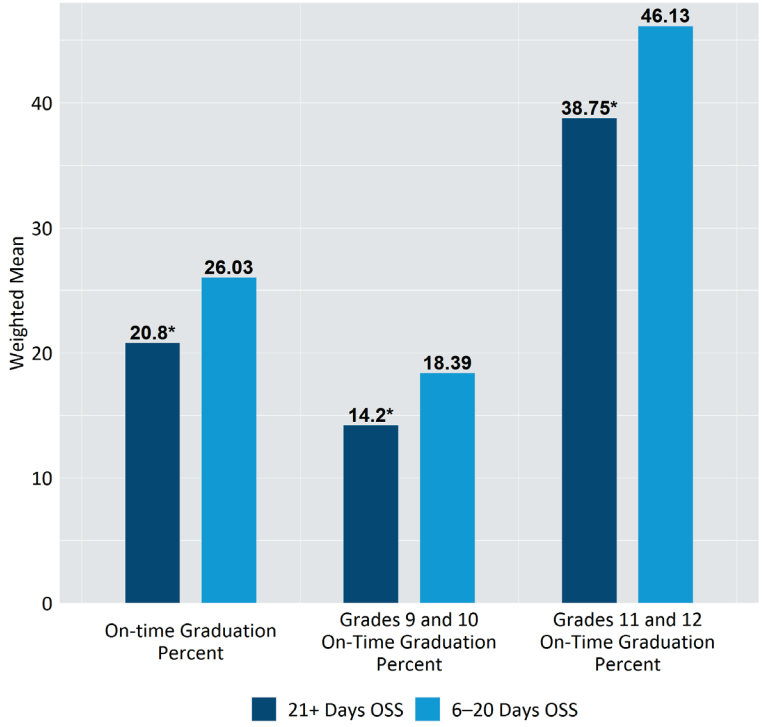
More severe exclusionary discipline also has a consistent and substantial negative effect on high school students' likelihood of graduating from high school (see Appendix A, Table A.26, for full results). Every increase in discipline severity decreases a students' likelihood of graduating by at least 2 percentage points, with the largest negative effect coming from receiving an OSS of 21+ days, as shown in Figure 17. For example, students who received a suspension of 21+ days during 11th or 12th grade graduated high school on time at a rate that was more than 7 percentage points, or nearly 20% lower, than they would have had they been given a suspension that was between 6 and 20 days, as shown in Figures 19 and 20. To put this number in perspective, for every twelve 11th- and 12th-grade students who are given a suspension of 21+ days instead of a suspension of 6–20 days, one additional student will fail to graduate, and for every twenty-four 9th- and 10th-grade students who are given a suspension of 21+ days, an additional student will fail to graduate.

Figure 19. Effect of Receiving Subsequently More Severe Exclusionary Discipline on the Percentage of Students Graduating on Time: High School



Note. ISS = in-school suspension; OSS = out-of-school suspension.

Figure 20. Effect of Receiving an OSS of 21+ Days Instead of 6–20 Days on the Probability of Graduating on Time by Grade Level: High School



Note. OSS = out-of-school suspension; * Indicates differences are statistically significant at the .05 level.

Peer Spillover Effects

We also examined the effect of the type and length of exclusionary discipline on the same-school same-grade peers of students involved in a disciplinary incident. Although we examined later outcomes of the students themselves, we examined contemporaneous outcomes of their peers. This is because one of the most common rationales for the need for exclusionary discipline is that removing disruptive or dangerous students from the classroom is necessary in order to have a classroom conducive to learning for the other students and a school where students feel safe attending. Additionally, students disperse over time, and longitudinal effects are less meaningful.

We did not find any consistent effect of the severity of discipline a student receives on the overall attendance rate, ELA or math standardized test scores, or credit accumulation of the students in the same grade within their school (see Appendix A, Tables A.27 and A.28, for middle school and high school full results, respectively). For middle school students, receiving an OSS rather than an ISS did have a small but statistically significant negative effect on peers' ELA and math standardized test scores and attendance, suggesting that, if anything, excluding students more severely might negatively affect peers.

School Climate

Finally, we examined the effect of the severity of exclusionary disciplinary response on students' and teachers' measures of school climate during the year of the incident, limiting analyses to the 2015–16 through 2018–19 school year due to consistency of data measures. Incidents which occurred during May or June were dropped from these analyses because students and teachers complete the school climate survey during these months and there incidents during these months did not necessarily precede these outcome measures.

We did not find any evidence that more exclusionary discipline improved students' or teachers' perceptions of classroom or school climate. In fact, for middle school students, receiving an OSS rather than an ISS was associated with slightly lower student and teacher reports of school climate. Other comparisons found no effect. Appendix A, Tables A.29 and A.30, present the full results for middle and high school students' and teachers' perceptions of school climate, respectively.

Student Subgroup Effects

All analyses of the effect of discipline type and severity on students' later behavioral, attendance, and academic outcomes were examined separately for subgroups of students in order to explore whether the type and length of exclusionary discipline affected different students differently. We examined Black and Hispanic middle school students separately and Black, Hispanic, and White high school students separately. Given our statistical models, there were not enough students from other ethnic subgroups at either grade level to enable a separate examination of these students, nor were there enough White middle school students. We also examined SWD and students from economically disadvantaged backgrounds. We found that there was no statistical evidence that the effects of exclusionary discipline differed for students with these characteristics or backgrounds. More exclusionary discipline had similarly null or negative effects for all students, with OSS rather than ISS and OSS of 21+ days being the most consistently harmful for all students. All student subgroup analyses can be found in Appendix B.

Discussion

In this paper, we set out to build on the current evidence by examining, using quasi-experimental machine learning methods and a decade's worth of data from NYCDOE, the effects of the type and length of exclusionary disciplinary responses on (a) high school students' educational outcomes, (b) their same-school same-grade peers' educational outcomes, and (c) teachers' and students' reports of school climate. We found that more severe exclusionary

discipline does not result in fewer future behavioral incidents, and for younger students it may instead lead to an increase in reports of negative behavior. This finding is in contradiction to deterrent theory and does not support claims that exclusionary discipline will reduce future misbehavior either by teaching students a lesson or by simply serving as a severe enough punishment to deter misbehavior.

Most importantly, more severe exclusionary discipline has a consistent negative effect on many other long-run educational outcomes for students. Receiving a more severe exclusionary disciplinary response to an incident increases the number of days students miss due to absence during subsequent school years, increases the number of days they miss due to suspension in subsequent school years, decreases their likelihood of earning both ELA and math credits throughout their high school career, and decreases their likelihood of graduating. We found that receiving an OSS rather than an ISS and receiving a suspension of 21+ days had particularly severe and consistent negative effects on students' educational outcomes.

In addition, these negative educational effects on students are not offset by any improvements to their peers' outcomes or their peers' or teachers' reports of their school's climate. The severity of exclusionary disciplinary response had no effect on same-school same-grade peers' academic outcomes or attendance, nor did it have any effect on students' or teachers' perceptions of school climate. As such, these results do not support the other common claim by proponents of exclusionary discipline—that removing misbehaving students from the classroom is necessary to ensure that their peers are able to learn and feel safe within their school.

Subgroup analyses of the data included in our study suggest that the effects of exclusionary discipline severity are similarly negative across all groups of students. However, as our risk ratio analyses indicate, Black students are overrepresented among students who receive more severe exclusionary discipline (as noted earlier, Black middle school students have a risk of receiving a suspension of 21+ days that is 8 times that of White students). These results suggest that the disparate use of exclusionary discipline by student race contributes to the racial achievement gap and the racial gap in high school graduation.

Contributions to the Literature

Our study employed a rich and robust longitudinal data set to empirically examine what cannot ethically be experimentally examined at scale: the academic and behavioral consequences of suspension in schools. Because these analyses are not experimental, we cannot rule out that there is some level of unobservable variable bias at play in these estimates. Even these advanced machine learning methods are only able to account for that which has been

observed. However, the inclusion of such detailed and specific data on the incident a student was involved in greatly reduces this bias, and as such, our estimates are an important addition to the current literature. For example, when estimating the effect of OSS versus ISS, we are comparing groups of students who are balanced not only on race, gender, economic disadvantage, and prior behavior, but also on the specific incidents they were involved in, the time of year those incidents occurred, the number of students involved in those incidents, the prior incidents the students were involved in, the type and length of discipline they received for those prior incidents, and numerous other factors. In addition, they are balanced on school-level variables, including average test scores, racial and socioeconomic makeup, school climate, and the types of behavioral incidents and disciplinary responses received among all students within the focal student's school. A total of 80 variables are included in the propensity score–weighting models, ensuring that the comparison group of students receiving the less punitive exclusionary response in each comparison is optimally comparable to the students receiving the more punitive exclusionary response in myriad ways.

Additionally, our results are generalizable in a way that many other studies are not. Although working with only one district's data, these analyses are not based on just one short period of time, nor are they examining differences in outcomes due to a change in suspension policy regarding particular types of behavioral incidents. These analyses use the universe of reported behavioral incidents over a 10-year period in the largest school district in the nation. The source of variation that allows us to observe the effect of differences in discipline severity comes from natural variation in the discipline assigned to students by school administrators, rather than being based on a discrete policy change. These natural sources of variation are implicit among disciplinary decisions in districts throughout the country. In addition, effect estimates for a specific comparison—for example, OSS compared to ISS—are based on all incidents that resulted in either an OSS or an ISS over 10 years, with the types of ISS incidents for which there is the largest likelihood of having possibly received an OSS being weighted the greatest. This means that effect estimates for this comparison represent the difference in the actual observed outcomes the students receiving an OSS had in the data, as compared to what they would have had if they had received an ISS; and this is based on the outcomes of those students who received an ISS but who, given the incident they were involved in and everything else we observed about them and their school, had a high likelihood of having received an OSS.

In addition, although the correlational literature has documented long-term differences in the outcomes of suspended students as compared to their nonsuspended peers, the causal literature has focused nearly exclusively on shorter term outcomes and, in many cases, contemporaneous outcomes (i.e., behavior and achievement during the year before a policy change for students who were suspended that year as compared to their behavior and

achievement the following year after the policy shift, or comparing academic achievement in a school quarter during which a student received a suspension as compared to a school quarter when the student did not). By relying on the variation in response to students involved in similar incidents and by employing 10 years' worth of data, we were able to observe that decisions made about whether and how severely to use exclusionary discipline as a response to a student's behavior, as early as when the student is still in middle school, have long-term deleterious effects on that student's likelihood of graduating from high school. These findings are consistent with those found by Sorensen et al. (2021), despite relying on different forms of variation to arrive at the estimates—although we used propensity weights to create optimally matched groups of students involved in the same infractions but receiving different levels of punishment across district schools, Sorensen et al. exploited principals' propensity to exclude students to different degrees and their movement between schools at the state level.

Study Limitations

This study is an important addition to the causal literature on student and peer effects of exclusionary discipline given its scope, length, and the specificity of the data. However, it is important to note limitations to our current study. Broadly, we discuss five main limitations that speak directly to the inability to extrapolate our effect estimates to particular incidents, students, and contexts; the potential impact of educator bias in incident reporting; and the unexplored role of moderators in determining the effects of exclusionary discipline.

First, all effects are estimated based on students who had some non-0 and non-1 propensity to receive either disciplinary response in each contrast. If a particular level of discipline is required as a response to a particular type of incident, it is not possible to estimate the effect of that level of discipline for that incident because there is no group of students who were involved in the same incident but who received a less exclusionary response. We are only able to estimate effects when there are similar students who received each type of disciplinary response for a particular type of incident. In reality, however, events that result in a near 0 or near 1 propensity are not able to be balanced between groups, particularly if they are low-frequency events. For example, in estimates for the effect of OSS versus ISS, all students who receive an OSS have a propensity weight of 1. Because incidents involving a weapon nearly exclusively result in an OSS and are low frequency among middle schools students (1.2% of incident records), we were not able to obtain balance on this variable—there were simply too few instances of incidents involving weapons that resulted in an ISS to upweight those cases enough to achieve balance. As such, all cases involving a weapon were dropped from middle school analyses, and our estimates cannot be extrapolated to these types of incidents. Because of this,

we also dropped cases in which the NYPD was notified (4.5% and 7.5% of incident records for middle school and high school, respectively) and when an EMS was called (0.3% and 0.5% of incident records for middle school and high school, respectively).

Second, because we estimate the average effect of receiving the more exclusionary disciplinary response compared to the less exclusionary disciplinary response for each comparison, the estimates are more heavily influenced by the types of incidents, for the types of students, and in the types of schools that have the most variation in responses. They are average estimates, and as such, they cannot be said to represent the effect of the response equally for each and every incident and student. Using the example comparison of OSS versus ISS again, Figure 3 demonstrates that the most variation in whether a middle school student receives an ISS or an OSS happens for incidents that are categorized as Level 4. There are a very small number of Level 3 incidents that result in an OSS and a similarly small number of Level 5 incidents that result in an ISS. As such, average estimates are almost entirely based on Level 4 incidents and are likely less accurate for incidents at either extreme. Additionally, if a student with no prior incidents is far more likely to receive an ISS than an OSS for this type of incident, that student will be downweighted in the comparison because of their low propensity to receive an OSS. As such, average estimates are likely less accurate for these types of students because they contribute less to the comparison. Future work in this area should explore the effects of these types of decisions separately for various types of incidents and in various contexts in order to obtain estimates that are more specific to particular cases.

Third, because comparisons could only be made when there was sufficient overlap between students who received the lesser and more severe responses, there were certain comparisons that our models were not able to make. Specifically, in the main analyses, we were not able to compare students who received an ISS to those who received a classroom removal. These students were too different, and there were too few cases of classroom removal to allow for our required balance to be achieved. Additionally, we were not able to conduct subgroup analyses for White middle schoolers or for students who identified as a race or ethnicity other than White, Black, and Hispanic due to the relatively smaller number of students in these racial/ethnic categories and therefore the increased difficulty achieving balance between groups of students in each comparison. Also due to smaller numbers as the data are broken up into more categories, we were not able to compare every increment of suspension length (e.g., 1 day compared to 2 days, 2 days compared to 3 days, 3 days compared to 4 days) and instead needed to group exclusions into broader severity categories.

Fourth, all propensity score weights are based on what was reported about a particular incident. Incident reports are made by humans and are subject to explicit and implicit human biases. A consistent body of research suggests that bias, including but not limited to, explicit

and implicit racial and gender bias, affects disciplinary referrals (Girvan et al., 2019; Osher et al., 2015; Rocque, 2010; Santiago-Rosario et al., 2021). It is not only possible, but likely, that there are underlying differences in students that lead to differences in the type of incident for which they are reported. As such, a Black student and a White student who did the exact same thing may be reported for incidents of varying severity if racial biases cause the teacher or principal to view the incident differently. By including student demographic characteristics, prior behavioral history, prior academics and school demographics, academics, and climate in weighting models, we seek to remedy this by creating a group of comparison students who are similar to students receiving the more severe punishment; however, it is important to note that no controls are perfect.

Finally, although this study gets closer to the ground than many similar studies and is able to compare students using more specific information about their behavioral incident and school context (e.g., Anderson et al., 2019), this study did not examine what factors in the students' ecology may be affecting particular coactions between and among students and teachers and how those actions and perceptions affect student trajectories and school climate (Osher et al., 2020). For example, we do not explore moderating factors, such as teacher-student racial match, the quality of mental health services available to teachers and students, the existence and quality of restorative practices, the consistency or variability of disciplinary responses within the school, or peers' perceptions of a student's removal, particularly when there is a shared identity with the suspended peer. Future research is needed to provide a more nuanced picture of how ecological context interacts with exclusionary discipline policy and practice.

References

Academy Council on School Health. (2013). Out-of-school suspension and expulsion. *Pediatrics*, *131*(3), e1000-e1007.

Aladjem, D. K., Birman, B. F., Orland, M., Harr-Robins, J., Heredia, A., Parrish, T. B., Ruffini, S. J. (2010). Achieving dramatic school improvement: An exploratory study. A cross-site analysis from the Evaluation of Comprehensive School Reform Program Implementation and Outcomes study. Office of Planning, Evaluation and Policy Development, & U.S. Department of Education.

American Association of School Administrators (AASA; The School Superintendents Association). (n.d.). School discipline: Dismantle the pre-K to prison pipeline. https://www.aasa.org/schooldiscipline.aspx

American Civil Liberties Union. (2017, April 12). *Bullies in blue: Origins and consequences of school policing*. https://www.aclu.org/report/bullies-blue-origins-and-consequences-school-policing

American Federation of Teachers. (2020). *State of the Union 2018–2020: We care. We fight. We show up. We vote*. https://www.aft.org/sites/default/files/sotu_2020.pdf

American Psychological Association Zero Tolerance Task Force. (2008). Are zero tolerance policies effective in the schools?: An evidentiary review and recommendations. *American Psychologist, 63*, 852–862. http://dx.doi.org/10.1037/0003-066X.63.9.852

Anderson, K. P., & Ritter, G. W. (2017). Disparate use of exclusionary discipline: Evidence on inequities in school discipline from a U.S. state. *Education Policy Analysis Archives, 25*(49). http://dx.doi.org/10.14507/epaa.25.2787

Anderson, K. P., Ritter, G. W., & Zamarro, G. (2019). Understanding a vicious cycle: The relationship between student discipline and student academic outcomes. *Educational Researcher*, *48*(5), 251–262.

Artiles, A. J., Kozleski, E. B., Trent, S. C., Osher, D., & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968–2008: A critique of underlying views of culture. *Exceptional Children*, *76*(3), 279–299.

Arum, R. (2003). *Judging school discipline: The crisis of moral authority.* Harvard University Press.

Astor, R. A., Noguera, P., Fergus, E., Gadsden, V., & Benbenishty, R. (2020). A call for the conceptual integration of opportunity structures within school safety research. *School Psychology Review*, 1–19.

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

Balfanz, R., Byrnes, V., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5*(2), Article 13.

Balfanz, R., Byrnes, V., & Fox, J. H. (2015). Sent home and put off track. In D. J. Losen, *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 17–30). Teachers College Press.

Bear, G. G. (2014). Preventive and classroom-based strategies. In *Handbook of classroom management* (pp. 25-49). Routledge.

Benbenishty, R., Astor, R. A., & Astor, R. (2005). *School violence in context: Culture, neighborhood, family, school, and gender*. Oxford University Press on Demand.

Benbenishty, R., & Astor, R. A. (2019). Conceptual foundations and ecological influences of school violence, bullying, and safety. In M. J. Mayer & S. R. Jimerson (Eds.), *School safety and violence prevention: Science, practice, policy* (pp. 19–44). American Psychological Association. https://doi.org/10.1037/0000106-002

Borman, G. D., Hewes, G. M., Overman, L. T., Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. Review of Educational Research, 73(2), 125–230. https://doi.org/10.3102/00346543073002125

Bradshaw, C. P., Sawyer, A. L., & O'Brennan, L. M. (2009). A social disorganization perspective on bullying-related attitudes and behaviors: The influence of school context. *American Journal of Community Psychology, 43*(3–4), 204–220. https://doi.org/10.1007/s10464-009-9240-1

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics*, *130*(5), e1136–e1145. https://doi.org/10.1542/peds.2012-0243

Carpenter, D. M., & Ramirez, A. (2007). More than one gap: Dropout rate gaps between and among Black, Hispanic, and White students. *Journal of Advanced Academics*, *19*(1), 32–64. https://doi.org/10.4219/jaa-2007-705

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics, 2*(1), 211–228. https://doi.org/10.3386/w14246

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi:10.1145/2939672.2939785

Chu, E. M., & Ready, D. D. (2018). Exclusion and urban public high schools: Short-and long-term consequences of school suspensions. *American Journal of Education*, *124*(4), 479–509.

Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology, 36*(1), 59–82. https://doi.org/10.1016/S0022-4405(97)00050-2

Deming, D. J. (2011). Better schools, less crime? *The Quarterly Journal of Economics, 126*(4), 2063–2115. https://doi.org/10.1093/qje/qjr036

Diamond, J. B., & Lewis, A. E. (2016). Race and discipline at a racially mixed suburban high school: Status, capital, and the practice of organizational routines. *Peabody Journal of Education, 94*(2). https://doi.org/10.1080/0161956X.2016.1184951

Dishion, T. J., & Patterson, G. R. (2006). The development and ecology of antisocial behavior in children and adolescents. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Risk, disorder, and adaptation* (pp. 503–541). John Wiley & Sons, Inc.

Dodge, K. A., Dishion, T. J., & Lansford, J. E. (2006). Deviant peer influences in intervention and public policy for youth. *Social Policy Report, 20*(1), 1–20. https://doi.org/10.1002/j.2379-3988.2006.tb00046.x

Dwyer, K., Osher, D., & Hoffman, C. C. (2000). Creating responsive schools: Contextualizing early warning, timely response. *Exceptional Children, 66*(3), 347–365.

Eccles, J. S., Midgley, C., Wigeld, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Iver, D. M. (1993). Development during adolescence: The impact of stage-environment on young adolescents' experiences in schools and in families. *American Psychologist*, *48*(2), 90–101.

Eden, M. (2017). School discipline reform and disorder: Evidence from New York City Public Schools, 2012–16. *The Education Digest*, *83*(1), 22.

Eden, M. (2019). *Safe and orderly schools: Updated guidance on school discipline: Report.* Manhattan Institute for Policy Research.

Educators for Excellence. (2018). Voices from the classroom: A survey of America's educators. https://e4e.org/sites/default/files/2018_voices_from_the_classroom_teacher_survey.pdf

Ewing, C. P. (2000, Jan.–Feb.). *Sensible zero tolerance protects students. Harvard Education Letter: Research online* (ED456938). https://files.eric.ed.gov/fulltext/ED456938.pdf

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement.* Council of State Governments Justice Center.

Ferguson, A. (2000). *Bad boys.* University of Michigan Press.

Figlio, D. N. (2007). Boys named Sue: Disruptive children and their peers. *Education Finance and Policy, 2*(4), 376–394. https://doi.org/10.3386/w11277

Fletcher, J. (2010). Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management*, *29*(1), 69–83. https://doi.org/10.2139/ssrn.1424191

Girvan, E. J. (2019). The law and social psychology of racial disparities in school discipline. In B. Bornstein & M. Miller (Eds.), *Advances in psychology and law*, vol. 4 (pp. 235–276). Springer, Cham. https://doi.org/10.1007/978-3-030-11042-0_8

Gopalan, M. (2019). Understanding the linkages between racial/ethnic discipline gaps and racial/ethnic achievement gaps in the United States. *Education Policy Analysis Archives, 27*(154). https://doi.org/10.14507/epaa.27.4469

Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, *47*(2), 434–465. https://doi.org/10.3102/0002831209350494

Gottfried, M. A. (2014). Classmates with disabilities and students' noncognitive outcomes. *Educational Evaluation and Policy Analysis*, *36*(1), 20–43. https://doi.org/10.3102/0162373713493130

Gregory, A., Osher, D., Bear, G. G., Jagers, R. L., & Sprague, J. R. (2021). Good intentions are not enough: Centering equity in school discipline reform. *School Psychology Review.* https://doi.org/10.1080/2372966X.2020.1861911

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59–68. https://doi.org/10.3102%2F0013189X09357621

Gresham, F. M., & Lane, K. L. (2019). Supporting students with intensive educational needs in a comprehensive, integrated, three-tiered model of prevention: Considerations for the field. In Osher, D., Mayer, M. J., Jagers, R. J., Kendziora, K., & Wood, L. (Eds.), *Keeping students safe and helping them thrive: A collaborative handbook on school safety, mental health, and wellness (II, 142–160).* Praeger/ABC-CLIO.

Guarino, K., & Caverly, S. (2019). The impact of trauma and adversities. In D. Osher, M. J. Mayer, R. J. Jagers, K. Kendziora, & L. Wood (Eds.), *Keeping students safe and helping them thrive: A collaborative handbook on school safety, mental health, and wellness (II, 80–115).* Praeger/ABC-CLIO.

Hammond, Z. (2020). Looking at SoLD through an equity lens: Will the science of learning and development be used to advance critical pedagogy or will it be used to maintain inequity by design? *Applied Developmental Science*, *24*(2), 151–158.

Hemphill, S. A., Kotevski, A., Herrenkohl, T. I., Smith, R., Toumbourou, J. W., & Catalano, R. F. (2013). Does school suspension affect subsequent youth nonviolent antisocial behavior? A longitudinal study of students in Victoria, Australia and Washington State, United States. *Australian Journal of Psychology, 65*(4), 236–249. https://doi.org/10.1111/ajpy.12026

Hinze-Pifer, R., & Sartain, L. (2018). Rethinking universal suspension for severe student behavior. *Peabody Journal of Education*, *93*(2), 228–243. https://doi.org/10.1080/0161956X.2018.1435051

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

Hwang, N. (2018). Suspensions and achievement: Varying links by type, frequency, and subgroup. *Educational Researcher, 47*(6), 363–374. https://doi.org/10.3102/0013189X18779579

Hwang, N., & Domina, T. (2020). Peer disruption and learning: Links between suspensions and the educational achievement of non-suspended students. *Education Finance and Policy,* 1–44. https://doi.org/10.1162/edfp_a_00308

Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review, 102*(5), 2048–2082. https://doi.org/10.1257/aer.102.5.2048

Interdisciplinary Group on Preventing School and Community Violence. (2013). December 2012 Connecticut school shooting position statement. *Journal of School Violence*, *12*(2), 119–133.

Jennings, P., Domenici, A., & Yoder, N. (2019). Creating the working conditions to enhance teacher social and emotional well-being. In D. Osher, M. J. Mayer, R. J. Jagers, K. Kendziora, & L. Wood (Eds.), *Keeping students safe and helping them thrive: A collaborative handbook on school safety, mental health, and wellness* (I, 210–239). Praeger/ABC-CLIO.

Kamenetz, A. (2017, March 23). *School suspensions have plunged: We don't yet know if that's good news*. NRP Ed. https://www.npr.org/sections/ed/2017/03/23/521070924/school-suspensions-have-plunged-we-don-t-yet-know-if-that-s-good-news

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., Petras, H., Ford, C., Windham, A., & Wilcox, H. C. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95,* S5–S28. https://doi.org/10.1016/j.drugalcdep.2008.01.004

Kirk, D. S., & Matsuda, M. (2011). Legal cynicism, collective efficacy, and the ecology of arrest. *Criminology, 49*(2), 443–472. https://doi.org/10.1111/j.1745-9125.2011.00226.x

Kirk, D. S., & Papachristos, A. V. (2011). Cultural mechanisms and the persistence of neighborhood violence. *American Journal of Sociology, 116*(4), 1190–1233. https://doi.org/10.1086/655754

Kristoffersen, J. H. G., Krægpøth, M. V., Nielsen, H. S., & Simonsen, M. (2015). Disruptive school peers and student outcomes. *Economics of Education Review*, *45*, 1–13. https://doi.org/10.1016/j.econedurev.2015.01.004

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

Lacoe, J., & Steinberg, M. P. (2018). Rolling back zero tolerance: The effect of discipline policy reform on suspension usage and student outcomes. *Peabody Journal of Education*, *93*(2), 207–227. https://doi.org/10.1080/0161956X.2018.1435047

Lacoe, J., & Steinberg, M. P. (2019). Do suspensions affect student outcomes? *Educational Evaluation and Policy Analysis*, *41*(1), 34–62. https://doi.org/10.3102/0162373718794897

Lacoe, J. R. (2015). Unequally safe: The race gap in school safety. *Youth Violence and Juvenile Justice*, *13*(2), 143–168. https://doi.org/10.1177/1541204014532659

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, *29*(3), 337–346. https://doi.org/10.1002/sim.3782

Linick, M., & D'Amico, N. (2014). *Student absences and academic performance*: *An examination of the non-linear relationship between student absences and different academic outcomes.* Cleveland Metropolitan School District Senior Leadership Team Meeting, Cleveland, OH.

Losen, D. J. (2011). *Discipline policies, successful schools, and racial justice*. National Education Policy Center. http://nepc.colorado.edu/publication/discipline-policies

Losen, D. J., & Whitaker, A. (2018, August 24). *11 million days lost: Race, discipline, and safety at U.S. public schools (Part 1).* American Civil Liberties Union. https://www.aclu.org/report/11-million-days-lost-race-discipline-and-safety-us-public-schools-part-1

Mader, N., Sartain, L., & Steinberg, M. (2016). *When suspensions are shorter: The effects on school climate and student learning* [Manuscript in preparation].

Mauer, M., & Chesney-Lind, M. (2002). *Invisible punishment: The collateral consequences of mass imprisonment.* The New Press.

Mayer, G. R., Butterworth, T., Nafpaktitis, M., & Sulzer-Azaroff, B. (1983). Preventing school vandalism and improving discipline: A three-year study. *Journal of Applied Behavior Analysis*, *16*(4), 355–369.

Mayer, G. R. (1995). Preventing antisocial behavior in the schools. *Journal of Applied Behavior Analysis*, *28*(4), 467–478.

Mayer, M. J., & Salinger, T. (2019). Students with school failure trajectories. In D. Osher, M. J. Mayer, R. J. Jagers, K. Kendziora, & L. Wood (Eds.), *Keeping students safe and helping them thrive: A collaborative handbook on school safety, mental health, and wellness (II, 116–141).* Praeger/ABC-CLIO.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine, 32*(19), 3388–3414. https://doi.org/10.1002/sim.5753

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403–425. https://doi.org/10.1037/1082-989X.9.4.403

Mendez-Raffaele, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children*, *26*(1), 30–51.

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. https://doi.org/10.1007/s10964-014-0103-1

Morgan, E., Salomon, N., Plotkin, M., & Cohen, R. (2014). *The school discipline consensus report: Strategies from the field to keep students engaged in school and out of the juvenile justice system.* Council of State Governments Justice Center.

Morris, E. W., & Perry, B. L. (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems*, *63*(1), 68–86. https://doi.org/10.1093/socpro/spv026

Nasir, N. I. S. (2020). Teaching for equity: Where developmental needs meet racialized structures. *Applied Developmental Science*, *24*(2), 146–150.

Nese, R. N. T., Nese, J. F. T., McIntosh, K., Mercer, S. H., & Kittelman, A. (2019). Predicting latency of reaching adequate implementation of tier 1 schoolwide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions, 21*(2), 106–116. https://doi.org/10.1177/1098300718783755

Nolan, K. (2011). *Police in the hallways: Discipline in an urban high school*. University of Minnesota Press.

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

Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A vicious cycle: A social–psychological account of extreme racial disparities in school discipline. *Perspectives on Psychological Science*, *11*(3), 381–398.

Osher, D., Bear, G., Sprague, J. R., & Doyle, W. (2010). How can we improve school discipline? *Educational Researcher, 39*(1), 48–58. https://doi.org/10.3102/0013189X09357618

Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2020). Drivers of human development: How relationships and context shape learning and development. *Applied Developmental Science*, *24*(1), 6–36. https://doi.org/10.1080/10888691.2017.1398650

Osher, D., Cantor, P., & Caverly, S. (2019). The relational, ecological, and phenomenological foundations of school safety, mental health, wellness, and learning. In D. Osher, M. J. Mayer, R. J. Jagers, K. Kendziora, & Wood, L. (Eds.), *Keeping students safe and helping them thrive: A collaborative handbook on school safety, mental health, and wellness (I, pp. 29–58)*. Praeger/ABC-CLIO.

Osher, D., Coggshall, J., Colombi, G., Woodruff, D., Francois, S., & Osher, T. W. (2012). Building school and teacher capacity to eliminate the school-to-prison pipeline. *Teaching Exceptional Children*, *32*(3), 30–37.

Osher, D., Fisher, D., Amos, L., Katz, J., Dwyer, K., Duffey, T., & Colombi, G. D. (2015). *Addressing the root causes of disparities in school discipline: An educator's action planning guide.* National Center on Safe Supportive Learning Environments.

Osher, D., Kidron, Y., DeCandia, C. J., Kendziora, K., & Weissberg, R. P. (2016). Interventions to promote safe and supportive school climate. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of Social Influences in School Contexts* (pp. 384–404). Routledge.

Osher, D., Sprague, J., Weissberg, R. P., Axelrod, J., Keenan, S., Kendziora, K., & Zins, J. E. (2008). A comprehensive approach to promoting social, emotional, and academic growth in contemporary schools. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (vol. 4, pp. 1263–1278). National Association of School Psychologists.

Osher, D., VanAker, R., Morrison, G., Gable, R., Dwyer, K., & Quinn, M. (2004). Warning signs of problems in schools: Ecological perspectives and effective practices for combating school aggression and violence. *Journal of School Violence, 2*(3), 13–37. https://doi.org/10.1300/J202v03n02_03

Osher, D., Woodruff, D., & Sims, A. (2002). Schools make a difference: The relationship between education services for African American children and youth and their overrepresentation in the juvenile justice system. In D. Losen (Ed.), *Minority issues in special education* (pp. 93–116). The Civil Rights Project, Harvard Education Publishing Group, Harvard University.

Perry, B. L., & Morris, E. W. (2014). Suspending progress: Collateral consequences of exclusionary punishment in public schools. *American Sociological Review*, *79*(6), 1067–1087. https://doi.org/10.1177/0003122414556308

Public Agenda. (2004). *Teaching interrupted: Do discipline policies in today's public schools foster the common good?*

Ramchand, R., Griffin, B. A., Harris, K. M., & Morral, A. R. (2011). Using a cross-study design to assess the efficacy of MET/CBT-5 in treating adolescents with cannabis-related disorders. *Journal of Studies on Alcohol and Drugs*, *72*(3), 380–389. https://doi.org/10.15288/jsad.2011.72.380

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., Tabor, J., Beuhring, T., Sieving, R. E., Shew, M., Ireland, M., Bearinger, L. H., & Udry, J. R. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *Journal of the American Medical Association*, *278*(10), 823–832.

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

Rocque, M. (2010). Office discipline and student behavior: Does race matter?. *American Journal of Education*, *116*(4), 557–581.

Rosenbaum, J. E. (2020). Educational and criminal justice outcomes 12 years after school suspension. *Youth & Society*, *52*(4), 515–547. https://doi.org/10.1177/0044118X17752208

Ryberg, R., Her, S., Temkin, D., & Harper, K. (2021). *Despite reductions since 2011-12, black students and students with disabilities remain more likely to experience suspension*. Child Trends. https://www.childtrends.org/publications/despite-reductions-black-students-and-students-with-disabilities-remain-more-likely-to-experience-suspension

Santiago-Rosario, M. R., Whitcomb, S. A., Pearlman, J., & McIntosh, K. (2021). Associations between teacher expectations and racial disproportionality in discipline referrals. *Journal of School Psychology*, *85*, 80–93.

Sartain, L., Allensworth, E. M., Porter, S., Levenstein, R., Johnson, D. W., Huynh, M. H., Anderson, E., Mader, N., & Steinberg, M. P. (2015). *Suspending Chicago's students: Differences in discipline practices across schools.* University of Chicago Consortium on Chicago School Research. https://consortium.uchicago.edu/publications/suspending-chicagos-students-differences-discipline-practices-across-schools

Skiba, R. J. (2014). The failure of zero tolerance. *Reclaiming Children and Youth*, *22*(4), 27–33.

Skiba, R. J., Mediratta, K., & Rausch, M. K. (2016). *Inequality in school discipline: Research and practice to reduce disparities*. Palgrave Macmillan.

Sorensen, L. C., Bushway, S. D., & Gifford, E. J. (2021). *Getting tough? The effects of discretionary principal discipline on student outcomes* [EdWorkingPaper: 20-216]. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai20-216

Steinberg, M. P., Allensworth, E., & Johnson, D. W. (2011). *Student and teacher safety in Chicago Public Schools: The roles of community context and school social organization.* Consortium on Chicago School Research.

Steinberg, M. P., & Lacoe, J. (2017). *The academic and behavioral consequences of discipline policy reform: Evidence from Philadelphia.* Thomas B. Fordham Institute.

Steinberg, M. P., & Lacoe, J. (2018). Reforming school discipline: School-level policy implementation and the consequences for suspended students and their peers. *American Journal of Education*, *125*(1), 29–77. https://doi.org/10.1086/699811

Suh, S., & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling, 10*(3), 297–306. https://doi.org/10.1177/2156759X0701000312

Swanson, E., Erickson, H. H., & Ritter, G. W. (2017). *Examining the impacts of middle school disciplinary policies on 9th grade retention* (EDRE Working Paper No. 2017–11). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2952972

U.S. Department of Education & Department of Justice. (2014). *Notice of language assistance: Dear colleague letter on the nondiscriminatory administration of school discipline* [Joint letter, Office for Civil Rights, Civil Rights Division]. https://www2.ed.gov/about/offices/list/ocr/letters/colleague-201401-title-vi.html

U.S. Department of Education, Office for Civil Rights. (2018). *2015–16 Civil Rights Data Collection: School climate and safety – Data highlights on school climate and safety in our nation's public schools.* https://www2.ed.gov/about/offices/list/ocr/docs/school-climate-and-safety.pdf

Valdebenito, S., Eisner, M., Farrington, D. P., Ttofi, M. M., & Sutherland, A. (2019). What can we do to reduce disciplinary school exclusion? A systematic review and meta-analysis. *Journal of Experimental Criminology*, *15*(3), 253–287. https://doi.org/10.1007/s11292-018-09351-0

Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development, 2003*(99), 9–15. https://doi.org/10.1002/yd.51

Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices*. Thomson/Wadsworth.

Watanabe, T. (2013, May 14). L.A. Unified bans suspension for "willful defiance." *Los Angeles Times.* https://www.latimes.com/local/la-xpm-2013-may-14-la-me-lausd-suspension-20130515-story.html

Welsh, R. O., & Little, S. (2018). The school discipline dilemma: A comprehensive review of disparities and alternative approaches. *Review of Educational Research*, *88*(5), 752–794.

Wilson, J. Q., & Kelling, G. L. (1982). Broken windows: The police and neighborhood safety. *Atlantic Monthly*, *249*(3), 29–36, 38.

Wolf, K. C., & Kupchik, A. (2017). School suspensions and adverse experiences in adulthood. *Justice Quarterly, 34*(3), 407–430. https://doi.org/10.1080/07418825.2016.1168475

Appendix A: Data Supplements

Table A.1. Distribution of Incidents by Level and Year: Middle School Students

Year	Level 1	Level 2	Level 3	Level 4	Level 5
2009	7,594	4,807	20,939	26,173	3,285
2010	7,071	5,726	23,792	26,771	3,219
2011	10,060	6,776	25,837	28,330	3,463
2012	15,136	6,219	26,822	28,256	2,669
2013	13,401	8,510	21,155	24,521	2,286
2014	14,190	9,169	23,261	20,779	2,398
2015	10,595	6,972	19,843	18,695	2,302
2016	9,224	6,227	19,403	18,019	2,158
2017	8,468	6,020	18,576	18,345	2,165
2018	12,412	7,514	23,524	21,120	2,111

Table A.2. Distribution of Incidents by Level and Year: High School Students

Year	Level 1	Level 2	Level 3	Level 4	Level 5
2009	5,569	6,726	21,328	20,978	3,941
2010	7,048	6,924	22,885	21,758	3,785
2011	12,492	7,775	25,670	24,404	3,675
2012	20,578	7,421	29,058	23,841	3,353
2013	16,767	8,327	21,919	22,389	2,940
2014	16,541	8,314	25,411	21,185	3,236
2015	13,794	8,612	21,845	21,877	3,371
2016	17,877	8,728	19,397	22,713	3,544
2017	15,365	8,021	18,974	21,873	3,122
2018	15,612	7,953	19,298	23,063	3,596

Table A.3. Distribution of Discipline by Incident Level: Middle School Students

Exclusion	Type	Duration	Level 1	Level 2	Level 3	Level 4
No exclusion			95,657	55,116	144,263	108,911
Exclusion	Removal	1–5 days	12,428	7,275	20,219	6,346
ISS	ISS	1 day	<100	1,075	11,132	9,524
		2–3 days	<100	2,728	29,576	40,708
		4–5 days	<100	1,111	12,892	26,661
OSS	OSS	1–5 days	0	0	273	3,588
	6–20 days	6–20 days	0	0	1,193	14,402
	1–5 days	21+ days	0	0	433	14,630

Note. ISS = in-school suspension; OSS = out-of-school suspension.

Table A.4. Distribution of Discipline by Incident Level: High School Students

Exclusion	Type	Duration	Level 1	Level 2	Level 3	Level 4
No exclusion			135,067	58,568	113,921	92,114
Exclusion	Removal	1–5 days	6,560	3,833	5,616	1,295
ISS	ISS	1 day	<100	3,151	15,119	6,329
	ISS	2–3 days	<100	8,140	50,460	39,295
	ISS	4–5 days	<100	5,220	47,921	83,788
OSS	OSS	1–5 days	0	0	185	2,623
	OSS	6–20 days	0	0	827	10,172
	OSS	21+ days	0	0	251	11,102

Note. ISS = in-school suspension; OSS = out-of-school suspension.

Table A.5. Race/Ethnicity Distribution of Students, by Group

Incident or Exclusion	Black	Hispanic	Other	White
All students	29.0%	39.8%	17.5%	13.8%
Students with an incident	49.4%	36.8%	5.0%	8.7%
Students with a Level 4 or 5 incident	50.5%	37.2%	4.9%	7.3%
Students with any type of exclusion	50.4%	37.5%	4.7%	7.3%
Students assigned ISS	49.5%	37.5%	5.1%	7.9%
Students assigned OSS	55.7%	35.6%	3.9%	4.7%
Students assigned OSS 21+ days	57.3%	31.4%	7.7%	3.6%

Note. ISS = in-school suspension; OSS = out-of-school suspension.

Table A.6. School Climate Survey Items, by Survey Year and Construct

Item	2014–15 survey	2015–16 survey	2016–17 survey	2017–18 survey
Classroom experience				
Adults at this school help students aspiring to enter the workforce, develop a plan to reach their future employment goals.	X			
Adults at this school provide students with information about the college enrollment process.	X			
Adults at this school help keep me on track for college or career.	X			
Adults at this school support students in navigating the post-secondary process.	X			
This school programs students with appropriate courses to achieve their postsecondary goals.	X			
There is an adult who is helping me plan for my next steps after graduation (career planning, college selection and application process, financial aid process, etc.)	X			
My school is kept clean.		X	X	X
Most students at my school treat each other with respect.		X	X	X
My teachers call on students of different races, ethnicities, cultures, and backgrounds.		X		
I feel safe outside around this school.		X	X	X
I feel safe traveling between home and this school.		X	X	X
I feel safe in the hallways, bathrooms, locker rooms, and cafeteria of this school.		X	X	X
I feel safe in my classes at this school.		X	X	X
At this school students harass, bully or intimidate other students.		X	X	X
At this school students harass, bully, or intimidate each other because of their race or ethnicity.		X		
At this school students harass, bully, or intimidate each other because of their gender, gender identity, gender expression, or sexual orientation.		X	X	X
At this school students harass, bully, or intimidate each other because of other differences, like national origin, citizenship/immigration status, religion, disability, or weight.		X		
At my school students get into physical fights.		X	X	X
At my school there is gang activity.		X	X	X
Adults at this school (including teachers, administrators, counselors, and the principal) help keep me on track for college or career.		X		

Item	2014–15 survey	2015–16 survey	2016–17 survey	2017–18 survey
Adults at this school (including teachers, administrators, counselors, and the principal) provide me with information about the college enrollment process.		X		
Adults at this school (including teachers, administrators, counselors, and the principal) help me plan for my next steps after graduation (career planning, college selection and application process, financial aid process, etc.).		X		
Adults at this school (including teachers, administrators, counselors, and the principal) help me choose which colleges to apply to.		X		
Adults at this school (including teachers, administrators, counselors, and the principal) talk with me about how to pay for college.		X		
Adults at this school (including teachers, administrators, counselors, and the principal) help me plan for how to meet my future career goals		X	X	X
At this school students harass, bully, or intimidate each other because of their race, religion, ethnicity, national origin, or citizenship/immigration status.			X	X
At this school students harass, bully, or intimidate each other because of other differences, like disability or weight.			X	X
Adults at this school (including teachers, administrators, counselors, and the principal) encourage me to continue my education after high school.			X	X
Adults at this school (including teachers, administrators, counselors, and the principal) provide me with information about the college enrollment process (college selection and application process, financial aid process, course registration, etc.).			X	
Adults at this school (including teachers, administrators, counselors, and the principal) help me consider which colleges to apply to.			X	X
Adults at this school (including teachers, administrators, counselors, and the principal) show me options for how to pay for college (scholarship, grants, loans, work study programs, etc.).			X	X
Adults at this school (including teachers, administrators, counselors, and the principal) talk to me about what I plan to do after high school.				X
Adults at this school (including teachers, administrators, counselors, and the principal) provide me with information about the college application process.				X

Item	2014–15 survey	2015–16 survey	2016–17 survey	2017–18 survey
School Environment				
My classes at this school really make me think.	X	X		
How many of the students in your class(es) feel it is important to come to school every day?	X			
How many of the students in your class(es) feel it is important to pay attention in class?	X			
How many of the students in your class(es) think doing homework is important?	X			
How many of the students in your class(es) try hard to get good grades?	X			
In my classes, my teachers expect students to work hard.	X			
In my classes, my teachers expect me to do my best all the time.	X			
In my classes, my teachers expect me to learn from my peers.	X			
In my classes, my teachers want students to become better thinkers, not just memorize things.	X			
I feel safe outside around this school.	X			
I feel safe traveling between home and this school.	X			
I feel safe in the hallways, bathrooms, locker rooms, and cafeteria of this school.	X			
I feel safe in my classes at this school.	X			
Students in this school get to know each other well in classes.	X			
Students in this school are very interested in getting to know other students.	X			
Students in this school enjoy doing things with each other during school activities.	X			
Students in this school get to know each other really well.	X			
Students in this school enjoy working together on projects in classes.	X			
The teachers at this school help me catch up if I am behind.	X			
The teachers at this school are willing to give extra help on schoolwork if I need it.	X			
The teachers at this school notice if I have trouble learning something.	X			
The teachers at this school give me specific suggestions about how I can improve my work in class.	X			
The teachers at this school compliment me if I do good work.	X			
The teachers at this school explain things a different way if I don't understand something in class.	X			

Item	2014–15 survey	2015–16 survey	2016–17 survey	2017–18 survey
The teachers at this school notice when I am upset or having emotional difficulty.	X			
My teachers will always listen to students' ideas.	X	X	X	X
My teachers always keep their promises.	X	X		
My teachers incorporate students' cultures/ backgrounds into the curriculum to make learning more meaningful.	X			
I see people of many cultures/backgrounds represented in the curriculum.	X			
My teachers connect to students of different cultures/backgrounds.	X			
My school offers a wide enough variety of programs, classes and activities to keep me interested in school.	X	X	X	X
The programs, classes, and activities at this school encourage students to develop talent outside academics.	X	X	X	X
My school is kept clean.	X			
Most students at my school treat each other with respect.	X			
At my school students harass or bully other students.	X			
At my school students harass or bully each other based on differences (such as race, color, ethnicity, national origin, citizenship/immigration status, religion, gender, gender identity, gender expression, sexual orientation, disability or weight).	X			
At my school students get into physical fights.	X			
At my school there is gang activity.	X			
This school provides useful information to students about the application/enrollment process to high school.	X	X	X	
This school provides guidance for the application process for high school.	X	X	X	X
I'm learning a lot in my classes at this school.	X			
I feel safe and comfortable with my teachers at this school.	X			
There is at least one adult in the school that I can confide in.	X	X	X	X
My teachers treat me with respect.	X	X	X	X
When my teachers tell me not to do something, I know they have a good reason.	X	X	X	X
My teachers appreciate my culture/background.	X			
It's clear what I need to do to get a good grade.		X	X	X
The work we do in class is good preparation for our class tests.		X	X	X
The homework assignments help me learn the course material.		X	X	X
In how many of your classes are you challenged?		X	X	X

Item	2014–15 survey	2015–16 survey	2016–17 survey	2017–18 survey
My teachers use examples of students' different cultures/backgrounds/families in their lessons to make learning more meaningful for me.		X	X	X
I see people of many races, ethnicities, cultures, and backgrounds represented in the curriculum.		X	X	X
In how many of your classes do your teachers ask difficult questions on tests?		X	X	X
In how many of your classes do your teachers ask difficult questions in class?		X	X	X
In how many of your classes do you work in small groups?		X	X	X
In how many of your classes do your teachers want students to become better thinkers, not just memorize things?		X	X	X
In how many of your classes at this school do students listen carefully when the teacher gives directions?		X		
In how many of your classes at this school do students follow the rules in class?		X		
In how many of your classes at this school do students pay attention when they are supposed to?		X		
In how many of your classes at this school do students work when they are supposed to?		X		
In how many of your classes at this school do students behave well even when the teacher isn't watching?		X		
In how many of your classes at this school do students feel it is important to come to school every day?		X		
In how many of your classes at this school do students feel it is important to pay attention in class?		X		
In how many of your classes at this school do students think doing homework is important?		X		
In how many of your classes at this school do students try hard to get good grades?		X		
I feel that my teachers respect my culture/background.		X	X	X
I learn a lot from feedback on my work.		X	X	X
I know what my teacher wants me to learn in class.		X	X	X
My teachers help me catch up if I am behind.		X	X	X
My teachers notice if I have trouble learning something.		X	X	X
My teachers give me specific suggestions about how I can improve my work in class.		X	X	X
My teachers explain things a different way if I don't understand something in class.		X	X	X
My teachers notice when I am upset.		X		

Item	2014–15 survey	2015–16 survey	2016–17 survey	2017–18 survey
I'm learning a lot in my classes at this school to prepare me for the next level or grade.		X	X	X
Discipline is applied fairly in my school.		X	X	X
My classes at this school really make me think critically.			X	X
In general, my teachers treat students from different cultures or backgrounds equally.			X	X
In general, my teachers make their lessons relevant to my everyday life experiences.			X	X
In general, my teachers present positive images of people from a variety of races, ethnicities, cultures, and backgrounds.			X	X
My teachers always do what they say they will do.			X	X
In how many of your classes at this school do YOU feel most students pay attention when they are supposed to?			X	
In how many of your classes at this school do YOU feel most students work when they are supposed to?			X	X
In how many of your classes at this school do YOU feel most students behave well even when the teacher isn't watching?			X	X
In how many of your classes at this school do YOU feel most students feel it is important to come to school every day?			X	X
In how many of your classes at this school do YOU feel most students feel it is important to pay attention in class?			X	X
In how many of your classes at this school do YOU feel most students think doing homework is important?			X	X
In how many of your classes at this school do YOU feel most students try hard to get good grades?			X	X
My teachers support me when I am upset.			X	X
In how many of your classes at this school do YOU feel most students follow the rules in class?				X
In how many of your classes at this school do YOU feel most students pay attention when they are supposed to?				X

Table A.7. Variables Included in Propensity Weighting Models

Variable level	Variable	Included in middle school model	Included in high school model
Incident	NYCDOE assigned incident level	X	X
	Calendar month	X	X
	School year	X	X
	Number of students involved	X	X
	Location (e.g., "hall," "classroom")	X	X
	Incident codes	X	X
	Number of incident codes	X	X
	Incident involved weapons		X
	Primary incident code (applies to the entire incident, across students)	X	X
Student	Overage 1+ years	X	X
	Overage 2+ years	X	X
	Number of incidents in the past 90 calendar days	X	X
	Number of incidents in the past 90 calendar days of Level 4 or 5	X	X
	Number of incidents in the past 90 calendar days resulting in a suspension	X	X
	Gender	X	X
	Race/ethnicity	X	X
	SWD	X	X
	ELL	X	X
	Poverty	X	X
	Grade	X	X
	Prior year absences	X	X
	Prior year state assessment scale scores in math/ELA	X	
	Grade 8 state assessment scaled score in math/ELA		X
	GPA in prior year	X	X
School	Prior year climate, classroom and school constructs	X	X
	Percent Black	X	X
	Percent Asian	X	X

Variable level	Variable	Included in middle school model	Included in high school model
	Percent Hispanic	X	X
	Percent White	X	X
	Percent SWD	X	X
	Percent ELL	X	X
	Percent poverty	X	X
	Prior year absences	X	X
	Prior year school absences missing	X	X
	Prior year state assessment scaled score in math/ELA	X	

Note. All state assessment scaled scored were *z*-scored by year and subject. ELA = English language arts; ELL = English-language learner; GPA = grade point average; NYCDOE = New York City Department of Education; SWD = students with disabilities.

Table A.8. Most Predictive Variables, by Model

Exclusion versus no exclusion	OSS versus ISS	ISS days	OSS days
Middle school students			
Incident level	Incident level	School climate	School climate
School climate	Count of incident codes	Classroom climate	School total absences, prior year
Classroom climate	Year	School percent White	Prior year ELA score
School percent Hispanic	School percent Black	School percent Black	Prior year math score
School percent SWD	School climate	School percent Asian	Total absences in the prior year
School percent Black	Classroom climate	School percent poverty	Classroom climate
School percent poverty	School percent Asian	School percent SWD	School percent Hispanic
School percent Asian	School percent Hispanic	School percent SWD	School percent SWD
Year	School percent White	School percent ELL	School percent poverty
School percent ELL	School percent poverty	School percent Hispanic	School percent ELL
High school students			
Incident level	Incident level	Classroom climate	Weapons involved in incident
School climate	School year	School climate	Total incident codes in prior 90 days
Classroom climate	School climate	Incident level	School climate
School year	Classroom climate	Number of students involved in incident	Incident level
Persistent Level 1 behavior	Total incident codes in prior 90 days	Total absences in the prior year	Classroom climate
Number of students involved in incident	Number of students involved in incident	Prior year ELA score	Total absences in the prior year
Leaving class or school premises without permission	Location: classroom	Prior year math score	School year
Location: classroom	Was NYPD contacted?	Prior year GPA	Location: scanning area
Profane, obscene, vulgar language or gestures	Total absences in the prior year	School year	Using force against/inflicting to/inflicting serious injury to students
Insubordination	Using force against/inflicting to/inflicting serious injury to students	Total incidents in prior 90 days	Prior year math score

Note. ELA = English language arts; ELL = English-language learner; GPA = grade point average; ISS = in-school suspension; NYPD = New York Police Department; OSS = out-of-school suspension; SWD = students with disabilities.

Table A.9. Covariates With Large Imbalances and Resolution, by Model

Model	Level	Coefficient	Resolution
Any exclusion	Middle school		
	High school		
Type of exclusion, removal versus ISS	Middle school	Incident level	Model not reported
	High school	Incident level	Model not reported
Type of exclusion, ISS versus OSS	Middle school		
	High school		
ISS 1 day versus 2–3 days	Middle school		
	High school		
ISS 2–3 days versus 4–5 days	Middle school		
	High school		
OSS 1–5 days versus 6–20 days	Middle school		
	High school		
OSS 6–20 days versus 21+ days	Middle school	Incident level	Model reported (reader advised to interpret with caution)
	High school		

Note. To avoid imbalance in incidents in which the NYPD was contacted (7.5% of cases) or an EMS was contacted (0.5% of cases), these cases were excluded from analysis in high school. For the same reason, cases were excluded from the middle school analysis when NYPD was contacted (4.5% of cases), an EMS was contacted (0.3% of cases), or incidents involved a weapon (1.2% of cases).

A dash indicates that there was no coefficient with a Cox index or ASMD above 0.25, and no resolution was necessary. ASMD = absolute standardized mean difference; EMS = emergency medical service; ISS = in-school suspension; NYPD = New York Police Department; OSS = out-of-school suspension.

Table A.10. Balance for OSS Versus ISS in Middle School

Variable	Raw OSS versus ISS	ATT weighted OSS versus ISS	Unweighted combined mean
Level 1			
Level 2			
Level 3	**−1.45**	*−0.18*	0.38
Level 4	**1.54**	*0.18*	0.56
Level 5			
Overage 1+ years	**0.29**	0.04	0.42
Overage 2+ years	**0.31**	*0.05*	0.11
Weapons			
SWD	*0.16*	0.03	0.38
ELL	*−0.20*	*−0.05*	0.10
Poverty	*0.17*	0.02	0.81
Total absences in the prior year	**0.33**	*0.11*	18.52
Prior year ELA *z* score	*−0.22*	*−0.05*	−0.53
Prior year math *z* score	*−0.25*	−0.05	−0.61
Black	**0.31**	*0.06*	0.47
White	**−0.55**	−0.03	0.08
Hispanic	*−0.11*	*−0.06*	0.40
Other race	**−0.44**	−0.02	0.05
Total incidents in prior 90 days	**0.36**	*0.16*	1.27
Total incidents of Level 4 or 5 in prior 90 days	**0.48**	*0.14*	0.41
Total suspensions in prior 90 days	**0.37**	*0.13*	0.44
School-level average ELA credits in prior year	*−0.13*	−0.03	0.80
School-level average math credits in prior year	*−0.15*	−0.03	0.80
School-level ELA *z* score	*−0.12*	−0.02	−0.17
School-level math *z* score	*−0.15*	−0.02	−0.21
Classroom experience in prior year	*−0.08*	−0.03	−0.08
School experience in prior year	−0.02	0.00	0.08
School percent Black	**0.32**	*0.09*	0.34
School percent Asian	**−0.28**	0.00	0.11
School percent Hispanic	0.02	*−0.10*	0.44
School percent White	**−0.32**	0.01	0.11
School percent ELL	0.05	0.02	0.13
School percent SWD	*0.22*	0.03	0.22
School percent poverty	*0.23*	0.01	0.76
School absences	**0.27**	0.05	14.08

Note. Regression adjustment was used for variables in the ATT weighted column in italics. Cells with values over the cutoff of 0.25 are bolded. ATT = average treatment effects on the treated; ELA = English language arts; ELL = English-language learner; ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities.

Table A.11. Balance for OSS Versus ISS in High School

	Raw OSS versus ISS	ATT weighted OSS versus ISS	Unweighted combined mean
Level 1			
Level 2			
Level 3	**−1.68**	*−0.20*	0.46
Level 4	**1.85**	*0.20*	0.47
Level 5			
Overage 1+ years	*0.13*	0.03	0.56
Overage 2+ years	*0.14*	0.05	0.23
Weapons			
SWD	*0.10*	0.03	0.30
ELL	*−0.14*	*0.06*	0.09
Poverty	*0.08*	−0.01	0.76
Total absences in the prior year	*0.22*	*0.10*	25.68
Prior year ELA z score	*−0.07*	−0.04	−0.50
Prior year Math z score	*−0.07*	−0.03	−0.58
GPA	*0.06*	0.03	0.34
Black	*0.13*	0.03	0.53
White	*−0.25*	*0.08*	0.07
Hispanic	*−0.07*	−0.04	0.36
Other race	*−0.11*	−0.01	0.04
Total incidents in prior 90 days	**0.25**	*0.13*	0.83
Total incidents of Level 4 or 5 in prior 90 days	**0.44**	*0.13*	0.21
Total suspensions in prior 90 days	**0.27**	*0.11*	0.36
School-level average ELA credits in prior year	0.02	−0.01	0.83
School-level average math credits in prior year	0.00	−0.01	0.78
Classroom experience in prior year	*−0.05*	−0.03	0.13
School experience in prior year	*0.14*	0.03	−0.13
School-level ELA *z* -core	*−0.07*	0.03	−0.20
School-level Math *z*-score	*−0.05*	*0.06*	−0.29
School percent Black	*0.11*	0.04	0.39
School percent Asian	*−0.13*	0.01	0.09
School percent Hispanic	*0.06*	*−0.07*	0.42
School percent White	*−0.20*	0.03	0.08
School percent ELL	−0.04	0.04	0.12
School percent SWD	*0.09*	0.01	0.20
School percent poverty	*0.18*	−0.02	0.74
School absences	*0.07*	0.02	27.12

Note. Regression adjustment was used for variables in the ATT weighted column in italics. Cells with values over the cutoff of 0.25 are bolded. ATT = average treatment effects on the treated; ELA = English language arts; ELL = English-language learner; ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities.

Table A.12. Balance for OSS Dose in Middle School

Variable	Raw 6–20 versus 1–5	Raw 21+ versus 6–20	ATT weighted 6–20 versus 1–5	ATT weighted 21+ versus 6–20	Unweighted combined mean
Level 1					
Level 2					
Level 3	*0.15*	**−0.67**	*0.13*	0.03	0.05
Level 4	*0.09*	*−0.20*	*−0.17*	**−0.25**	0.77
Level 5	*−0.19*	**0.44**	*0.17*	**0.27**	0.18
Overage 1+ years	0.03	*0.20*	−0.01	0.04	0.53
Overage 2+ years	0.00	*0.18*	*0.06*	0.04	0.16
Weapons					
SWD	*−0.13*	*−0.08*	−0.01	−0.02	0.43
ELL	*0.23*	*−0.09*	**0.25**	0.03	0.09
Poverty	*0.11*	*0.07*	0.02	−0.02	0.84
Total absences in the prior year	*−0.10*	**0.28**	0.04	*0.12*	23.81
Prior year ELA *z* score	0.04	*−0.11*	0.01	−0.05	−0.67
Prior year math *z* score	0.03	*−0.14*	−0.01	−0.05	−0.79
Black	*−0.12*	*0.18*	*−0.10*	0.02	0.57
White	*−0.10*	*−0.16*	*0.14*	0.03	0.04
Hispanic	*0.16*	*−0.14*	*0.06*	−0.03	0.36
Other race	*−0.09*	*−0.21*	*0.16*	*0.08*	0.03
Total incidents in prior 90 days	−0.03	*0.19*	*0.09*	*0.09*	1.88
Total incidents of Level 4 or 5 in prior 90 days	*−0.13*	*0.20*	0.04	*0.08*	0.74
Total suspensions in prior 90 days	−0.05	*0.21*	*0.06*	*0.10*	0.70
School-level average ELA credits in prior year	−0.03	0.01	0.04	0.04	0.77
School-level average math credits in prior year	−0.01	0.01	*0.06*	0.04	0.76
School-level ELA *z* score	−0.01	−0.04	0.01	0.00	−0.27
School-level Math z score	0.01	−0.05	0.02	0.00	−0.33
Classroom experience in prior year	−0.01	*−0.21*	*0.12*	0.00	−0.16
School experience in prior year	*−0.06*	−0.01	*−0.07*	−0.01	0.08
School percent Black	*−0.09*	*0.15*	*−0.07*	0.03	0.41
School percent Asian	0.00	*−0.08*	*0.06*	0.04	0.08
School percent Hispanic	*0.12*	*−0.10*	−0.01	*−0.09*	0.43

Variable	Raw 6–20 versus 1–5	Raw 21+ versus 6–20	ATT weighted 6–20 versus 1–5	ATT weighted 21+ versus 6–20	Unweighted combined mean
School percent White	*−0.05*	−0.04	*0.11*	*0.07*	0.06
School percent ELL	*0.15*	0.01	*0.09*	0.01	0.14
School percent SWD	*−0.09*	−0.01	0.00	0.00	0.24
School percent poverty	*0.10*	*0.07*	−0.04	*−0.06*	0.79
School absences	−0.04	*0.18*	−0.01	0.01	15.48

Note. Regression adjustment was used for variables in the ATT weighted column in italics. Cells with values over the cutoff of 0.25 are bolded. ATT = average treatment effects on the treated; ELA = English language arts; ELL = English-language learner; OSS = out-of-school suspension; SWD = students with disabilities.

Table A.13. Balance for OSS Dose in High School

Variable	Raw 6–20 versus 1–5	Raw 21+ versus 6–20	ATT weighted 6–20 versus 1–5	ATT weighted 21+ versus 6–20	Unweighted combined mean
Level 1					
Level 2					
Level 3	0.03	**−0.66**	−0.02	−0.03	0.03
Level 4	*−0.21*	**0.29**	*−0.12*	*−0.10*	0.58
Level 5	*0.21*	*−0.22*	*0.13*	*0.10*	0.38
Overage 1+ years	−0.03	**0.28**	−0.02	*0.05*	0.60
Overage 2+ years	*−0.09*	**0.26**	−0.02	*0.06*	0.26
Weapons	**0.57**	**−1.27**	*0.15*	0.04	0.19
SWD	**−0.26**	0.02	−0.01	−0.03	0.32
ELL	*0.13*	*−0.13*	*0.09*	−0.02	0.08
Poverty	0.00	*0.05*	0.00	−0.01	0.78
Total absences in the prior year	*−0.13*	**0.34**	0.02	*0.05*	29.98
Prior year ELA z score	*0.12*	*−0.12*	0.02	−0.01	−0.53
Prior year math z score	*0.13*	*−0.14*	−0.01	−0.01	−0.61
gpa_gpa_m1_hs	*0.13*	*−0.18*	*0.09*	0.00	0.41
Black	*−0.15*	*0.23*	−0.04	0.05	0.57
White	*0.11*	**−0.32**	*0.10*	0.02	0.05
Hispanic	*0.12*	*−0.14*	0.01	*−0.05*	0.34
Other race	*0.09*	**−0.27**	*0.06*	−0.01	0.04
Total incidents in prior 90 days	*−0.11*	**0.29**	0.03	*0.07*	0.98
Total incidents of Level 4 or 5 in prior 90 days	*−0.19*	**0.30**	−0.01	*0.07*	0.37
Total suspensions in prior 90 days	*−0.14*	**0.31**	0.02	*0.07*	0.44
School-level average ELA credits in prior year	*0.14*	*−0.13*	−0.04	−0.01	0.83
School-level average math credits in prior year	*0.17*	*−0.07*	−0.05	−0.02	0.78
Classroom experience in prior year	−0.03	*−0.11*	−0.05	−0.03	0.05
School experience in prior year	−0.03	*−0.10*	−0.03	−0.03	−0.01
School-level ELA *z* score	*0.14*	*−0.18*	*0.11*	−0.02	−0.21
School-level math *z* score	*0.12*	−0.04	*0.05*	0.03	−0.31
School percent Black	*−0.16*	*0.12*	*−0.06*	0.01	0.42
School percent Asian	*0.15*	*−0.09*	*0.13*	0.03	0.08

Variable	Raw 6–20 versus 1–5	Raw 21+ versus 6–20	ATT weighted 6–20 versus 1–5	ATT weighted 21+ versus 6–20	Unweighted combined mean
School percent Hispanic	*0.06*	−0.04	*−0.06*	−0.04	0.43
School percent White	*0.08*	*−0.09*	*0.13*	0.02	0.06
School percent ELL	0.03	0.02	0.03	0.01	0.12
School percent SWD	**−0.34**	*0.07*	−0.01	0.00	0.21
School percent poverty	−0.05	*0.10*	*−0.08*	0.00	0.76
School absences	*−0.19*	*0.17*	0.00	0.02	28.32

Note. Regression adjustment was used for variables in the ATT weighted column in italics. Cells with values over the cutoff of 0.25 are bolded. ATT = average treatment effects on the treated; ELA = English language arts; ELL = English-language learner; OSS = out-of-school suspension; SWD = students with disabilities.

Table A.14. Balance for ISS Dose in Middle School

Variable	Raw 2–3 versus 1	Raw 4–5 versus 2–3	ATT weighted 2–3 versus 1	ATT weighted 4–5 versus 2–3	Unweighted combined mean
Level 1	−1.03	−0.75			0.00
Level 2	*−0.24*	*−0.19*	*0.15*	*0.05*	0.02
Level 3	*−0.25*	−0.27	*−0.11*	*−0.06*	0.41
Level 4	0.27	0.26	*0.10*	*0.06*	0.56
Level 5	0.65	0.77			0.00
Overage 1+ years	*0.09*	*0.11*	*0.05*	0.02	0.40
Overage 2+ years	*0.14*	*0.12*	*0.11*	0.03	0.10
Weapons					0.00
SWD	−0.02	−0.03	0.00	−0.01	0.38
ELL	*0.08*	0.00	0.04	0.00	0.11
Poverty	−0.01	0.01	−0.01	0.00	0.82
Total absences in the prior year	0.05	*0.09*	*0.06*	0.04	17.06
Prior year ELA *z* score	*−0.08*	−0.04	−0.02	−0.01	−0.54
Prior year math *z* score	*−0.09*	*−0.07*	−0.03	−0.02	−0.59
Black	*0.05*	*0.11*	0.03	0.02	0.44
White	*−0.15*	−0.28	−0.02	−0.01	0.10
Hispanic	0.04	0.01	−0.04	−0.02	0.40
Other race	*−0.11*	*−0.16*	*0.08*	0.03	0.06
Total incidents in prior 90 days	*−0.13*	0.05	0.02	0.04	1.17
Total incidents of Level 4 or 5 in prior 90 days	−0.04	*0.06*	0.03	0.04	0.35
Total suspensions in prior 90 days	*−0.07*	*0.08*	0.02	0.03	0.38
School-level average ELA credits in prior year	0.04	*−0.08*	−0.03	−0.03	0.84
School-level average math credits in prior year	0.05	*−0.09*	−0.03	−0.04	0.84
School-level ELA *z* score	−0.03	−0.05	−0.02	−0.01	−0.15
School-level math *z* score	−0.03	*−0.06*	−0.02	−0.01	−0.18
Classroom experience in prior year	*−0.07*	−0.04	0.04	0.01	0.01
School experience in prior year	−0.03	−0.01	0.04	−0.02	0.08
School percent Black	*0.07*	*0.16*	*0.07*	*0.06*	0.31
School percent Asian	−0.02	*−0.15*	−0.03	−0.04	0.13
School percent Hispanic	0.05	*0.07*	−0.03	−0.04	0.43

Variable	Raw 2–3 versus 1	Raw 4–5 versus 2–3	ATT weighted 2–3 versus 1	ATT weighted 4–5 versus 2–3	Unweighted combined mean
School percent White	*−0.15*	*−0.22*	−0.04	0.00	0.12
School percent ELL	*0.11*	*0.06*	*0.05*	0.01	0.13
School percent SWD	*−0.07*	0.01	0.04	0.03	0.22
School percent poverty	−0.01	*0.05*	−0.01	0.00	0.77
School absences	0.04	*0.12*	0.05	0.04	13.24

Note. Regression adjustment was used for variables in the ATT weighted column in italics. Cells with values over the cutoff of 0.25 are bolded. ATT = average treatment effects on the treated; ELA = English language arts; ELL = English-language learner; ISS = in-school suspension; SWD = students with disabilities.

Table A.15. Balance for ISS Dose in High School

Variable	Raw 2–3 versus 1	Raw 4–5 versus 2–3	ATT weighted 2–3 versus 1	ATT weighted 4–5 versus 2–3	Unweighted combined mean
Level 1					0.00
Level 2	**−0.32**	**−0.51**	0.04	0.01	0.03
Level 3	**−0.31**	**−0.42**	*−0.10*	*−0.09*	0.50
Level 4	**0.40**	**0.47**	*0.10*	*0.09*	0.47
Level 5					0.00
Overage 1+ years	*0.13*	*0.14*	0.00	0.01	0.55
Overage 2+ years	*0.13*	*0.16*	*0.05*	0.03	0.23
Weapons					0.00
SWD	0.03	0.00	−0.01	0.00	0.31
ELL	0.02	−0.04	*0.13*	*0.06*	0.10
Poverty	0.04	0.05	−0.02	−0.01	0.78
Total absences in the prior year	*0.07*	*0.16*	*0.07*	0.05	25.09
ela_scale_scoreZ8	*−0.06*	*−0.05*	−0.01	−0.02	−0.50
math_scale_scoreZ8	*−0.08*	*−0.06*	−0.01	−0.01	−0.58
gpa_gpa_m1_hs	0.02	−0.01	*0.10*	0.04	0.47
Black	*0.17*	*0.12*	−0.04	−0.01	0.53
White	*−0.17*	**−0.27**	−0.04	0.01	0.07
Hispanic	*−0.10*	−0.04	0.01	0.00	0.36
Other race	*−0.12*	*−0.10*	*0.22*	0.04	0.04
Total incidents in prior 90 days	0.00	*0.05*	*0.07*	0.04	0.85
Total incidents of Level 4 or 5 in prior 90 days	0.01	*0.11*	0.05	0.04	0.21
Total suspensions in prior 90 days	*−0.07*	*0.07*	0.04	0.01	0.34
School-level average ELA credits in prior year	*−0.10*	*−0.08*	−0.03	−0.01	0.85
School-level average math credits in prior year	*−0.08*	*−0.09*	−0.04	−0.03	0.80
Classroom experience in prior year	*−0.13*	*−0.12*	*−0.06*	0.00	0.05
School experience in prior year	*−0.10*	−0.03	*0.05*	0.03	0.00
School-level ELA *z* score	*−0.09*	*−0.12*	*−0.10*	−0.04	−0.20
School-level math *z* score	0.00	*−0.09*	0.02	−0.01	−0.28
School percent Black	*0.10*	*0.12*	0.01	0.00	0.39
School percent Asian	*0.05*	−0.04	*0.09*	−0.01	0.09

Variable	Raw 2–3 versus 1	Raw 4–5 versus 2–3	ATT weighted 2–3 versus 1	ATT weighted 4–5 versus 2–3	Unweighted combined mean
School percent Hispanic	*−0.08*	−0.02	−0.03	0.00	0.42
School percent White	*−0.10*	*−0.14*	−0.03	0.01	0.09
School percent ELL	0.05	0.03	*0.08*	0.04	0.12
School percent SWD	−0.01	−0.02	−0.01	0.00	0.20
School percent poverty	*0.07*	*0.13*	−0.01	−0.01	0.76
School absences	*0.10*	*0.11*	0.03	0.01	26.66

Note. Regression adjustment was used for variables in the ATT weighted column in italics. Cells with values over the cutoff of 0.25 are bolded. ATT = average treatment effects on the treated; ELA = English language arts; ELL = English-language learner; ISS = in-school suspension; SWD = students with disabilities.

Table A.16. Effect of Discipline Type and Length on Middle School Students' Later Behavioral Incidents

		Total incidents				Incidents, Level 4+			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	457,347	379,766	305,827	227,986	456,446	378,720	304,625	226,143
	Estimate	0.041	0.076*	0.043*	0.008	0.024*	0.026*	0.012*	0
	SD	0.024	0.022	0.016	0.019	0.008	0.008	0.005	0.007
OSS versus ISS	*N*	123,791	106,183	88,431	67,891	123,595	105,894	88,063	67,315
	Estimate	0.232*	0.250*	0.136*	0.124*	0.122*	0.127*	0.101*	0.051*
	SD	0.05	0.049	0.055	0.042	0.022	0.023	0.02	0.018
ISS 2–3 versus 1	*N*	48,110	39,908	31,921	23,095	48,057	39,825	31,796	22,911
	Estimate	−0.103	−0.057	0.032	−0.108*	−0.025	0.014	0.007	−0.023
	SD	0.085	0.066	0.068	0.053	0.031	0.023	0.022	0.026
ISS 4–5 versus 2–3	*N*	55,378	45,804	36,602	26,199	55,314	45,697	36,445	25,958
	Estimate	0.078	0.056	0.097*	0.059	0.061*	0.021	0.027	0.028
	SD	0.061	0.04	0.042	0.035	0.019	0.014	0.016	0.015
OSS 6–20 versus OSS 1–5	*N*	14,385	11,862	9,544	6,976	14,344	11,820	9,488	6,896
	Estimate	−0.186	−0.117	0.036	0.004	−0.039	−0.042	−0.024	0.045
	SD	0.133	0.132	0.103	0.09	0.065	0.042	0.044	0.041
OSS 21+ versus OSS 6–20	*N*	24,023	19,849	15,779	11,267	23,949	19,746	15,662	11,118
	Estimate	0.274*	0.231*	0.056	0.125	0.045	0.097*	0.008	0.022
	SD	0.065	0.076	0.071	0.068	0.032	0.029	0.034	0.029

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.17. Effect of Discipline Type and Length on High School Students' Later Behavioral Incidents

		Total incidents			Incidents, Level 4+		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	420,362	241,867	104,801	415,850	237,798	102,014
	Estimate	−0.060*	−0.035*	−0.042*	−0.006	−0.003	−0.008
	SD	0.018	0.017	0.021	0.006	0.006	0.008
OSS versus ISS	*N*	143,234	84,316	37,203	142,031	82,908	36,190
	Estimate	0.023	0.026	0.119*	0.029	0.036*	0.047*
	SD	0.035	0.037	0.053	0.015	0.016	0.022
ISS 2–3 versus 1	*N*	56,445	33,279	14,643	56,024	32,725	14,258
	Estimate	0.031	0.089	0.067	0.01	0.022	−0.001
	SD	0.06	0.052	0.057	0.024	0.02	0.024
ISS 4–5 versus 2–3	*N*	85,223	48,099	20,337	84,532	47,236	19,707
	Estimate	0.009	−0.005	0.02	0.019	0.011	0.029
	SD	0.027	0.028	0.033	0.011	0.011	0.016
OSS 6–20 versus OSS 1–5	*N*	10,759	5,935	2,470	10,671	5,838	2,404
	Estimate	0.007	0.026	−0.025	−0.049	−0.035	−0.001
	SD	0.099	0.103	0.125	0.038	0.047	0.047
OSS 21+ versus OSS 6–20	*N*	16,749	9,098	3,654	16,609	8,940	3,524
	Estimate	0.016	−0.095	−0.043	0.031	−0.02	−0.05
	SD	0.082	0.067	0.129	0.03	0.029	0.044

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.18. Effect of Discipline Type and Length on Middle School Students' Later Number and Length of Suspensions

		Total suspensions				Total days of suspension			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,786	378,035	303,757	225,297	455,786	378,035	303,757	225,295
	Estimate	0.069*	0.037*	0.017*	−0.001	0.682*	0.498*	0.280*	0.065
	SD	0.009	0.009	0.006	0.007	0.164	0.144	0.118	0.156
OSS versus ISS	*N*	123,455	105,668	87,825	67,058	123,455	105,668	87,825	67,056
	Estimate	0.073*	0.105*	0.071*	0.034*	3.476*	2.647*	1.657*	1.339*
	SD	0.02	0.021	0.031	0.017	0.869	0.381	0.466	0.339
ISS 2–3 versus 1	*N*	48,002	39,762	31,748	22,804	48,002	39,762	31,748	22,804
	Estimate	−0.003	0.014	0.015	−0.013	−0.299	0.189	−0.159	−0.188
	SD	0.031	0.022	0.022	0.019	0.525	0.402	0.283	0.386
ISS 4–5 versus 2–3	*N*	55,255	45,626	36,373	25,840	55,255	45,626	36,373	25,840
	Estimate	0.038*	0.017	0.02	0.015	1.215*	0.531	0.713*	0.457
	SD	0.018	0.015	0.016	0.013	0.338	0.294	0.231	0.246
OSS 6–20 versus OSS 1–5	*N*	14,327	11,770	9,445	6,862	14,327	11,770	9,445	6,861
	Estimate	−0.014	−0.019	−0.004	0.032	−1.45	−0.998	0.608	1.554*
	SD	0.051	0.039	0.047	0.031	1.758	0.976	1.028	0.759
OSS 21+ versus OSS 6–20	*N*	23,899	19,653	15,580	11,032	23,899	19,653	15,580	11,030
	Estimate	0.059*	0.062	0.006	0.025	3.281*	1.316	0.525	0.062
	SD	0.026	0.034	0.035	0.026	0.786	0.796	0.676	0.676

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.19. Effect of Discipline Type and Length on High School Students' Later Number and Length of Suspensions

		Total suspensions			Total days of suspension		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	414,870	236,394	100,956	414,867	236,393	100,955
	Estimate	0.024*	0.005	−0.02	0.323*	−0.077	−0.14
	SD	0.009	0.008	0.014	0.113	0.112	0.13
OSS versus ISS	*N*	141,840	82,433	35,733	141,839	82,432	35,732
	Estimate	−0.01	0.024	0.059*	2.468*	1.533*	1.310*
	SD	0.018	0.018	0.024	0.456	0.337	0.381
ISS 2–3 versus 1	*N*	55,957	32,569	14,093	55,957	32,568	14,092
	Estimate	0.019	0.018	0.007	−0.19	0.818*	0.397
	SD	0.028	0.027	0.027	0.727	0.338	0.439
ISS 4–5 versus 2–3	*N*	84,354	46,892	19,393	84,354	46,891	19,392
	Estimate	0.009	0.004	0.021	0.981*	0.22	0.23
	SD	0.013	0.013	0.013	0.256	0.218	0.222
OSS 6–20 versus OSS 1–5	*N*	10,603	5,780	2,377	10,603	5,780	2,377
	Estimate	0.018	0.039	0.038	0.083	−0.129	0.697
	SD	0.043	0.045	0.036	0.709	0.642	0.974
OSS 21+ versus OSS 6–20	*N*	16,526	8,852	3,492	16,525	8,852	3,492
	Estimate	−0.003	−0.029	−0.049	3.017*	1.382*	−1.33
	SD	0.031	0.033	0.045	0.618	0.603	1.021

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.20. Effect of Discipline Type and Length on Middle School Students' Later Absences

		Total absences			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,303	377,109	302,477	223,509
	Estimate	1.028*	1.680*	1.727*	1.089*
	SD	0.219	0.26	0.408	0.384
OSS versus ISS	*N*	123,311	105,398	87,452	66,494
	Estimate	5.877*	5.325*	5.635*	5.556*
	SD	0.504	0.731	0.793	1.036
ISS 2–3 versus 1	*N*	47,962	39,673	31,637	22,620
	Estimate	0.881	1.491*	3.066*	3.050*
	SD	0.463	0.516	0.937	1.295
ISS 4–5 versus 2–3	*N*	55,189	45,508	36,225	25,623
	Estimate	2.192*	3.092*	2.905*	3.752*
	SD	0.392	0.435	0.639	0.782
OSS 6–20 versus OSS 1–5	*N*	14,298	11,718	9,400	6,785
	Estimate	−0.871	−0.575	−0.56	1.246
	SD	1.043	1.293	1.737	2.109
OSS 21+ versus OSS 6–20	*N*	23,842	19,562	15,478	10,889
	Estimate	6.082*	6.860*	5.911*	5.265*
	SD	0.522	0.806	1.072	1.496

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.21. Effect of Discipline Type and Length on High School Students' Later Absences

		Total absences		
		1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	410,724	232,463	98,289
	Estimate	1.015*	1.199*	1.768*
	SD	0.379	0.572	0.727
OSS versus ISS	*N*	140,254	80,846	34,605
	Estimate	6.118*	6.940*	8.438*
	SD	0.598	0.896	1.803
ISS 2–3 versus 1	*N*	55,378	32,006	13,732
	Estimate	2.829*	4.118*	4.225*
	SD	1.112	1.538	1.924
ISS 4–5 versus 2–3	*N*	83,392	45,968	18,795
	Estimate	5.466*	5.662*	6.925*
	SD	0.533	0.836	1.46
OSS 6–20 versus OSS 1–5	*N*	10,498	5,669	2,301
	Estimate	1.129	1.396	1.604
	SD	1.535	2.252	4.301
OSS 21+ versus OSS 6–20	*N*	16,342	8,643	3,354
	Estimate	6.469*	8.797*	10.344*
	SD	1.261	1.742	3.116

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.22. Effect of Discipline Type and Length on Middle School Students' Later ELA and Math Standardized Test Scores

		ELA test score (*z* score)		Math test score (*z* score)	
		1 year after incident	2 years after incident	1 year after incident	2 years after incident
Exclusion versus no exclusion	*N*	271,520	105,494	261,131	97,988
	Estimate	0.003	0.002	−0.002	−0.007
	SD	0.007	0.01	0.006	0.011
OSS versus ISS	*N*	73,356	28,973	71,007	27,416
	Estimate	−0.033*	−0.04	−0.035*	−0.045*
	SD	0.011	0.026	0.014	0.019
ISS 2–3 versus 1	*N*	29,542	11,620	28,331	10,635
	Estimate	0.004	−0.021	0.039*	0.015
	SD	0.015	0.027	0.018	0.025
ISS 4–5 versus 2–3	*N*	33,059	12,577	31,699	11,548
	Estimate	0	−0.002	0.009	−0.04
	SD	0.015	0.021	0.011	0.022
OSS 6–20 versus OSS 1–5	*N*	7,841	2,799	7,542	2,655
	Estimate	0.042	0.034	−0.01	−0.029
	SD	0.029	0.045	0.031	0.046
OSS 21+ versus OSS 6–20	*N*	12,615	4,358	12,105	4,127
	Estimate	0.003	−0.041	0.011	−0.006
	SD	0.018	0.033	0.021	0.031

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.23. Effect of Discipline Type and Length on Likelihood of Middle School Students' Later ELA and Math Credit Accumulation

		Percentage of students earning ELA credit				Percentage of students earning math credit			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	147,779	246,285	273,646	203,862	147,779	246,285	273,646	203,862
	Estimate	−1.257*	−1.382*	−1.451*	−0.972*	−0.55	−1.074*	−1.451*	−0.997*
	SD	0.408	0.333	0.289	0.369	0.388	0.353	0.321	0.393
OSS versus ISS	*N*	41,428	70,216	79,173	60,779	41,428	70,216	79,173	60,779
	Estimate	−1.794	−2.024*	−4.171*	−3.637*	−2.736*	−3.114*	−3.438*	−2.158*
	SD	0.924	0.76	0.87	1.028	1.079	0.7	0.812	0.954
ISS 2–3 versus 1	*N*	15,399	26,074	29,127	20,954	15,399	26,074	29,127	20,954
	Estimate	−1.961	−1.852*	−0.589	−2.671*	−1.752	−2.132*	−3.017*	−1.748
	SD	1.081	0.844	0.909	1.132	1.315	0.957	0.88	1.027
ISS 4–5 versus 2–3	*N*	18,523	30,581	33,277	23,656	18,523	30,581	33,277	23,656
	Estimate	−1.704*	−1.377*	−1.648*	−0.88	−1.309	−1.224*	−2.407*	−1.864*
	SD	0.739	0.616	0.707	0.749	1.017	0.483	0.64	0.627
OSS 6–20 versus OSS 1–5	*N*	5,288	8,106	8,405	6,106	5,288	8,106	8,405	6,106
	Estimate	−0.054	−1.93	0.181	−1.521	2.685	0.765	−1.319	−2.208
	SD	1.808	1.765	1.735	1.695	2.082	1.563	1.607	1.672
OSS 21+ versus OSS 6–20	*N*	8,941	13,614	13,651	9,580	8,941	13,614	13,651	9,580
	Estimate	−4.837*	−4.145*	−4.055*	−2.530*	−5.189*	−4.443*	−5.361*	−3.483*
	SD	1.25	0.999	0.968	1.144	1.227	1.015	1.022	1.191

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.24. Effect of Discipline Type and Length on Likelihood of High School Students' Later ELA and Math Credit Accumulation

		Percentage of students earning ELA credit			Percentage of students earning math credit		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	378,117	205,514	81,557	378,117	205,514	81,557
	Estimate	−0.687*	−0.769	−0.82	−0.529	−0.569	−0.614
	SD	0.302	0.387	0.513	0.274	0.488	0.536
OSS versus ISS	*N*	130,584	72,437	29,477	130,584	72,437	29,477
	Estimate	−2.743*	−4.396*	−4.209*	−3.112*	−4.678*	−1.208
	SD	0.747	0.879	1.483	0.683	0.891	1.494
ISS 2–3 versus 1	*N*	52,195	29,033	11,866	52,195	29,033	11,866
	Estimate	−3.165*	−3.092*	−2.07	−5.097*	−2.693*	−1.848
	SD	0.958	1.172	1.83	1.041	1.015	2.047
ISS 4–5 versus 2–3	*N*	77,720	41,207	15,926	77,720	41,207	15,926
	Estimate	−2.899*	−2.860*	−3.351*	−2.692*	−1.931*	−3.507*
	SD	0.49	0.581	0.823	0.466	0.568	1.003
OSS 6–20 versus OSS 1–5	*N*	9,689	5,014	1,917	9,689	5,014	1,917
	Estimate	−4.729*	−2.632	−5.471	−1.995	−2.837	−1.726
	SD	1.612	2.069	3.751	1.741	2.288	4.372
OSS 21+ versus OSS 6–20	*N*	14,911	7,521	2,710	14,911	7,521	2,710
	Estimate	−4.771*	−7.378*	−4.606	−6.237*	−6.028*	−5.087
	SD	0.927	1.624	2.713	1.12	1.604	2.808

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.25. Effect of Discipline Type and Length on Middle School Students' Likelihood of On-Time High School Graduation

		Percentage of students graduating on time
Exclusion versus no exclusion	*N*	200,653
	Estimate	−0.897*
	SD	0.341
OSS versus ISS	*N*	62,414
	Estimate	−3.385*
	SD	0.784
ISS 2–3 versus 1	*N*	17,280
	Estimate	−1.384
	SD	1.337
ISS 4–5 versus 2–3	*N*	20,029
	Estimate	−2.336*
	SD	0.775
OSS 6–20 versus OSS 1–5	*N*	6,713
	Estimate	−0.85
	SD	1.6
OSS 21+ versus OSS 6–20	N	11,138
	Estimate	−4.512*
	SD	0.952

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.26. Effect of Discipline Type and Length on High School Students' Likelihood of On-Time High School Graduation

		Percentage of students graduating on time		
		Overall	9th or 10th grade during year of incident	11th or 12th grade during year of incident
Exclusion versus no exclusion	*N*	435,138	281,211	153,927
	Estimate	−0.353	−0.857*	−0.31
	SD	0.333	0.327	0.568
OSS versus ISS	*N*	145,480	99,713	45,767
	Estimate	−3.338*	−3.661*	−3.241*
	SD	0.587	0.517	1.283
ISS 2–3 versus 1	*N*	56,029	37,152	18,877
	Estimate	−3.230*	−3.954*	−2.695
	SD	0.933	0.924	1.547
ISS 4–5 versus 2–3	*N*	84,238	56,549	27,689
	Estimate	−2.469*	−2.759*	−2.699*
	SD	0.411	0.478	0.682
OSS 6–20 versus OSS 1–5	*N*	10,994	7,025	3,969
	Estimate	−2.875*	−2.132	−3.882
	SD	1.409	1.561	2.074
OSS 21+ versus OSS 6–20	*N*	16,629	11,426	5,203
	Estimate	−5.229*	−4.190*	−7.378*
	SD	0.911	0.836	2.064

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.27. Effect of Discipline Type and Length on Middle School Students' Same-School Same-Grade Peers' Credit Accumulation and Absences During the Year of the Incident

		Same-school same-grade peers' educational outcomes during the year of the incident		
		Likelihood of peers earning ELA credit	Likelihood of peers earning math credit	Average days absent of peers
Exclusion versus no exclusion	*N*	487,047	485,985	488,794
	Estimate	0.006	0.011*	−0.087
	SD	0.004	0.005	0.05
OSS versus ISS	*N*	132,000	131,814	132,097
	Estimate	−0.022*	−0.013*	0.209*
	SD	0.005	0.006	0.081
ISS 2–3 versus 1	*N*	51,327	51,252	51,358
	Estimate	0.009	0.014	−0.041
	SD	0.009	0.013	0.134
ISS 4–5 versus 2–3	*N*	59,095	58,992	59,125
	Estimate	0.002	0.005	0.061
	SD	0.006	0.008	0.091
OSS 6–20 versus OSS 1–5	*N*	15,616	15,576	15,629
	Estimate	0.009	0.016	−0.188
	SD	0.007	0.009	0.156
OSS 21+ versus OSS 6–20	*N*	25,864	25,806	25,879
	Estimate	−0.002	0.007	0.103
	SD	0.005	0.006	0.077

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.28. Effect of Discipline Type and Length on High School Students' Same-School Same-Grade Peers' Credit Accumulation and Absences During the Year of the Incident

		Same-school same-grade peers' educational outcomes during the year of the incident		
		Likelihood of peers earning ELA credit	Likelihood of peers earning math credit	Average days absent of peers
Exclusion versus no exclusion	*N*	570,642	570,642	571,698
	Estimate	0.003*	0.002	−0.22
	SD	0.001	0.001	0.133
OSS versus ISS	*N*	172,869	172,869	172,900
	Estimate	0	0	0.202
	SD	0.002	0.002	0.249
ISS 2–3 versus 1	*N*	67,820	67,820	67,837
	Estimate	−0.003	−0.006*	0.47
	SD	0.003	0.003	0.328
ISS 4–5 versus 2–3	*N*	103,106	103,106	103,123
	Estimate	−0.001	−0.002	0.118
	SD	0.001	0.002	0.188
OSS 6–20 versus OSS 1–5	*N*	13,365	13,365	13,370
	Estimate	−0.006	−0.007	0.289
	SD	0.004	0.004	0.384
OSS 21+ versus OSS 6–20	*N*	20,390	20,390	20,393
	Estimate	−0.004	−0.003	0.222
	SD	0.003	0.002	0.319

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.29. Effect of Discipline Type and Length on Middle School Students' and Teachers' Perceptions of School Climate During the Year of the Incident

		Standardized differences in perception of school climate measures		
		Students' perception of classroom climate	Students' perception of schoolwide climate	Teachers' perception of school climate
Exclusion versus no exclusion	*N*	126,879	126,879	82,424
	Estimate	0.016	0.042*	0.014*
	SD	0.01	0.01	0.004
OSS versus ISS	*N*	26,976	26,976	16,797
	Estimate	-0.004	-0.085*	-0.029*
	SD	0.021	0.029	0.011
ISS 2–3 versus 1	*N*	15,453	15,453	9,612
	Estimate	0.083*	0.047	0.025
	SD	0.032	0.033	0.015
ISS 4–5 versus 2–3	*N*	17,534	17,534	10,887
	Estimate	0.02	-0.002	0.005
	SD	0.022	0.019	0.01
OSS 6–20 versus OSS 1–5	*N*	3,774	3,774	2,462
	Estimate	0.033	-0.023	-0.022
	SD	0.033	0.033	0.015
OSS 21+ versus OSS 6–20	*N*	6,314	6,314	4,079
	Estimate	0.007	0.008	0.005
	SD	0.02	0.027	0.011

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table A.30. Effect of Discipline Type and Length on High School Students' and Teachers' Perceptions of School Climate During the Year of the Incident

		Standardized differences in perception of school climate measures		
		Students' perception of classroom climate	Students' perception of schoolwide climate	Teachers' perception of school climate
Exclusion versus no exclusion	*N*	214,103	214,103	160,535
	Estimate	0.025*	0.031*	0.014*
	SD	0.011	0.01	0.004
OSS versus ISS	*N*	45,755	45,755	27,507
	Estimate	-0.036	-0.035	0.007
	SD	0.026	0.023	0.008
ISS 2–3 versus 1	*N*	22,660	22,660	12,996
	Estimate	0.017	-0.028	0.008
	SD	0.023	0.021	0.013
ISS 4–5 versus 2–3	*N*	37,007	37,007	22,244
	Estimate	0.007	-0.025	0.005
	SD	0.01	0.014	0.007
OSS 6–20 versus OSS 1–5	*N*	4,420	4,420	3,038
	Estimate	0.019	-0.134*	0.003
	SD	0.054	0.054	0.016
OSS 21+ versus OSS 6–20	*N*	6,865	6,865	4,620
	Estimate	-0.017	-0.017	0.011
	SD	0.029	0.035	0.01

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Appendix B: Student Subgroup Analyses

Table B.1. Effect of Discipline Type and Length on Middle School Students' Later Behavioral Incidents, by Race

		Total incidents (all)				Total incidents (Black)				Total incidents (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	457,347	379,766	305,827	227,986	217,159	183,739	150,193	112,560	171,467	141,415	112,156	82,122
	Estimate	0.041	0.076*	0.043*	0.008	−0.01	0.073*	0.049*	0.006	−0.083*	−0.009	0.012	0.017
	SD	0.024	0.022	0.016	0.019	0.036	0.028	0.02	0.02	0.031	0.022	0.022	0.028
OSS versus ISS	*N*	123,791	106,183	88,431	67,891	58,165	50,495	42,496	32,634	48,705	41,551	34,107	25,855
	Estimate	0.232*	0.250*	0.136*	0.124*	0.139*	0.229*	0.136*	0.081	0.203*	0.215*	0.105	0.140*
	SD	0.05	0.049	0.055	0.042	0.064	0.057	0.059	0.053	0.075	0.068	0.06	0.056
ISS 2–3 versus 1	*N*	48,110	39,908	31,921	23,095	20,389	17,337	14,175	10,342	19,372	15,968	12,604	8,953
	Estimate	−0.103	−0.057	0.032	−0.108*	−0.268*	−0.164	0.044	−0.038	−0.017	−0.059	−0.063	−0.142
	SD	0.085	0.066	0.068	0.053	0.123	0.087	0.093	0.076	0.07	0.08	0.096	0.09
ISS 4–5 versus 2–3	*N*	55,378	45,804	36,602	26,199	24,632	20,831	16,972	12,244	22,500	18,414	14,459	10,176
	Estimate	0.078	0.056	0.097*	0.059	−0.024	0.008	0.132*	0.038	0.01	0.047	−0.041	0.006
	SD	0.061	0.04	0.042	0.035	0.09	0.044	0.051	0.059	0.063	0.066	0.058	0.059
OSS 6–20 versus OSS 1–5	*N*	14,385	11,862	9,544	6,976	7,898	6,531	5,300	3,856	5,373	4,432	3,507	2,588
	Estimate	−0.186	−0.117	0.036	0.004	−0.107	0.118	0.199	0.092	−0.183	−0.28	−0.062	−0.032
	SD	0.133	0.132	0.103	0.09	0.209	0.176	0.12	0.115	0.174	0.218	0.142	0.211
OSS 21+ versus OSS 6–20	*N*	24,023	19,849	15,779	11,267	13,826	11,510	9,243	6,629	8,637	7,090	5,518	3,916
	Estimate	0.274*	0.231*	0.056	0.125	0.247*	0.14	−0.038	0.141	0.246*	0.409*	0.084	−0.019
	SD	0.065	0.076	0.071	0.068	0.097	0.104	0.165	0.09	0.103	0.108	0.099	0.143

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.2. Effect of Discipline Type and Length on Middle School Students' Later Behavioral Incidents, by Special Status

		Total incidents (all)				Total incidents (SWD)				Total incidents (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	457,347	379,766	305,827	227,986	183,937	150,448	119,304	87,013	373,053	311,228	251,020	186,274
	Estimate	0.041	0.076*	0.043*	0.008	−0.070*	0.036	−0.005	−0.011	−0.004	0.053*	0.033	0.014
	SD	0.024	0.022	0.016	0.019	0.031	0.028	0.026	0.025	0.023	0.018	0.017	0.022
OSS versus ISS	*N*	123,791	106,183	88,431	67,891	46,896	39,553	32,286	24,069	100,816	86,654	72,010	54,839
	Estimate	0.232*	0.250*	0.136*	0.124*	0.199*	0.272*	0.096	0.125*	0.171*	0.214*	0.086	0.111*
	SD	0.05	0.049	0.055	0.042	0.088	0.076	0.068	0.047	0.059	0.061	0.064	0.045
ISS 2–3 versus 1	*N*	48,110	39,908	31,921	23,095	18,684	15,275	11,995	8,340	39,526	32,993	26,501	19,191
	Estimate	−0.103	−0.057	0.032	−0.108*	−0.119	−0.1	−0.008	−0.250*	−0.138	−0.125	0.036	−0.083
	SD	0.085	0.066	0.068	0.053	0.102	0.115	0.095	0.086	0.079	0.068	0.064	0.059
ISS 4–5 versus 2–3	*N*	55,378	45,804	36,602	26,199	21,187	17,264	13,522	9,289	45,519	37,851	30,279	21,615
	Estimate	0.078	0.056	0.097*	0.059	0.037	−0.033	−0.016	0.054	0.062	0.083	0.057	0.053
	SD	0.061	0.04	0.042	0.035	0.094	0.072	0.086	0.06	0.058	0.044	0.048	0.037
OSS 6–20 versus OSS 1–5	*N*	14,385	11,862	9,544	6,976	6,424	5,217	4,145	2,957	12,065	9,957	7,993	5,794
	Estimate	−0.186	−0.117	0.036	0.004	−0.439	−0.052	−0.204	0.043	−0.222	−0.001	0.034	0.023
	SD	0.133	0.132	0.103	0.09	0.236	0.166	0.151	0.162	0.143	0.111	0.123	0.114
OSS 21+ versus OSS 6–20	*N*	24,023	19,849	15,779	11,267	10,042	8,124	6,340	4,462	20,462	16,911	13,404	9,502
	Estimate	0.274*	0.231*	0.056	0.125	0.223*	0.220*	0.218*	0.208*	0.269*	0.244*	0.04	0.142*
	SD	0.065	0.076	0.071	0.068	0.105	0.092	0.094	0.092	0.075	0.073	0.083	0.069

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.3. Effect of Discipline Type and Length on Middle School Students' Later Serious Behavioral Incidents, by Race

		Total incidents (all)				Total incidents (Black)				Total incidents (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	456,446	378,720	304,625	226,143	216,642	183,084	149,433	111,367	171,151	141,139	111,773	81,535
	Estimate	0.024*	0.026*	0.012*	0	0.006	0.019	0.016*	0.01	−0.011	0.01	−0.002	−0.005
	SD	0.008	0.008	0.005	0.007	0.01	0.011	0.007	0.008	0.011	0.008	0.009	0.007
OSS versus ISS	*N*	123,595	105,894	88,063	67,315	58,058	50,315	42,281	32,269	48,626	41,464	33,971	25,673
	Estimate	0.122*	0.127*	0.101*	0.051*	0.088*	0.130*	0.093*	0.043	0.133*	0.097*	0.082*	0.046
	SD	0.022	0.023	0.02	0.018	0.029	0.025	0.027	0.022	0.027	0.028	0.023	0.025
ISS 2–3 versus 1	*N*	48,057	39,825	31,796	22,911	20,364	17,287	14,099	10,231	19,347	15,947	12,563	8,891
	Estimate	−0.025	0.014	0.007	−0.023	−0.116*	−0.039	0.005	−0.013	0.004	0.013	−0.024	−0.069
	SD	0.031	0.023	0.022	0.026	0.049	0.037	0.036	0.033	0.03	0.028	0.031	0.038
ISS 4–5 versus 2–3	*N*	55,314	45,697	36,445	25,958	24,600	20,764	16,879	12,090	22,469	18,383	14,406	10,106
	Estimate	0.061*	0.021	0.027	0.028	0.036	0.012	0.044*	0.04	0.015	0.027	−0.024	−0.003
	SD	0.019	0.014	0.016	0.015	0.027	0.021	0.021	0.022	0.027	0.024	0.02	0.017
OSS 6–20 versus OSS 1–5	*N*	14,344	11,820	9,488	6,896	7,873	6,500	5,269	3,807	5,359	4,423	3,483	2,559
	Estimate	−0.039	−0.042	−0.024	0.045	−0.005	0.025	0.044	0.068	−0.022	−0.082	−0.076	0.089
	SD	0.065	0.042	0.044	0.041	0.094	0.055	0.057	0.056	0.078	0.082	0.069	0.059
OSS 21+ versus OSS 6–20	*N*	23,949	19,746	15,662	11,118	13,783	11,445	9,173	6,530	8,609	7,059	5,473	3,873
	Estimate	0.045	0.097*	0.008	0.022	0.016	0.054	0.015	0.025	0.044	0.142*	−0.021	−0.002
	SD	0.032	0.029	0.034	0.029	0.049	0.039	0.047	0.039	0.048	0.046	0.044	0.042

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.4. Effect of Discipline Type and Length on Middle School Students' Later Serious Behavioral Incidents, by Special Status

		Total incidents (all)				Total incidents (SWD)				Total incidents (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	456,446	378,720	304,625	226,143	183,355	149,914	118,813	86,280	372,303	310,392	250,099	184,690
	Estimate	0.024*	0.026*	0.012*	0	0.001	0.007	−0.011	−0.006	0.01	0.017*	0.009	−0.001
	SD	0.008	0.008	0.005	0.007	0.011	0.01	0.009	0.009	0.007	0.008	0.006	0.008
OSS versus ISS	*N*	123,595	105,894	88,063	67,315	46,766	39,425	32,138	23,852	100,648	86,420	71,734	54,364
	Estimate	0.122*	0.127*	0.101*	0.051*	0.101*	0.137*	0.080*	0.028	0.105*	0.119*	0.083*	0.042*
	SD	0.022	0.023	0.02	0.018	0.037	0.026	0.025	0.021	0.024	0.029	0.024	0.02
ISS 2–3 versus 1	*N*	48,057	39,825	31,796	22,911	18,643	15,243	11,940	8,284	39,477	32,935	26,402	19,033
	Estimate	−0.025	0.014	0.007	−0.023	−0.054	0	0.001	−0.087*	−0.044	−0.009	0.004	−0.023
	SD	0.031	0.023	0.022	0.026	0.042	0.048	0.037	0.038	0.03	0.023	0.02	0.025
ISS 4–5 versus 2–3	*N*	55,314	45,697	36,445	25,958	21,145	17,216	13,461	9,206	45,458	37,773	30,157	21,426
	Estimate	0.061*	0.021	0.027	0.028	0.053	−0.026	0.021	0.027	0.065*	0.032*	0.02	0.032*
	SD	0.019	0.014	0.016	0.015	0.031	0.024	0.024	0.028	0.019	0.016	0.018	0.014
OSS 6–20 versus OSS 1–5	*N*	14,344	11,820	9,488	6,896	6,393	5,197	4,120	2,915	12,030	9,922	7,950	5,725
	Estimate	−0.039	−0.042	−0.024	0.045	−0.036	−0.099	−0.137	0.086	−0.02	−0.019	−0.021	0.089*
	SD	0.065	0.042	0.044	0.041	0.095	0.078	0.069	0.075	0.06	0.048	0.052	0.04
OSS 21+ versus OSS 6–20	*N*	23,949	19,746	15,662	11,118	9,999	8,079	6,295	4,397	20,398	16,824	13,315	9,371
	Estimate	0.045	0.097*	0.008	0.022	0.037	0.090*	0.076	0.063	0.034	0.092*	0.012	0.024
	SD	0.032	0.029	0.034	0.029	0.049	0.045	0.044	0.036	0.037	0.031	0.031	0.028

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.5. Effect of Discipline Type and Length on Middle School Students' Later Number of Suspensions, by Race

		Total suspensions (all)				Total suspensions (Black)				Total suspensions (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,786	378,035	303,757	225,297	216,232	182,714	148,818	110,760	170,959	140,873	111,538	81,329
	Estimate	0.069*	0.037*	0.017*	−0.001	0.062*	0.034*	0.019*	0.006	0.034*	0.018	0	−0.003
	SD	0.009	0.009	0.006	0.007	0.01	0.011	0.007	0.008	0.013	0.011	0.01	0.009
OSS versus ISS	*N*	123,455	105,668	87,825	67,058	57,964	50,165	42,128	32,074	48,592	41,394	33,892	25,626
	Estimate	0.073*	0.105*	0.071*	0.034*	0.04	0.097*	0.088*	0.036	0.072*	0.104*	0.057*	0.042
	SD	0.02	0.021	0.031	0.017	0.025	0.025	0.027	0.025	0.023	0.029	0.022	0.023
ISS 2–3 versus 1	*N*	48,002	39,762	31,748	22,804	20,331	17,244	14,076	10,145	19,329	15,931	12,540	8,873
	Estimate	−0.003	0.014	0.015	−0.013	−0.075	−0.011	0.005	−0.013	0.021	−0.021	−0.03	−0.058
	SD	0.031	0.022	0.022	0.019	0.047	0.035	0.031	0.033	0.029	0.03	0.031	0.034
ISS 4–5 versus 2–3	*N*	55,255	45,626	36,373	25,840	24,561	20,714	16,845	12,012	22,454	18,366	14,370	10,069
	Estimate	0.038*	0.017	0.02	0.015	0.013	0.014	0.022	0.019	−0.002	0.002	−0.014	−0.009
	SD	0.018	0.015	0.016	0.013	0.025	0.018	0.02	0.02	0.028	0.024	0.02	0.017
OSS 6–20 versus OSS 1–5	*N*	14,327	11,770	9,445	6,862	7,858	6,464	5,238	3,782	5,359	4,409	3,471	2,553
	Estimate	−0.014	−0.019	−0.004	0.032	0.04	0.089*	0.051	0.045	0.049	−0.137	0.008	0.091
	SD	0.051	0.039	0.047	0.031	0.059	0.042	0.061	0.057	0.079	0.093	0.053	0.055
OSS 21+ versus OSS 6–20	*N*	23,899	19,653	15,580	11,032	13,747	11,385	9,105	6,465	8,601	7,028	5,460	3,859
	Estimate	0.059*	0.062	0.006	0.025	0.047	0.005	−0.022	0.011	0.066	0.122*	−0.02	0.006
	SD	0.026	0.034	0.035	0.026	0.032	0.044	0.058	0.037	0.036	0.039	0.044	0.038

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.6. Effect of Discipline Type and Length on Middle School Students' Later Number of Suspensions, by Special Status

		Total suspensions (all)				Total suspensions (SWD)				Total suspensions (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,786	378,035	303,757	225,297	182,927	149,486	118,392	85,975	371,680	309,816	249,380	184,001
	Estimate	0.069*	0.037*	0.017*	−0.001	0.054*	0.030*	−0.001	0.001	0.059*	0.029*	0.012*	−0.005
	SD	0.009	0.009	0.006	0.007	0.011	0.012	0.009	0.008	0.008	0.009	0.006	0.007
OSS versus ISS	*N*	123,455	105,668	87,825	67,058	46,700	39,303	32,014	23,772	100,521	86,222	71,528	54,157
	Estimate	0.073*	0.105*	0.071*	0.034*	0.035	0.090*	0.069*	0.03	0.048*	0.100*	0.062*	0.033
	SD	0.02	0.021	0.031	0.017	0.03	0.027	0.024	0.018	0.02	0.024	0.023	0.018
ISS 2–3 versus 1	*N*	48,002	39,762	31,748	22,804	18,613	15,208	11,916	8,246	39,430	32,875	26,372	18,947
	Estimate	−0.003	0.014	0.015	−0.013	−0.024	0.01	0.018	−0.051	−0.024	−0.018	0.009	−0.022
	SD	0.031	0.022	0.022	0.019	0.052	0.052	0.037	0.032	0.03	0.021	0.02	0.02
ISS 4–5 versus 2–3	*N*	55,255	45,626	36,373	25,840	21,117	17,174	13,420	9,167	45,409	37,708	30,113	21,336
	Estimate	0.038*	0.017	0.02	0.015	−0.001	−0.024	−0.01	0.027	0.037	0.025	0.02	0.019
	SD	0.018	0.015	0.016	0.013	0.026	0.027	0.024	0.022	0.019	0.015	0.018	0.012
OSS 6–20 versus OSS 1–5	*N*	14,327	11,770	9,445	6,862	6,387	5,162	4,098	2,902	12,012	9,880	7,913	5,702
	Estimate	−0.014	−0.019	−0.004	0.032	−0.054	−0.022	−0.094	0.095	0.001	−0.035	0.012	0.058
	SD	0.051	0.039	0.047	0.031	0.085	0.066	0.072	0.074	0.049	0.045	0.05	0.033
OSS 21+ versus OSS 6–20	*N*	23,899	19,653	15,580	11,032	9,980	8,028	6,254	4,369	20,352	16,740	13,241	9,306
	Estimate	0.059*	0.062	0.006	0.025	0.01	0.069	0.072	0.076*	0.063*	0.068	0.01	0.033
	SD	0.026	0.034	0.035	0.026	0.037	0.049	0.037	0.035	0.029	0.034	0.034	0.026

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.7. Effect of Discipline Type and Length on Middle School Students' Total Days Suspended, by Race

		Total days suspended (all)				Total days suspended (Black)				Total days suspended (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,786	378,035	303,757	225,295	216,232	182,714	148,818	110,758	170,959	140,873	111,538	81,329
	Estimate	0.682*	0.498*	0.280*	0.065	0.603*	0.668*	0.313	0.361	0.28	0.161	0.071	−0.18
	SD	0.164	0.144	0.118	0.156	0.194	0.239	0.192	0.192	0.194	0.128	0.158	0.156
OSS versus ISS	*N*	123,455	105,668	87,825	67,056	57,964	50,165	42,128	32,072	48,592	41,394	33,892	25,626
	Estimate	3.476*	2.647*	1.657*	1.339*	3.393*	2.428*	1.576*	1.229*	3.845*	2.424*	1.579*	1.036*
	SD	0.869	0.381	0.466	0.339	0.666	0.461	0.534	0.445	0.558	0.474	0.374	0.385
ISS 2–3 versus 1	*N*	48,002	39,762	31,748	22,804	20,331	17,244	14,076	10,145	19,329	15,931	12,540	8,873
	Estimate	−0.299	0.189	−0.159	−0.188	−0.951	−0.029	0.17	−0.252	0.282	0.103	−0.181	−0.724
	SD	0.525	0.402	0.283	0.386	0.996	0.723	0.626	0.623	0.4	0.528	0.414	0.471
ISS 4–5 versus 2–3	*N*	55,255	45,626	36,373	25,840	24,561	20,714	16,845	12,012	22,454	18,366	14,370	10,069
	Estimate	1.215*	0.531	0.713*	0.457	0.847	0.046	0.956*	0.766	1.057*	0.541	0.04	−0.101
	SD	0.338	0.294	0.231	0.246	0.531	0.522	0.298	0.411	0.459	0.361	0.292	0.261
OSS 6–20 versus OSS 1–5	*N*	14,327	11,770	9,445	6,861	7,858	6,464	5,238	3,781	5,359	4,409	3,471	2,553
	Estimate	−1.45	−0.998	0.608	1.554*	−1.536	−0.946	2.167	2.701*	1.437	−1.461	−0.514	0.604
	SD	1.758	0.976	1.028	0.759	2.126	1.452	1.317	1.09	1.262	1.645	0.977	0.891
OSS 21+ versus OSS 6–20	*N*	23,899	19,653	15,580	11,030	13,747	11,385	9,105	6,463	8,601	7,028	5,460	3,859
	Estimate	3.281*	1.316	0.525	0.062	3.160*	1.132	−0.206	−0.875	2.544*	1.873*	0.686	0.122
	SD	0.786	0.796	0.676	0.676	1.36	0.908	1.078	0.946	0.971	0.777	0.898	0.795

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.8. Effect of Discipline Type and Length on Middle School Students' Total Days Suspended, by Special Status

		Total suspensions (all)				Total days suspended (SWD)				Total days suspended (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,786	378,035	303,757	225,295	182,927	149,486	118,392	85,975	371,680	309,816	249,380	183,999
	Estimate	0.682*	0.498*	0.280*	0.065	0.394*	0.300*	0.033	0.108	0.523*	0.385*	0.173	0.029
	SD	0.164	0.144	0.118	0.156	0.181	0.135	0.179	0.198	0.146	0.137	0.132	0.173
OSS versus ISS	*N*	123,455	105,668	87,825	67,056	46,700	39,303	32,014	23,772	100,521	86,222	71,528	54,155
	Estimate	3.476*	2.647*	1.657*	1.339*	3.138*	2.282*	1.424*	1.334*	2.924*	2.486*	1.685*	1.543*
	SD	0.869	0.381	0.466	0.339	0.519	0.493	0.461	0.441	1.106	0.47	0.464	0.356
ISS 2–3 versus 1	*N*	48,002	39,762	31,748	22,804	18,613	15,208	11,916	8,246	39,430	32,875	26,372	18,947
	Estimate	−0.299	0.189	−0.159	−0.188	−0.585	0.09	−0.037	−1.797*	−0.481	0	−0.132	−0.089
	SD	0.525	0.402	0.283	0.386	0.621	0.681	0.516	0.72	0.558	0.429	0.399	0.378
ISS 4–5 versus 2–3	*N*	55,255	45,626	36,373	25,840	21,117	17,174	13,420	9,167	45,409	37,708	30,113	21,336
	Estimate	1.215*	0.531	0.713*	0.457	1.242*	0.009	0.577	0.371	1.442*	0.661*	0.651*	0.38
	SD	0.338	0.294	0.231	0.246	0.515	0.383	0.352	0.361	0.363	0.326	0.273	0.253
OSS 6–20 versus OSS 1–5	*N*	14,327	11,770	9,445	6,861	6,387	5,162	4,098	2,902	12,012	9,880	7,913	5,701
	Estimate	−1.45	−0.998	0.608	1.554*	−0.782	−1.271	−1.197	1.181	0.25	−1.273	0.392	2.095*
	SD	1.758	0.976	1.028	0.759	1.458	1.525	1.821	1.215	1.023	1.174	1.315	0.85
OSS 21+ versus OSS 6–20	*N*	23,899	19,653	15,580	11,030	9,980	8,028	6,254	4,369	20,352	16,740	13,241	9,304
	Estimate	3.281*	1.316	0.525	0.062	1.944	1.46	1.473	0.142	3.353*	1.598	0.565	0.202
	SD	0.786	0.796	0.676	0.676	1.23	1.6	1.011	0.958	0.997	0.844	0.784	0.726

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.9. Effect of Discipline Type and Length on Middle School Students' Later Absences, by Race

		Total absences (all)				Total absences (Black)				Total absences (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,303	377,109	302,477	223,509	215,916	182,016	147,817	109,425	170,809	140,647	111,287	80,937
	Estimate	1.028*	1.680*	1.727*	1.089*	1.109*	1.421*	2.018*	0.817	0.863*	1.584*	0.935	1.016
	SD	0.219	0.26	0.408	0.384	0.238	0.311	0.447	0.498	0.313	0.403	0.584	0.745
OSS versus ISS	*N*	123,311	105,398	87,452	66,494	57,886	49,964	41,830	31,663	48,530	41,326	33,822	25,485
	Estimate	5.877*	5.325*	5.635*	5.556*	4.904*	5.373*	5.306*	6.651*	6.649*	6.278*	5.816*	6.517*
	SD	0.504	0.731	0.793	1.036	0.622	0.67	0.777	1.023	0.55	0.683	1.079	1.343
ISS 2–3 versus 1	*N*	47,962	39,673	31,637	22,620	20,310	17,177	13,987	10,009	19,311	15,909	12,521	8,828
	Estimate	0.881	1.491*	3.066*	3.050*	0.426	1.845*	2.836	3.08	1.622*	2.056*	4.749*	1.868
	SD	0.463	0.516	0.937	1.295	0.715	0.807	1.522	1.837	0.739	0.852	1.177	1.534
ISS 4–5 versus 2–3	*N*	55,189	45,508	36,225	25,623	24,525	20,626	16,720	11,843	22,425	18,336	14,351	10,026
	Estimate	2.192*	3.092*	2.905*	3.752*	1.479*	2.984*	2.471*	2.296*	2.839*	3.466*	2.723*	4.743*
	SD	0.392	0.435	0.639	0.782	0.49	0.711	0.907	1.017	0.529	0.814	0.981	1.305
OSS 6–20 versus OSS 1–5	*N*	14,298	11,718	9,400	6,785	7,842	6,429	5,202	3,724	5,346	4,392	3,462	2,536
	Estimate	−0.871	−0.575	−0.56	1.246	−1.098	−1.976	−0.865	4.7	3.201	2.773	0.522	1.286
	SD	1.043	1.293	1.737	2.109	1.58	2.082	2.455	2.856	1.688	1.575	3.103	4.159
OSS 21+ versus OSS 6–20	*N*	23,842	19,562	15,478	10,889	13,712	11,317	9,023	6,356	8,579	7,005	5,440	3,831
	Estimate	6.082*	6.860*	5.911*	5.265*	5.318*	6.884*	5.125*	4.133*	5.592*	5.552*	6.093*	1.794
	SD	0.522	0.806	1.072	1.496	0.645	1.035	1.578	2.022	1.048	1.217	1.469	2.14

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.10. Effect of Discipline Type and Length on Middle School Students' Later Absences, by Special Status

		Total absences (all)				Total absences (SWD)				Total absences (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	455,303	377,109	302,477	223,509	182,625	149,040	117,958	85,401	371,315	309,027	248,314	182,509
	Estimate	1.028*	1.680*	1.727*	1.089*	1.119*	1.583*	1.726*	1.194*	0.800*	1.196*	1.308*	0.795*
	SD	0.219	0.26	0.408	0.384	0.293	0.405	0.547	0.585	0.216	0.252	0.391	0.397
OSS versus ISS	*N*	123,311	105,398	87,452	66,494	46,605	39,170	31,910	23,610	100,412	86,006	71,242	53,696
	Estimate	5.877*	5.325*	5.635*	5.556*	6.256*	5.460*	5.019*	6.379*	5.196*	4.878*	5.541*	5.214*
	SD	0.504	0.731	0.793	1.036	0.533	0.78	0.948	1.204	0.539	0.717	0.865	1.197
ISS 2–3 versus 1	*N*	47,962	39,673	31,637	22,620	18,589	15,157	11,882	8,193	39,398	32,800	26,274	18,787
	Estimate	0.881	1.491*	3.066*	3.050*	1.497	2.731*	5.536*	6.580*	0.853	1.657*	3.124*	3.511*
	SD	0.463	0.516	0.937	1.295	0.765	1.102	1.431	2.103	0.525	0.611	1.111	1.292
ISS 4–5 versus 2–3	*N*	55,189	45,508	36,225	25,623	21,074	17,111	13,380	9,101	45,356	37,606	29,988	21,149
	Estimate	2.192*	3.092*	2.905*	3.752*	2.203*	3.711*	1.126	0.581	2.411*	3.142*	3.257*	3.719*
	SD	0.392	0.435	0.639	0.782	0.56	0.812	1.166	1.298	0.412	0.487	0.727	0.907
OSS 6–20 versus OSS 1–5	*N*	14,298	11,718	9,400	6,785	6,364	5,139	4,086	2,881	11,993	9,843	7,883	5,638
	Estimate	−0.871	−0.575	−0.56	1.246	1.029	1.078	0.988	0.729	0.148	0.544	−0.245	2.405
	SD	1.043	1.293	1.737	2.109	1.49	2.164	2.82	4.002	1.213	1.397	1.842	2.063
OSS 21+ versus OSS 6–20	*N*	23,842	19,562	15,478	10,889	9,941	7,986	6,231	4,336	20,309	16,672	13,161	9,185
	Estimate	6.082*	6.860*	5.911*	5.265*	6.022*	7.784*	4.504*	6.132*	5.769*	6.969*	5.916*	4.878*
	SD	0.522	0.806	1.072	1.496	0.882	1.137	1.641	2.053	0.593	0.926	1.246	1.535

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.11. Effect of Discipline Type and Length on Middle School Students' Later ELA and Math Standardized Test Scores, by Race

		ELA test score (*z* score) (all)		Math test score (*z* score) (all)		ELA test score (*z* score) (Black)		Math test score (*z* score) (Black)		ELA test score (*z* score) (Hispanic)		Math test score (*z* score) (Hispanic)	
		1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident
Exclusion versus no exclusion	*N*	271,520	105,494	261,131	97,988	130,475	52,209	126,257	49,129	100,005	38,379	96,547	35,926
	Estimate	0.003	0.002	−0.002	−0.007	0.006	0.002	0.007	−0.007	0.005	0.015	0.008	−0.001
	SD	0.007	0.01	0.006	0.011	0.008	0.012	0.007	0.011	0.008	0.014	0.007	0.016
OSS versus ISS	*N*	73,356	28,973	71,007	27,416	35,103	14,500	34,035	13,822	28,274	10,938	27,412	10,321
	Estimate	−0.033*	−0.04	−0.035*	−0.045*	−0.02	−0.034	−0.025	−0.02	−0.03	−0.026	−0.03	−0.009
	SD	0.011	0.026	0.014	0.019	0.012	0.02	0.014	0.022	0.017	0.027	0.017	0.027
ISS 2–3 versus 1	*N*	29,542	11,620	28,331	10,635	12,696	5,342	12,211	4,934	11,788	4,523	11,373	4,158
	Estimate	0.004	−0.021	0.039*	0.015	0.003	−0.027	0.025	0.02	−0.003	−0.026	0.032	−0.027
	SD	0.015	0.027	0.018	0.025	0.022	0.035	0.023	0.033	0.026	0.04	0.026	0.043
ISS 4–5 versus 2–3	*N*	33,059	12,577	31,699	11,548	14,940	6,055	14,372	5,569	13,285	4,928	12,760	4,554
	Estimate	0	−0.002	0.009	−0.04	0.028	0.007	0.044*	−0.008	−0.015	0.006	0	−0.06
	SD	0.015	0.021	0.011	0.022	0.017	0.026	0.014	0.031	0.019	0.029	0.02	0.032
OSS 6–20 versus OSS 1–5	*N*	7,841	2,799	7,542	2,655	4,425	1,610	4,245	1,533	2,882	1,030	2,780	978
	Estimate	0.042	0.034	−0.01	−0.029	0.009	−0.013	−0.048	−0.034	0.025	−0.107	−0.075	−0.083
	SD	0.029	0.045	0.031	0.046	0.032	0.058	0.038	0.057	0.047	0.066	0.044	0.057
OSS 21+ versus OSS 6–20	*N*	12,615	4,358	12,105	4,127	7,521	2,658	7,225	2,531	4,356	1,499	4,161	1,407
	Estimate	0.003	−0.041	0.011	−0.006	0.069*	−0.002	0.033	0.018	−0.032	−0.022	0.029	−0.037
	SD	0.018	0.033	0.021	0.031	0.025	0.03	0.025	0.036	0.028	0.058	0.032	0.052

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.12. Effect of Discipline Type and Length on Middle School Students' Later ELA and Math Standardized Test Scores, by Special Status

		ELA test score (*z* score) (all)		Math test score (*z* score) (all)		ELA test score (*z* score) (SWD)		Math test score (*z* score) (SWD)		ELA test score (*z* score) (students from economically disadvantaged groups)		Math test score (*z* score) (students from economically disadvantaged homes)	
		1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident	1 year after incident	2 years after incident
Exclusion versus no exclusion	*N*	271,520	105,494	261,131	97,988	105,413	40,281	101,886	37,861	222,804	87,794	214,658	82,058
	Estimate	0.003	0.002	−0.002	−0.007	0.006	−0.007	0.01	−0.004	0.005	0.009	0.002	0
	SD	0.007	0.01	0.006	0.011	0.009	0.014	0.008	0.013	0.007	0.01	0.006	0.011
OSS versus ISS	*N*	73,356	28,973	71,007	27,416	27,232	10,394	26,343	9,839	60,169	24,189	58,196	22,939
	Estimate	−0.033*	−0.04	−0.035*	−0.045*	−0.006	−0.031	−0.003	−0.008	−0.030*	−0.031	−0.033*	−0.031
	SD	0.011	0.026	0.014	0.019	0.016	0.027	0.016	0.031	0.011	0.025	0.013	0.018
ISS 2–3 versus 1	*N*	29,542	11,620	28,331	10,635	11,167	4,244	10,761	3,909	24,455	9,814	23,504	9,036
	Estimate	0.004	−0.021	0.039*	0.015	0.016	−0.073	0.042	−0.028	−0.002	−0.003	0.025	0.015
	SD	0.015	0.027	0.018	0.025	0.029	0.043	0.027	0.041	0.017	0.029	0.02	0.031
ISS 4–5 versus 2–3	*N*	33,059	12,577	31,699	11,548	12,407	4,600	11,947	4,265	27,361	10,597	26,266	9,788
	Estimate	0	−0.002	0.009	−0.04	−0.001	0.04	0.009	−0.051	−0.001	0.001	0.01	−0.028
	SD	0.015	0.021	0.011	0.022	0.021	0.032	0.019	0.033	0.015	0.02	0.013	0.025
OSS 6–20 versus OSS 1–5	*N*	7,841	2,799	7,542	2,655	3,523	1,213	3,397	1,150	6,694	2,446	6,429	2,319
	Estimate	0.042	0.034	−0.01	−0.029	0.019	−0.087	−0.028	−0.08	0.019	0.02	−0.031	−0.049
	SD	0.029	0.045	0.031	0.046	0.034	0.068	0.044	0.068	0.032	0.05	0.031	0.046
OSS 21+ versus OSS 6–20	*N*	12,615	4,358	12,105	4,127	5,261	1,713	5,016	1,609	10,920	3,838	10,433	3,617
	Estimate	0.003	−0.041	0.011	−0.006	−0.01	−0.057	0.039	0.018	−0.009	−0.045	0.019	−0.006
	SD	0.018	0.033	0.021	0.031	0.03	0.047	0.031	0.052	0.023	0.033	0.023	0.038

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.13. Effect of Discipline Type and Length on Likelihood of Middle School Students' Later ELA Credit Accumulation, by Race

		Percentage of students earning ELA credit (all)				Percentage of students earning ELA credit (Black)				Percentage of students earning ELA credit (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	147,779	246,285	273,646	203,862	68,742	118,025	133,335	100,213	56,786	92,335	100,142	73,191
	Estimate	−1.257*	−1.382*	−1.451*	−0.972*	−1.579*	−1.303*	−1.394*	−0.652	−0.929	−1.078	−0.493	−0.955
	SD	0.408	0.333	0.289	0.369	0.503	0.401	0.359	0.475	0.677	0.551	0.477	0.561
OSS versus ISS	*N*	41,428	70,216	79,173	60,779	18,809	32,552	37,649	29,012	16,745	27,829	30,527	23,090
	Estimate	−1.794	−2.024*	−4.171*	−3.637*	−1.619	−2.138*	−3.330*	−3.695*	−1.739	−2.692*	−4.519*	−2.888*
	SD	0.924	0.76	0.87	1.028	1.117	0.804	0.755	0.931	1.328	0.971	1.172	1.232
ISS 2–3 versus 1	*N*	15,399	26,074	29,127	20,954	6,352	11,005	12,775	9,296	6,310	10,575	11,520	8,108
	Estimate	−1.961	−1.852*	−0.589	−2.671*	−2.651	−1.498	−0.901	−2.059	0.319	−0.716	−0.671	−2.019
	SD	1.081	0.844	0.909	1.132	2.09	1.073	1.644	1.764	1.9	1.568	1.342	1.644
ISS 4–5 versus 2–3	*N*	18,523	30,581	33,277	23,656	8,019	13,563	15,305	10,972	7,651	12,401	13,150	9,182
	Estimate	−1.704*	−1.377*	−1.648*	−0.88	−0.949	−1.633	0.116	0.369	−2.978*	−0.895	−3.397*	−1.546
	SD	0.739	0.616	0.707	0.749	1.22	0.902	0.982	1.074	1.301	1.116	1.109	1.131
OSS 6–20 versus OSS 1–5	*N*	5,288	8,106	8,405	6,106	2,804	4,415	4,644	3,380	2,002	3,010	3,090	2,252
	Estimate	−0.054	−1.93	0.181	−1.521	0.994	−3.175	0.16	−1.223	−3.121	−0.236	1.771	−4.895
	SD	1.808	1.765	1.735	1.695	2.346	2.555	2.501	2.261	3.256	2.729	3.612	3.607
OSS 21+ versus OSS 6–20	*N*	8,941	13,614	13,651	9,580	4,953	7,816	7,990	5,634	3,329	4,864	4,769	3,331
	Estimate	−4.837*	−4.145*	−4.055*	−2.530*	−5.218*	−3.209	−2.541	−2.219	−3.417*	−5.437*	−4.474*	0.866
	SD	1.25	0.999	0.968	1.144	1.914	1.612	1.53	1.665	1.543	1.527	1.964	2.248

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.14. Effect of Discipline Type and Length on Likelihood of Middle School Students' Later ELA Credit Accumulation, by Special Status

		Percentage of students earning ELA credit (all)				Percentage of students earning ELA credit (SWD)				Percentage of students earning ELA credit (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	147,779	246,285	273,646	203,862	57,175	95,099	105,052	75,167	118,985	200,275	223,819	165,940
	Estimate	−1.257*	−1.382*	−1.451*	−0.972*	−0.541	−1.314*	−1.486*	−0.786	−1.005*	−1.142*	−1.500*	−0.758
	SD	0.408	0.333	0.289	0.369	0.596	0.496	0.431	0.573	0.46	0.332	0.283	0.383
OSS versus ISS	*N*	41,428	70,216	79,173	60,779	15,592	26,124	28,881	21,167	33,125	56,610	64,195	48,883
	Estimate	−1.794	−2.024*	−4.171*	−3.637*	−0.751	−1.563	−2.910*	−1.736	−1.042	−1.912*	−4.157*	−3.207*
	SD	0.924	0.76	0.87	1.028	1.319	0.985	1.03	1.088	0.899	0.818	0.938	1.015
ISS 2–3 versus 1	*N*	15,399	26,074	29,127	20,954	5,930	9,978	10,866	7,413	12,415	21,302	24,104	17,355
	Estimate	−1.961	−1.852*	−0.589	−2.671*	−3.05	−1.856	−2.298	−4.424*	−2.219	−1.712	−1.032	−3.445*
	SD	1.081	0.844	0.909	1.132	2.338	1.569	1.744	1.9	1.316	1.028	1.098	1.165
ISS 4–5 versus 2–3	*N*	18,523	30,581	33,277	23,656	6,965	11,423	12,163	8,214	14,984	25,020	27,447	19,484
	Estimate	−1.704*	−1.377*	−1.648*	−0.88	−1.337	−0.592	−0.924	−0.544	−2.041*	−1.574*	−1.345	−0.974
	SD	0.739	0.616	0.707	0.749	1.219	1.227	1.067	1.366	0.708	0.655	0.708	0.809
OSS 6–20 versus OSS 1–5	*N*	5,288	8,106	8,405	6,106	2,259	3,546	3,665	2,570	4,310	6,695	7,025	5,050
	Estimate	−0.054	−1.93	0.181	−1.521	2.148	−2.277	−0.464	0.041	0.34	−3.875*	−0.772	−0.713
	SD	1.808	1.765	1.735	1.695	2.696	2.589	2.332	3.267	2.09	1.921	1.864	1.943
	N	8,941	13,614	13,651	9,580	3,691	5,608	5,559	3,773	7,428	11,460	11,578	8,064
	Estimate	−4.837*	−4.145*	−4.055*	−2.530*	−4.645*	−4.020*	−4.910*	−4.182*	−5.236*	−3.309*	−4.316*	−2.556

		Percentage of students earning ELA credit (all)				Percentage of students earning ELA credit (SWD)				Percentage of students earning ELA credit (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
OSS 21+ versus OSS 6–20	SD	1.25	0.999	0.968	1.144	2.25	1.883	1.624	1.885	1.552	1.148	1.096	1.42

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.15. Effect of Discipline Type and Length on Likelihood of Middle School Students' Later Math Credit Accumulation, by Race

		Percentage of students earning math credit (all)				Percentage of students earning math credit (Black)				Percentage of students earning math credit (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	147,779	246,285	273,646	203,862	68,742	118,025	133,335	100,213	56,786	92,335	100,142	73,191
	Estimate	−0.55	−1.074*	−1.451*	−0.997*	−0.89	−1.312*	−1.425*	−0.901*	−0.503	−0.387	−0.744	−0.652
	SD	0.388	0.353	0.321	0.393	0.513	0.443	0.393	0.43	0.654	0.5	0.497	0.656
OSS versus ISS	*N*	41,428	70,216	79,173	60,779	18,809	32,552	37,649	29,012	16,745	27,829	30,527	23,090
	Estimate	−2.736*	−3.114*	−3.438*	−2.158*	−2.668*	−3.229*	−2.957*	−2.402*	−3.056*	−2.972*	−2.909*	−2.523*
	SD	1.079	0.7	0.812	0.954	1.29	0.897	0.86	0.982	1.121	0.893	1.105	1.132
ISS 2–3 versus 1	*N*	15,399	26,074	29,127	20,954	6,352	11,005	12,775	9,296	6,310	10,575	11,520	8,108
	Estimate	−1.752	−2.132*	−3.017*	−1.748	−2.169	−2.945*	−2.771	0.827	0.365	−3.135*	−4.060*	−4.616*
	SD	1.315	0.957	0.88	1.027	1.892	1.239	1.435	1.691	2.188	1.487	1.532	1.894
ISS 4–5 versus 2–3	*N*	18,523	30,581	33,277	23,656	8,019	13,563	15,305	10,972	7,651	12,401	13,150	9,182
	Estimate	−1.309	−1.224*	−2.407*	−1.864*	0.05	−1.562	−1.322	−1.037	−1.104	0.225	−1.891*	−0.279

		Percentage of students earning math credit (all)				Percentage of students earning math credit (Black)				Percentage of students earning math credit (Hispanic)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
	SD	1.017	0.483	0.64	0.627	1.367	0.891	1.059	0.992	1.746	0.981	0.932	1.202
OSS 6–20 versus OSS 1–5	*N*	5,288	8,106	8,405	6,106	2,804	4,415	4,644	3,380	2,002	3,010	3,090	2,252
	Estimate	2.685	0.765	−1.319	−2.208	0.953	−2.267	−2.028	−3.539	1.749	−1.037	1.281	−5.122
	SD	2.082	1.563	1.607	1.672	2.746	1.949	2.064	3.029	4.071	2.986	3.35	3.093
OSS 21+ versus OSS 6–20	*N*	8,941	13,614	13,651	9,580	4,953	7,816	7,990	5,634	3,329	4,864	4,769	3,331
	Estimate	−5.189*	−4.443*	−5.361*	−3.483*	−2.902	−4.960*	−3.052	−2.738	−7.022*	−3.272	−5.988*	−1.347
	SD	1.227	1.015	1.022	1.191	1.709	1.442	1.87	1.691	1.956	1.647	1.648	1.797

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.16. Effect of Discipline Type and Length on Likelihood of Middle School Students' Later Math Credit Accumulation, by Special Status

		Percentage of students earning math credit (all)				Percentage of students earning math credit (SWD)				Percentage of students earning math credit (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
Exclusion versus no exclusion	*N*	147,779	246,285	273,646	203,862	57,175	95,099	105,052	75,167	118,985	200,275	223,819	165,940
	Estimate	−0.55	−1.074*	−1.451*	−0.997*	−0.033	−0.752	−0.929	−0.767	−0.589	−0.584	−1.165*	−0.669
	SD	0.388	0.353	0.321	0.393	0.594	0.53	0.486	0.547	0.417	0.354	0.327	0.41
OSS versus ISS	*N*	41,428	70,216	79,173	60,779	15,592	26,124	28,881	21,167	33,125	56,610	64,195	48,883
	Estimate	−2.736*	−3.114*	−3.438*	−2.158*	−2.449	−2.545*	−1.936*	−2.340*	−2.819*	−3.267*	−3.763*	−2.230*

		Percentage of students earning math credit (all)				Percentage of students earning math credit (SWD)				Percentage of students earning math credit (students from economically disadvantaged homes)			
		1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident	1 year after incident	2 years after incident	3 years after incident	4 years after incident
	SD	1.079	0.7	0.812	0.954	1.355	0.851	0.941	1.106	1.127	0.758	0.869	0.941
ISS 2–3 versus 1	*N*	15,399	26,074	29,127	20,954	5,930	9,978	10,866	7,413	12,415	21,302	24,104	17,355
	Estimate	−1.752	−2.132*	−3.017*	−1.748	−3.664	−4.238*	−5.686*	−4.012	−1.133	−1.595	−3.638*	−2.313
	SD	1.315	0.957	0.88	1.027	1.881	1.731	1.622	2.383	1.551	1.08	0.994	1.267
ISS 4–5 versus 2–3	*N*	18,523	30,581	33,277	23,656	6,965	11,423	12,163	8,214	14,984	25,020	27,447	19,484
	Estimate	−1.309	−1.224*	−2.407*	−1.864*	0.875	−0.748	−0.638	−1.043	−1.4	−1.260*	−2.016*	−1.560*
	SD	1.017	0.483	0.64	0.627	1.27	1.209	1.16	1.255	1.054	0.581	0.668	0.719
OSS 6–20 versus OSS 1–5	*N*	5,288	8,106	8,405	6,106	2,259	3,546	3,665	2,570	4,310	6,695	7,025	5,050
	Estimate	2.685	0.765	−1.319	−2.208	2.596	−0.797	1.064	−2.191	2.245	0.389	−1.393	−2.104
	SD	2.082	1.563	1.607	1.672	3.482	2.134	2.513	3.36	2.228	1.601	1.648	1.889
OSS 21+ versus OSS 6–20	*N*	8,941	13,614	13,651	9,580	3,691	5,608	5,559	3,773	7,428	11,460	11,578	8,064
	Estimate	−5.189*	−4.443*	−5.361*	−3.483*	−5.803*	−6.031*	−4.631*	−3.424	−4.719*	−3.566*	−5.601*	−3.567*
	SD	1.227	1.015	1.022	1.191	1.985	1.558	1.394	1.737	1.385	1.095	1.231	1.275

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.17. Effect of Discipline Type and Length on Middle School Students' Likelihood of On-Time High School Graduation, by Race and Special Status

		Percentage of students graduating on time (all)	Percentage of students graduating on time (Black)	Percentage of students graduating on time (Hispanic)	Percentage of students graduating on time (SWD)	Percentage of students graduating on time (students from economically disadvantaged homes)
Exclusion versus no exclusion	*N*	200,653	97,989	74,559	75,601	162,266
	Estimate	−0.897*	−0.933*	−0.473	−0.453	−0.513
	SD	0.341	0.453	0.576	0.465	0.368
OSS versus ISS	*N*	62,414	29,518	24,585	22,232	49,803
	Estimate	−3.385*	−3.061*	−3.170*	−0.869	−3.314*
	SD	0.784	0.89	0.983	1.016	0.839
ISS 2–3 versus 1	*N*	17,280	7,575	6,884	6,305	14,085
	Estimate	−1.384	−0.808	−1.714	−1.725	−2.816
	SD	1.337	1.935	2.187	2.011	1.442
ISS 4–5 versus 2–3	*N*	20,029	9,223	7,956	7,137	16,286
	Estimate	−2.336*	−0.53	−1.952	0.111	−2.594*
	SD	0.775	0.953	1.127	1.043	0.804
OSS 6–20 versus OSS 1–5	*N*	6,713	3,641	2,583	2,845	5,543
	Estimate	−0.85	−3.105	−0.902	−1.266	−0.689
	SD	1.6	2.067	2.863	2.285	1.383
OSS 21+ versus OSS 6–20	*N*	11,138	6,372	4,067	4,380	9,296
	Estimate	−4.512*	−3.036*	−3.483*	−2.898*	−4.488*
	SD	0.952	1.401	1.369	1.178	0.979

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.18. Effect of Discipline Type and Length on High School Students' Later Behavioral Incidents, by Race

		Total incidents (all)			Total incidents (Black)			Total incidents (Hispanic)			Total incidents (White)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	420,362	241,867	104,801	219,937	127,342	55,510	151,358	87,296	37,998	30,659	16,939	6,990
	Estimate	−0.060*	−0.035*	−0.042*	−0.076*	−0.057*	−0.054	−0.139*	−0.059*	−0.002	−0.207*	−0.041	0.002
	SD	0.018	0.017	0.021	0.024	0.028	0.031	0.033	0.027	0.027	0.058	0.057	0.059
OSS versus ISS	*N*	143,234	84,316	37,203	76,561	45,203	20,190	51,226	30,214	13,237	9,485	5,456	2,284
	Estimate	0.023	0.026	0.119*	0.026	−0.057	0.056	−0.041	0.008	0.016	−0.16	−0.191	−0.14
	SD	0.035	0.037	0.053	0.045	0.06	0.09	0.055	0.074	0.067	0.172	0.16	0.308
ISS 2–3 versus 1	*N*	56,445	33,279	14,643	28,366	16,848	7,514	20,860	12,346	5,423	4,607	2,638	1,057
	Estimate	0.031	0.089	0.067	0.019	0.154*	0.042	−0.007	−0.001	−0.046	−0.166	−0.118	0.125
	SD	0.06	0.052	0.057	0.098	0.068	0.107	0.07	0.073	0.074	0.114	0.133	0.082
ISS 4–5 versus 2–3	*N*	85,223	48,099	20,337	46,147	26,196	11,184	30,050	16,933	7,181	5,498	3,026	1,180
	Estimate	0.009	−0.005	0.02	−0.014	−0.003	0.009	−0.022	−0.02	0.033	0.065	−0.045	0.217
	SD	0.027	0.028	0.033	0.039	0.038	0.052	0.041	0.049	0.034	0.12	0.084	0.155
OSS 6–20 versus OSS 1–5	*N*	10,759	5,935	2,470	5,815	3,220	1,371	3,862	2,132	869	610	312	124
	Estimate	0.007	0.026	−0.025	−0.034	0.091	0.164	−0.056	0.001	−0.082	−0.034	0.043	−0.293
	SD	0.099	0.103	0.125	0.116	0.135	0.123	0.153	0.153	0.185	0.255	0.551	0.495
OSS 21+ versus OSS 6–20	*N*	16,749	9,098	3,654	9,562	5,239	2,162	5,764	3,116	1,227	800	392	140
	Estimate	0.016	−0.095	−0.043	0.016	−0.153	−0.137	0.07	−0.055	0.197	0.224	−0.289	−0.862
	SD	0.082	0.067	0.129	0.104	0.114	0.198	0.112	0.111	0.134	0.255	0.277	0.497

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.19. Effect of Discipline Type and Length on High School Students' Later Behavioral Incidents, by Special Status

		Total incidents (all)			Total incidents (SWD)			Total incidents (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	420,362	241,867	104,801	138,576	83,190	39,597	324,382	189,355	83,213
	Estimate	−0.060*	−0.035*	−0.042*	−0.165*	−0.064	−0.075	−0.073*	−0.034	−0.039
	SD	0.018	0.017	0.021	0.037	0.032	0.048	0.022	0.018	0.023
OSS versus ISS	*N*	143,234	84,316	37,203	44,626	27,257	12,951	110,635	66,121	29,615
	Estimate	0.023	0.026	0.119*	−0.074	−0.114	−0.046	0.026	0.047	0.096
	SD	0.035	0.037	0.053	0.073	0.091	0.082	0.039	0.04	0.053
ISS 2–3 versus 1	*N*	56,445	33,279	14,643	18,274	11,245	5,422	44,215	26,532	11,839
	Estimate	0.031	0.089	0.067	−0.129	0.036	0.048	−0.013	0.072	0.09
	SD	0.06	0.052	0.057	0.093	0.095	0.105	0.073	0.056	0.063
ISS 4–5 versus 2–3	*N*	85,223	48,099	20,337	27,755	16,447	7,611	67,512	38,703	16,626
	Estimate	0.009	−0.005	0.02	−0.014	−0.027	−0.014	−0.004	−0.004	0.012
	SD	0.027	0.028	0.033	0.054	0.052	0.054	0.029	0.028	0.037
OSS 6–20 versus OSS 1–5	*N*	10,759	5,935	2,470	3,684	2,074	914	8,396	4,683	1,978
	Estimate	0.007	0.026	−0.025	0.09	0.197	0.038	−0.02	0.064	−0.041
	SD	0.099	0.103	0.125	0.164	0.147	0.174	0.107	0.129	0.118
OSS 21+ versus OSS 6–20	N	16,749	9,098	3,654	5,407	2,962	1,284	13,183	7,236	2,955
	Estimate	0.016	−0.095	−0.043	−0.064	−0.196	−0.19	0.024	−0.036	−0.056
	SD	0.082	0.067	0.129	0.135	0.113	0.146	0.083	0.086	0.133

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.20. Effect of Discipline Type and Length on High School Students' Later Serious Behavioral Incidents, by Race

		Total incidents (all)			Total incidents (Black)			Total incidents (Hispanic)			Total incidents (White)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	415,850	237,798	102,014	216,992	124,752	53,708	150,075	86,081	37,217	30,460	16,749	6,841
	Estimate	−0.006	−0.003	−0.008	−0.014	−0.008	−0.02	−0.02	−0.011	0.021	0.002	0.021	0.016
	SD	0.006	0.006	0.008	0.008	0.01	0.013	0.01	0.009	0.011	0.024	0.021	0.018
OSS versus ISS	*N*	142,031	82,908	36,190	75,828	44,303	19,529	50,837	29,786	12,954	9,432	5,398	2,239
	Estimate	0.029	0.036*	0.047*	0.022	−0.003	0.024	0.017	0.025	0.024	−0.043	−0.081	−0.044
	SD	0.015	0.016	0.022	0.02	0.025	0.03	0.03	0.032	0.043	0.067	0.099	0.115
ISS 2–3 versus 1	*N*	56,024	32,725	14,258	28,125	16,494	7,275	20,708	12,186	5,308	4,585	2,608	1,037
	Estimate	0.01	0.022	−0.001	0.023	0.04	−0.017	−0.046	−0.018	−0.032	−0.065	−0.095	0.021
	SD	0.024	0.02	0.024	0.035	0.031	0.039	0.035	0.028	0.037	0.063	0.091	0.044
ISS 4–5 versus 2–3	*N*	84,532	47,236	19,707	45,729	25,614	10,771	29,825	16,706	7,006	5,470	2,985	1,148
	Estimate	0.019	0.011	0.029	0.019	0.012	0.043*	0.011	−0.004	0.006	−0.01	−0.037	0.016
	SD	0.011	0.011	0.016	0.016	0.013	0.02	0.015	0.019	0.016	0.032	0.042	0.053
OSS 6–20 versus OSS 1–5	*N*	10,671	5,838	2,404	5,759	3,156	1,327	3,832	2,102	850	608	309	122
	Estimate	−0.049	−0.035	−0.001	−0.109*	−0.028	0.098	−0.064	−0.062	−0.079	0.035	0.015	−0.39
	SD	0.038	0.047	0.047	0.047	0.07	0.06	0.071	0.073	0.09	0.104	0.161	0.294
OSS 21+ versus OSS 6–20	*N*	16,609	8,940	3,524	9,467	5,133	2,070	5,726	3,069	1,192	795	387	137
	Estimate	0.031	−0.02	−0.05	0.035	−0.029	−0.062	0.043	−0.024	−0.012	0.038	−0.011	−0.124
	SD	0.03	0.029	0.044	0.04	0.044	0.07	0.039	0.046	0.053	0.092	0.123	0.237

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.21. Effect of Discipline Type and Length on High School Students' Later Serious Behavioral Incidents, by Special Status

		Total incidents (all)			Total incidents (SWD)			Total incidents (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	415,850	237,798	102,014	137,528	82,185	38,818	321,130	186,341	81,075
	Estimate	−0.006	−0.003	−0.008	−0.025*	−0.012	−0.02	−0.011	0.002	−0.008
	SD	0.006	0.006	0.008	0.01	0.01	0.02	0.007	0.006	0.007
OSS versus ISS	*N*	142,031	82,908	36,190	44,359	26,913	12,670	109,703	65,014	28,843
	Estimate	0.029	0.036*	0.047*	−0.006	−0.015	0.008	0.035	0.040*	0.039
	SD	0.015	0.016	0.022	0.029	0.038	0.039	0.018	0.017	0.025
ISS 2–3 versus 1	*N*	56,024	32,725	14,258	18,172	11,096	5,301	43,898	26,100	11,535
	Estimate	0.01	0.022	−0.001	−0.046	−0.021	−0.032	0.004	0.016	0.012
	SD	0.024	0.02	0.024	0.038	0.047	0.046	0.026	0.023	0.025
ISS 4–5 versus 2–3	*N*	84,532	47,236	19,707	27,578	16,209	7,417	66,974	38,013	16,133
	Estimate	0.019	0.011	0.029	−0.006	−0.008	0.014	0.018	0.011	0.026
	SD	0.011	0.011	0.016	0.022	0.022	0.028	0.012	0.012	0.017
OSS 6–20 versus OSS 1–5	*N*	10,671	5,838	2,404	3,656	2,044	895	8,326	4,605	1,922
	Estimate	−0.049	−0.035	−0.001	−0.077	0.011	0.027	−0.064	−0.008	−0.024
	SD	0.038	0.047	0.047	0.066	0.066	0.108	0.043	0.053	0.057
OSS 21+ versus OSS 6–20	*N*	16,609	8,940	3,524	5,366	2,921	1,252	13,070	7,117	2,852
	Estimate	0.031	−0.02	−0.05	0.038	−0.07	−0.081	0.036	0.012	−0.044
	SD	0.03	0.029	0.044	0.048	0.045	0.062	0.036	0.029	0.046

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.22. Effect of Discipline Type and Length on High School Students' Later Number of Suspensions, by Race

		Total suspensions (all)			Total suspensions (Black)			Total suspensions (Hispanic)			Total suspensions (White)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	414,870	236,394	100,956	216,385	123,644	52,920	149,753	85,863	36,993	30,443	16,706	6,835
	Estimate	0.024*	0.005	−0.02	0.023*	0.002	−0.028	0.003	0.006	0.022	0.023	0.001	−0.008
	SD	0.009	0.008	0.014	0.01	0.01	0.017	0.018	0.011	0.012	0.036	0.028	0.051
OSS versus ISS	*N*	141,840	82,433	35,733	75,687	43,919	19,186	50,803	29,727	12,881	9,426	5,378	2,223
	Estimate	−0.01	0.024	0.059*	−0.013	−0.01	0.041	−0.016	0.031	0.028	−0.11	−0.11	−0.101
	SD	0.018	0.018	0.024	0.022	0.026	0.043	0.027	0.033	0.031	0.065	0.11	0.139
ISS 2–3 versus 1	*N*	55,957	32,569	14,093	28,074	16,362	7,142	20,703	12,170	5,284	4,579	2,601	1,033
	Estimate	0.019	0.018	0.007	0.027	0.06	0.001	−0.015	−0.012	−0.061	−0.064	−0.104	0.047
	SD	0.028	0.027	0.027	0.044	0.035	0.042	0.037	0.034	0.044	0.124	0.081	0.032
ISS 4–5 versus 2–3	*N*	84,354	46,892	19,393	45,594	25,357	10,548	29,794	16,641	6,952	5,460	2,972	1,128
	Estimate	0.009	0.004	0.021	0.007	−0.001	0.025	−0.014	0.002	0.007	−0.081	−0.058	0.034
	SD	0.013	0.013	0.013	0.018	0.019	0.02	0.02	0.016	0.015	0.055	0.059	0.073
OSS 6–20 versus OSS 1–5	*N*	10,603	5,780	2,377	5,705	3,114	1,307	3,823	2,092	843	605	306	122
	Estimate	0.018	0.039	0.038	−0.038	0.052	0.127*	0.003	0.02	−0.052	0.025	0.009	−0.113
	SD	0.043	0.045	0.036	0.057	0.063	0.05	0.068	0.06	0.068	0.118	0.225	0.318
OSS 21+ versus OSS 6–20	*N*	16,526	8,852	3,492	9,407	5,066	2,047	5,709	3,055	1,187	791	383	137
	Estimate	−0.003	−0.029	−0.049	0.017	−0.052	−0.095	0.015	0.017	0.116	0.085	−0.162	−0.392
	SD	0.031	0.033	0.045	0.039	0.043	0.065	0.046	0.044	0.065	0.109	0.177	0.355

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.23. Effect of Discipline Type and Length on High School Students' Later Number of Suspensions, by Special Status

		Total suspensions (all)			Total suspensions (SWD)			Total suspensions (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	414,870	236,394	100,956	137,161	81,743	38,549	320,309	185,155	80,205
	Estimate	0.071*	0.032*	0.006	0.016	0.009	−0.03	0.025*	0.014	−0.022
	SD	0.01	0.007	0.009	0.02	0.013	0.037	0.011	0.007	0.015
OSS versus ISS	*N*	141,840	82,433	35,733	44,272	26,722	12,553	109,574	64,655	28,482
	Estimate	−0.014	0.022	0.054*	−0.037	−0.038	0.011	−0.006	0.032	0.057*
	SD	0.019	0.022	0.027	0.032	0.033	0.035	0.018	0.02	0.025
ISS 2–3 versus 1	*N*	55,957	32,569	14,093	18,134	11,047	5,262	43,840	25,970	11,403
	Estimate	0.018	0.018	−0.017	−0.055	−0.004	−0.019	0.008	0.009	0.001
	SD	0.023	0.021	0.024	0.052	0.065	0.054	0.036	0.026	0.028
ISS 4–5 versus 2–3	*N*	84,354	46,892	19,393	27,507	16,067	7,327	66,826	37,739	15,881
	Estimate	0.01	0.006	0.018	−0.027	−0.025	0.023	0.011	0.003	0.018
	SD	0.013	0.014	0.013	0.031	0.025	0.023	0.014	0.014	0.015
OSS 6–20 versus OSS 1–5	*N*	10,603	5,780	2,377	3,639	2,022	885	8,275	4,556	1,897
	Estimate	0.033	0.051	0.028	0.072	0.12	−0.065	0.02	0.071	0.009
	SD	0.042	0.043	0.043	0.059	0.071	0.092	0.045	0.044	0.043
OSS 21+ versus OSS 6–20	*N*	16,526	8,852	3,492	5,349	2,894	1,237	13,012	7,043	2,818
	Estimate	−0.006	−0.033	−0.046	0.025	−0.065	−0.059	0.01	−0.004	−0.035
	SD	0.032	0.036	0.044	0.052	0.053	0.069	0.032	0.031	0.046

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.24. Effect of Discipline Type and Length on High School Students' Total Days Suspended, by Race

		Total days suspended (all)			Total days suspended (Black)			Total days suspended (Hispanic)			Total days suspended (White)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	414,867	236,393	100,955	216,382	123,644	52,920	149,753	85,862	36,992	30,443	16,706	6,835
	Estimate	0.323*	−0.077	−0.14	0.225	−0.093	−0.313	−0.04	−0.064	0.370*	0.186	0.395	0.171
	SD	0.113	0.112	0.13	0.196	0.173	0.227	0.191	0.117	0.149	0.4	0.275	0.24
OSS versus ISS	*N*	141,839	82,432	35,732	75,686	43,919	19,186	50,803	29,726	12,880	9,426	5,378	2,223
	Estimate	2.468*	1.533*	1.310*	2.375*	1.055	0.882	1.977*	1.725*	1.515*	1.197	−6.064	−1.479
	SD	0.456	0.337	0.381	0.656	0.541	0.6	0.509	0.554	0.563	1.962	7.225	2.476
ISS 2–3 versus 1	*N*	55,957	32,568	14,092	28,074	16,362	7,142	20,703	12,169	5,283	4,579	2,601	1,033
	Estimate	−0.19	0.818*	0.397	0.113	0.876*	−0.26	−0.68	0.761	0.019	0.202	−0.051	0.212
	SD	0.727	0.338	0.439	0.849	0.43	1.061	0.953	0.429	0.495	0.486	0.58	0.205
ISS 4–5 versus 2–3	*N*	84,354	46,891	19,392	45,594	25,357	10,548	29,794	16,640	6,951	5,460	2,972	1,128
	Estimate	0.981*	0.22	0.23	1.105*	0.339	0.367	0.637	−0.052	−0.248	0.7	0.584	0.151
	SD	0.256	0.218	0.222	0.371	0.306	0.452	0.343	0.254	0.39	0.412	0.76	0.55
OSS 6–20 versus OSS 1–5	*N*	10,603	5,780	2,377	5,705	3,114	1,307	3,823	2,092	843	605	306	122
	Estimate	0.083	−0.129	0.697	−0.972	−0.414	1.529	−0.554	0.299	0.03	0.402	1.527	−0.152
	SD	0.709	0.642	0.974	1.148	1.119	1.488	1.365	0.845	1.465	2.625	3.614	2.248
OSS 21+ versus OSS 6–20	*N*	16,525	8,852	3,492	9,406	5,066	2,047	5,709	3,055	1,187	791	383	137
	Estimate	3.017*	1.382*	−1.33	3.497*	1.511	−1.968	2.316*	1.265	0.835	2.492	0.57	−1.656
	SD	0.618	0.603	1.021	1.081	0.758	1.362	0.856	1.084	1.34	1.88	2.602	2.673

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.25. Effect of Discipline Type and Length on High School Students' Total Days Suspended, by Special Status

		Total suspensions (all)			Total days suspended (SWD)			Total days suspended (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	414,867	236,393	100,955	137,161	81,743	38,549	320,307	185,154	80,204
	Estimate	0.323*	−0.077	−0.14	0.173	−0.098	−0.204	0.134	0.014	−0.234
	SD	0.113	0.112	0.13	0.165	0.254	0.226	0.141	0.146	0.181
OSS versus ISS	*N*	141,839	82,432	35,732	44,272	26,722	12,553	109,573	64,654	28,481
	Estimate	2.468*	1.533*	1.310*	1.985*	0.387	1.891*	2.345*	1.594*	1.549*
	SD	0.456	0.337	0.381	0.625	0.716	0.839	0.502	0.364	0.486
ISS 2–3 versus 1	*N*	55,957	32,568	14,092	18,134	11,047	5,262	43,840	25,969	11,402
	Estimate	−0.19	0.818*	0.397	−0.162	0.638	0.901	−0.065	0.578	0.337
	SD	0.727	0.338	0.439	0.587	0.534	0.501	0.581	0.346	0.518
ISS 4–5 versus 2–3	*N*	84,354	46,891	19,392	27,507	16,067	7,327	66,826	37,738	15,880
	Estimate	0.981*	0.22	0.23	0.913*	0.085	0.472	0.999*	0.065	0.227
	SD	0.256	0.218	0.222	0.371	0.366	0.433	0.265	0.214	0.245
OSS 6–20 versus OSS 1–5	*N*	10,603	5,780	2,377	3,639	2,022	885	8,275	4,556	1,897
	Estimate	0.083	−0.129	0.697	0.138	−0.394	−1.387	−0.279	0.287	0.25
	SD	0.709	0.642	0.974	1.317	1.027	3.243	0.88	0.733	1.114
OSS 21+ versus OSS 6–20	*N*	16,525	8,852	3,492	5,349	2,894	1,237	13,011	7,043	2,818
	Estimate	3.017*	1.382*	−1.33	3.249*	0.257	−2.435	2.916*	1.591*	−1.447
	SD	0.618	0.603	1.021	1.024	0.957	1.373	0.862	0.643	1.158

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.26. Effect of Discipline Type and Length on High School Students' Later Absences, by Race

		Total absences (all)			Total absences (Black)			Total absences (Hispanic)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	410,724	232,463	98,289	213,523	120,926	51,062	148,653	84,839	36,332
	Estimate	1.015*	1.199*	1.768*	0.912*	1.18	1.542*	0.945	0.825	0.161
	SD	0.379	0.572	0.727	0.447	0.601	0.71	0.52	0.675	1.009
OSS versus ISS	*N*	140,254	80,846	34,605	74,617	42,867	18,412	50,367	29,268	12,589
	Estimate	6.118*	6.940*	8.438*	7.287*	8.723*	10.962*	5.001*	3.888*	6.873*
	SD	0.598	0.896	1.803	0.946	1.171	2.157	0.776	1.619	3.278
ISS 2–3 versus 1	*N*	55,378	32,006	13,732	27,695	16,010	6,903	20,524	11,988	5,182
	Estimate	2.829*	4.118*	4.225*	2.191	2.946	0.423	1.195	2.221	5.114*
	SD	1.112	1.538	1.924	1.423	1.828	2.712	1.526	2.338	2.463
ISS 4–5 versus 2–3	*N*	83,392	45,968	18,795	44,935	24,731	10,137	29,531	16,387	6,797
	Estimate	5.466*	5.662*	6.925*	6.175*	6.555*	7.713*	3.485*	4.316*	4.540*
	SD	0.533	0.836	1.46	0.632	0.972	1.625	0.73	1.156	2.189
OSS 6–20 versus OSS 1–5	*N*	10,498	5,669	2,301	5,633	3,043	1,262	3,794	2,054	819
	Estimate	1.129	1.396	1.604	1.803	1.57	−1.715	1.438	−0.153	5.894
	SD	1.535	2.252	4.301	2.055	3.402	5.438	2.638	3.432	8.829
OSS 21+ versus OSS 6–20	*N*	16,342	8,643	3,354	9,266	4,910	1,951	5,673	3,004	1,151
	Estimate	6.469*	8.797*	10.344*	5.274*	9.055*	8.840*	7.064*	7.573*	10.777*
	SD	1.261	1.742	3.116	1.324	2.056	4.296	1.954	2.636	4.493

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.27. Effect of Discipline Type and Length on High School Students' Later Absences, by Special Status

		Total absences (all)			Total absences (SWD)			Total absences (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	410,724	232,463	98,289	136,093	80,708	37,902	317,257	182,122	78,169
	Estimate	1.015*	1.199*	1.768*	1.419*	2.197*	2.048*	0.75	0.924	1.394
	SD	0.379	0.572	0.727	0.502	0.711	1.015	0.4	0.612	0.756
OSS versus ISS	*N*	140,254	80,846	34,605	43,887	26,326	12,291	108,340	63,427	27,618
	Estimate	6.118*	6.940*	8.438*	7.272*	7.837*	7.702*	5.773*	6.766*	9.954*
	SD	0.598	0.896	1.803	0.979	1.645	2.488	0.708	0.969	1.907
ISS 2–3 versus 1	*N*	55,378	32,006	13,732	17,962	10,904	5,169	43,398	25,535	11,114
	Estimate	2.829*	4.118*	4.225*	2.905*	6.257*	3.708	2.309*	4.607*	2.237
	SD	1.112	1.538	1.924	1.294	1.649	2.741	0.954	1.408	2.116
ISS 4–5 versus 2–3	*N*	83,392	45,968	18,795	27,254	15,832	7,172	66,081	36,998	15,405
	Estimate	5.466*	5.662*	6.925*	6.829*	7.103*	7.958*	5.625*	5.626*	6.894*
	SD	0.533	0.836	1.46	0.78	1.041	1.598	0.636	0.87	1.604
OSS 6–20 versus OSS 1–5	*N*	10,498	5,669	2,301	3,608	1,992	873	8,189	4,473	1,841
	Estimate	1.129	1.396	1.604	1.686	−0.165	4.754	0.601	0.979	3.381
	SD	1.535	2.252	4.301	2.297	3.937	6.982	1.559	2.669	5.234
OSS 21+ versus OSS 6–20	*N*	16,342	8,643	3,354	5,302	2,839	1,206	12,863	6,878	2,718
	Estimate	6.469*	8.797*	10.344*	7.059*	5.325	2.649	6.262*	8.600*	11.991*
	SD	1.261	1.742	3.116	2.036	2.761	4.944	1.363	1.917	3.04

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.28. Effect of Discipline Type and Length on Likelihood of High School Students' Later ELA Credit Accumulation, by Race

		Percentage of students earning ELA credit (all)			Percentage of students earning ELA credit (Black)			Percentage of students earning ELA credit (Hispanic)			Percentage of students earning ELA credit (White)		
		1 year after incident	2 years after incident	3 years after incident	1 year after Incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	378,117	205,514	81,557	197,474	107,662	42,882	136,312	74,661	29,920	27,502	14,251	5,313
	Estimate	−0.687*	−0.769	−0.82	−0.197	−0.443	−0.797	−0.479	−0.004	−0.447	−0.865	−3.518*	0.523
	SD	0.302	0.387	0.513	0.428	0.42	0.746	0.428	0.518	0.814	0.971	0.958	2.051
OSS versus ISS	*N*	130,584	72,437	29,477	69,673	38,708	15,800	46,785	25,999	10,615	8,565	4,650	1,801
	Estimate	−2.743*	−4.396*	−4.209*	−3.667*	−5.495*	−5.713*	−2.017	−3.405*	−1.831	−0.146	3.076	8.194
	SD	0.747	0.879	1.483	0.785	1.105	2.024	1.025	1.481	2.613	2.889	3.159	7.145
ISS 2–3 versus 1	*N*	52,195	29,033	11,866	26,167	14,596	6,045	19,332	10,833	4,401	4,242	2,279	840
	Estimate	−3.165*	−3.092*	−2.07	−1.951	−1.901	0.547	−1.199	−3.1	−3.011	−1.594	−6.782*	−2.765
	SD	0.958	1.172	1.83	1.141	1.445	2.661	1.345	1.775	2.865	3.422	2.929	4.634
ISS 4–5 versus 2–3	*N*	77,720	41,207	15,926	42,029	22,359	8,693	27,430	14,565	5,655	4,982	2,543	903
	Estimate	−2.899*	−2.860*	−3.351*	−2.746*	−3.548*	−3.655*	−2.052*	−1.438	−2.05	−2.805	−2.807	−6.704
	SD	0.49	0.581	0.823	0.553	0.826	1.19	0.756	0.865	1.475	1.707	1.791	3.577
OSS 6–20 versus OSS 1–5	*N*	9,689	5,014	1,917	5,201	2,702	1,056	3,496	1,797	677	552	280	100
	Estimate	−4.729*	−2.632	−5.471	−5.901*	−2.232	−1.757	−2.432	−4.997	−9.773	−1.972	−0.634	11.796
	SD	1.612	2.069	3.751	2.202	3.102	5.177	2.807	4.061	7.387	6.492	6.675	20.405
OSS 21+ versus OSS 6–20	*N*	14,911	7,521	2,710	8,467	4,288	1,599	5,158	2,599	908	704	330	107
	Estimate	−4.771*	−7.378*	−4.606	−5.417*	−7.077*	−5.257	−5.213*	−7.203*	−6.697	−0.593	−12.44	−21.994
	SD	0.927	1.624	2.713	1.343	1.925	3.536	1.706	2.548	4.03	4.91	11.624	22.259

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.29. Effect of Discipline Type and Length on Likelihood of High School Students' Later ELA Credit Accumulation, by Special Status

		Percentage of students earning ELA credit (all)			Percentage of students earning ELA credit (SWD)			Percentage of students earning ELA credit (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	378,117	205,514	81,557	119,209	66,860	28,721	292,961	161,339	64,859
	Estimate	−0.687*	−0.769	−0.82	−1.061*	−0.892	−1.032	−0.453	−0.235	−0.472
	SD	0.302	0.387	0.513	0.482	0.729	0.894	0.373	0.476	0.519
OSS versus ISS	*N*	130,584	72,437	29,477	40,271	23,004	10,118	101,090	56,824	23,508
	Estimate	−2.743*	−4.396*	−4.209*	−3.367*	−4.420*	−4.028	−2.508*	−4.025*	−4.729*
	SD	0.747	0.879	1.483	1.042	1.593	2.261	0.812	1.085	1.754
ISS 2–3 versus 1	*N*	52,195	29,033	11,866	16,729	9,687	4,357	40,950	23,164	9,612
	Estimate	−3.165*	−3.092*	−2.07	−3.184*	−2.668	−4.369	−2.955*	−2.629*	−1.496
	SD	0.958	1.172	1.83	1.51	2.045	2.358	0.965	0.992	1.943
ISS 4–5 versus 2–3	*N*	77,720	41,207	15,926	25,079	13,831	5,911	61,674	33,174	13,064
	Estimate	−2.899*	−2.860*	−3.351*	−3.410*	−3.239*	−4.061*	−2.934*	−3.170*	−3.603*
	SD	0.49	0.581	0.823	0.744	0.937	1.703	0.592	0.63	0.966
OSS 6–20 versus OSS 1–5	*N*	9,689	5,014	1,917	3,221	1,682	691	7,572	3,950	1,527
	Estimate	−4.729*	−2.632	−5.471	−6.102*	−7.039*	−2.188	−3.402	−4.828*	−4.855
	SD	1.612	2.069	3.751	2.743	3.359	6.773	1.812	2.248	4.603
OSS 21+ versus OSS 6–20	*N*	14,911	7,521	2,710	4,755	2,406	940	11,773	5,987	2,184
	Estimate	−4.771*	−7.378*	−4.606	−5.595*	−7.386*	−0.833	−4.452*	−7.555*	−7.117*
	SD	0.927	1.624	2.713	1.749	2.843	4.954	1.054	1.95	3.045

Note. ELA = English language arts; ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.30. Effect of Discipline Type and Length on Likelihood of High School Students' Later Math Credit Accumulation, by Race

		Percentage of students earning math credit (all)			Percentage of students earning math credit (Black)			Percentage of students earning math credit (Hispanic)			Percentage of students earning math credit (White)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	378,117	205,514	81,557	197,474	107,662	42,882	136,312	74,661	29,920	27,502	14,251	5,313
	Estimate	−0.529	−0.569	−0.614	−0.09	−1.128*	−1.014	−0.527	0.548	0.382	−1.411	−3.068	−3.161
	SD	0.274	0.488	0.536	0.307	0.537	0.853	0.498	0.632	0.899	0.906	1.683	2.325
OSS versus ISS	*N*	130,584	72,437	29,477	69,673	38,708	15,800	46,785	25,999	10,615	8,565	4,650	1,801
	Estimate	−3.112*	−4.678*	−1.208	−4.185*	−5.503*	−4.797*	−1.623	−4.027*	1.698	−4.13	0.527	4.613
	SD	0.683	0.891	1.494	0.927	1.134	2.025	0.89	1.728	2.302	2.795	3.707	5.923
ISS 2–3 versus 1	*N*	52,195	29,033	11,866	26,167	14,596	6,045	19,332	10,833	4,401	4,242	2,279	840
	Estimate	−5.097*	−2.693*	−1.848	−4.625*	−1.445	0.656	−2.428	−2.595	−3.909	−2.146	−6.219	−3.22
	SD	1.041	1.015	2.047	1.228	1.43	2.942	1.394	1.746	2.764	2.239	3.218	5.601
ISS 4–5 versus 2–3	*N*	77,720	41,207	15,926	42,029	22,359	8,693	27,430	14,565	5,655	4,982	2,543	903
	Estimate	−2.692*	−1.931*	−3.507*	−2.599*	−3.231*	−4.479*	−1.166	−0.159	−0.724	−3.760*	−2.443	−2.981
	SD	0.466	0.568	1.003	0.582	0.8	1.238	0.703	0.942	1.702	1.681	2.521	3.836
OSS 6–20 versus OSS 1–5	*N*	9,689	5,014	1,917	5,201	2,702	1,056	3,496	1,797	677	552	280	100
	Estimate	−1.995	−2.837	−1.726	−1.518	0.044	3.286	−2.441	−1.39	−10.201	−1.742	−0.332	20.229
	SD	1.741	2.288	4.372	2.013	3.194	5.339	3.393	4.496	7.275	5.224	8.943	21.073
OSS 21+ versus OSS 6–20	*N*	14,911	7,521	2,710	8,467	4,288	1,599	5,158	2,599	908	704	330	107
	Estimate	−6.237*	−6.028*	−5.087	−7.821*	−6.636*	−7.377*	−2.926	−5.634*	−2.291	−3.635	−14.673	−38.098
	SD	1.12	1.604	2.808	1.456	2.232	3.571	1.69	2.542	4.257	5.348	9.385	22.107

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.31. Effect of Discipline Type and Length on Likelihood of High School Students' Later Math Credit Accumulation, by Special Status

		Percentage of students earning math credit (all)			Percentage of students earning math credit (SWD)			Percentage of students earning math credit (students from economically disadvantaged homes)		
		1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident	1 year after incident	2 years after incident	3 years after incident
Exclusion versus no exclusion	*N*	378,117	205,514	81,557	119,209	66,860	28,721	292,961	161,339	64,859
	Estimate	−0.529	−0.569	−0.614	−1.092*	−1.226	−1.538	−0.222	−0.247	−0.686
	SD	0.274	0.488	0.536	0.515	0.658	0.984	0.327	0.494	0.536
OSS versus ISS	*N*	130,584	72,437	29,477	40,271	23,004	10,118	101,090	56,824	23,508
	Estimate	−3.112*	−4.678*	−1.208	−3.093*	−3.770*	−3.405	−3.096*	−4.645*	−2.527
	SD	0.683	0.891	1.494	1.096	1.338	2.148	0.836	1.01	1.635
ISS 2–3 versus 1	*N*	52,195	29,033	11,866	16,729	9,687	4,357	40,950	23,164	9,612
	Estimate	−5.097*	−2.693*	−1.848	−6.003*	−3.863*	−3.643	−4.943*	−1.874*	−3.057
	SD	1.041	1.015	2.047	1.513	1.792	2.221	1.072	0.888	2.205
ISS 4–5 versus 2–3	*N*	77,720	41,207	15,926	25,079	13,831	5,911	61,674	33,174	13,064
	Estimate	−2.692*	−1.931*	−3.507*	−3.630*	−3.289*	−4.223*	−2.833*	−1.643*	−2.785*
	SD	0.466	0.568	1.003	0.682	0.858	1.688	0.55	0.736	1.017
OSS 6–20 versus OSS 1–5	*N*	9,689	5,014	1,917	3,221	1,682	691	7,572	3,950	1,527
	Estimate	−1.995	−2.837	−1.726	−6.739*	−2.84	5.062	−3.561	−3.381	−3.497
	SD	1.741	2.288	4.372	2.329	3.751	6.187	1.842	2.504	4.895
OSS 21+ versus OSS 6–20	*N*	14,911	7,521	2,710	4,755	2,406	940	11,773	5,987	2,184
	Estimate	−6.237*	−6.028*	−5.087	−5.985*	−4.325	−0.309	−5.435*	−7.602*	−6.663*
	SD	1.12	1.604	2.808	1.533	2.955	4.529	1.217	1.664	3.017

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

Table B.32. Effect of Discipline Type and Length on High School Students' Likelihood of On-Time High School Graduation, by Race

		Percentage of students graduating on time (all)			Percentage of students graduating on time (Black)			Percentage of students graduating on time (Hispanic)			Percentage of students graduating on time (White)		
		Overall	9th or 10th graders	11th or 12th graders	Overall	9th or 10th graders	11th or 12th graders	Overall	9th or 10th graders	11th or 12th graders	Overall	9th or 10th graders	11th or 12th graders
Exclusion versus no exclusion	*N*	435,138	281,211	153,927	230,625	149,604	81,021	153,700	102,255	51,445	31,835	18,484	13,351
	Estimate	−0.353	−0.857*	−0.31	−0.258	−1.062*	0.153	0.145	0.083	−1.047	−1.944	−3.067*	−0.995
	SD	0.333	0.327	0.568	0.442	0.475	0.627	0.511	0.55	0.782	1.068	1.186	1.434
OSS versus ISS	*N*	145,480	99,713	45,767	78,121	53,474	24,647	51,228	36,131	15,097	9,916	6,237	3,679
	Estimate	−3.338*	−3.661*	−3.241*	−3.793*	−4.468*	−3.368*	−2.447*	−2.819*	−1.927	−0.427	−2.338	1.141
	SD	0.587	0.517	1.283	0.621	0.736	1.299	0.866	0.886	1.743	2.394	2.849	3.114
ISS 2–3 versus 1	*N*	56,029	37,152	18,877	28,368	18,963	9,405	20,379	13,862	6,517	4,682	2,785	1,897
	Estimate	−3.230*	−3.954*	−2.695	−3.341*	−3.724*	−4.408	−1.201	−1.742	−0.346	−3.902	−7.295	0.998
	SD	0.933	0.924	1.547	1.454	1.53	2.368	1.072	1.318	1.943	3.291	5.4	2.246
ISS 4–5 versus 2–3	*N*	84,238	56,549	27,689	45,867	30,664	15,203	29,264	20,287	8,977	5,567	3,424	2,143
	Estimate	−2.469*	−2.759*	−2.699*	−3.231*	−3.331*	−3.545*	−0.873	−1.893*	−0.101	−1.968	−3.176	−0.115
	SD	0.411	0.478	0.682	0.564	0.701	0.869	0.574	0.567	1.223	1.826	2.504	2.467
OSS 6–20 versus OSS 1–5	*N*	10,994	7,025	3,969	5,900	3,802	2,098	3,955	2,594	1,361	648	346	302
	Estimate	−2.875*	−2.132	−3.882	−2.326	−2.198	−1.82	−4.218*	−2.324	−7.932*	−1.702	−6.709	0.802
	SD	1.409	1.561	2.074	2.011	2.475	3.044	2.105	2.675	3.568	4.159	5.953	5.816
OSS 21+ versus OSS 6–20	*N*	16,629	11,426	5,203	9,489	6,537	2,952	5,697	4,022	1,675	804	475	329
	Estimate	−5.229*	−4.190*	−7.378*	−4.511*	−2.783*	−8.753*	−5.674*	−4.801*	−6.507*	−7.314	−8.988	−2.769
	SD	0.911	0.836	2.064	1.155	0.964	2.63	1.422	1.447	3.049	3.667	4.565	10.813

Note. ISS = in-school suspension; OSS = out-of-school suspension; * indicates effects are statistically significant at the .05 level.

Table B.33. Effect of Discipline Type and Length on High School Students' Likelihood of On-Time High School Graduation, by Special Status

		Percentage of students graduating on time (all)			Percentage of students graduating on time (SWD)			Percentage of students graduating on time (students from economically disadvantaged homes)		
		Overall	9th or 10th grade during year of incident	11th or 12th grade during year of incident	Overall	9th or 10th grade during year of incident	11th or 12th grade during year of incident	Overall	9th or 10th grade during year of incident	11th or 12th grade during year of incident
Exclusion versus no exclusion	*N*	435,138	281,211	153,927	131,138	91,556	39,582	329,093	217,855	111,238
	Estimate	−0.353	−0.857*	−0.31	−0.061	−0.72	0.76	0.268	−0.334	0.284
	SD	0.333	0.327	0.568	0.435	0.44	0.938	0.377	0.369	0.653
OSS versus ISS	*N*	145,480	99,713	45,767	42,365	31,439	10,926	109,995	77,041	32,954
	Estimate	−3.338*	−3.661*	−3.241*	−2.252*	−3.244*	−0.935	−3.246*	−3.850*	−3.044*
	SD	0.587	0.517	1.283	0.72	0.908	1.721	0.666	0.587	1.377
ISS 2–3 versus 1	*N*	56,029	37,152	18,877	16,923	12,270	4,653	43,025	29,243	13,782
	Estimate	−3.230*	−3.954*	−2.695	−2.716	−3.571	−1.198	−2.614*	−3.627*	−1.273
	SD	0.933	0.924	1.547	1.629	1.795	2.512	0.959	1.088	1.567
ISS 4–5 versus 2–3	*N*	84,238	56,549	27,689	25,535	18,738	6,797	65,459	44,838	20,621
	Estimate	−2.469*	−2.759*	−2.699*	−2.319*	−2.680*	−2.305	−2.037*	−2.443*	−2.339*
	SD	0.411	0.478	0.682	0.667	0.551	1.549	0.481	0.537	0.76
OSS 6–20 versus OSS 1–5	*N*	10,994	7,025	3,969	3,469	2,449	1,020	8,387	5,474	2,913
	Estimate	−2.875*	−2.132	−3.882	−0.466	1.885	−7.875	−2.036	−1.799	−2.883
	SD	1.409	1.561	2.074	1.815	1.866	4.212	1.666	1.965	2.173
OSS 21+ versus OSS 6–20	*N*	16,629	11,426	5,203	5,007	3,729	1,278	12,810	8,961	3,849
	Estimate	−5.229*	−4.190*	−7.378*	−3.413*	−3.053*	−3.025	−5.188*	−4.210*	−7.307*
	SD	0.911	0.836	2.064	1.163	1.213	3.194	0.986	0.918	2.351

Note. ISS = in-school suspension; OSS = out-of-school suspension; SWD = students with disabilities; * indicates effects are statistically significant at the .05 level.

About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. The AIR family of organizations now includes IMPAQ, Maher & Maher, and Kimetrica. For more information, visit AIR.ORG.

American Institutes for Research®
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**