Racial disparities in school-based disciplinary actions are associated with county-level rates of racial bias

Travis Riddle[a,1] and Stacey Sinclair[a,b,c]

[a]Department of Psychology, Princeton University, Princeton, NJ 08544; [b]Department of African American Studies, Princeton University, Princeton, NJ 08544; and [c]Department of Public Affairs, Princeton University, Princeton, NJ 08544

There are substantial gaps in educational outcomes between black and white students in the United States. Recently, increased attention has focused on differences in the rates at which black and white students are disciplined, finding that black students are more likely to be seen as problematic and more likely to be punished than white students are for the same offense. Although these disparities suggest that racial biases are a contributor, no previous research has shown associations with psychological measurements of bias and disciplinary outcomes. We show that county-level estimates of racial bias, as measured using data from approximately 1.6 million visitors to the Project Implicit website, are associated with racial disciplinary disparities across approximately 96,000 schools in the United States, covering around 32 million white and black students. These associations do not extend to sexuality biases, showing the specificity of the effect. These findings suggest that acknowledging that racial biases and racial disparities in education go hand-in-hand may be an important step in resolving both of these social ills.

education | racial disparities | school discipline | racial bias | social psychology

In comparison with white Americans, black Americans exhibit poorer educational outcomes across a range of metrics. One outcome of particular concern is the gap in disciplinary actions (1, 2). Research using administrative datasets and longitudinal samples clearly show that black American students are far more likely to be suspended or expelled (3, 4) and, conditional on an office referral, more likely to receive stiffer punishments (5, 6). These disparities are particularly concerning as they are associated with long-term outcomes, including employment (7) and involvement in the criminal justice system (8).

As complex social phenomena, racial differences in disciplinary outcomes are multiply determined (2). However, racial bias is thought to be one such determinant. For instance, a controlled experiment using hypothetical vignettes found that in comparison with white students, teachers were more likely to view the same behavior from black students as being indicative of a long-term problem and deserving of suspension (9). Similarly, discipline data from an urban high school showed that black students were especially likely to be referred to the office for discipline on the basis of defiant behavior—a relatively subjective category of misbehavior in comparison with others they examined, including truancy or fighting (10). Overall, there is consistent evidence that black students' behaviors are both perceived as more problematic and are punished more harshly compared with white students. However, to our knowledge, there has been no work assessing whether racial bias is directly associated with disciplinary disparities. Additionally, there has been no work assessing how community-level racial bias is associated with educational disparities.

Psychological measurements of racial bias typically occur through one of two ways. Either individuals are asked to self-report their relative attitudes toward different racial groups (i.e., "explicit bias") or via methods designed to assess automatic associations with people of different races. The "implicit biases" assessed by the latter technique are thought to reflect cognitive and affective response components that are difficult to control. Accordingly, implicit attitudes should overcome some of the social-desirability biases associated with self-report (11). Recently, researchers have begun aggregating these measures up to geographical regions such as counties or states, finding that regional-level measures of implicit and explicit racial bias are associated with racial disparities in key social outcomes, although the relative contributions are not consistent across studies (12–14). For example, one study found that black Americans had reduced access to healthcare and increased rates of death due to circulatory disease in comparison with whites in counties with higher levels of explicit racial bias against blacks (12). They found no such associations for implicit racial bias. In contrast, other work found that the disproportionate use of lethal force by police on black Americans was associated with regional implicit biases but not with explicit biases (14). As such, it is important to assess both types of bias when seeking to understand the relationship between regional-level bias and behavioral outcomes.

Regional levels of bias could be associated with the size of racial student disciplinary disparities for a number of reasons. We highlight several that are likely to be driven by intergroup contact and/or sociopolitical power of the majority group. First, being in an area with elevated racial bias likely means encountering individuals who have negative feelings and beliefs about one's group and whose actions within and/or outside of an educational setting could contribute to disciplinary disparities. For example, if teachers and administrators are biased, then they may be more likely to make decisions that are unfavorable to black students, such as deciding that a given misbehavior is worthy of disciplinary action. Similarly, if members of the community

Significance

Black students in the United States are subject to disciplinary action at rates much higher than their white counterparts. These disciplinary actions put students at higher risk for negative life outcomes, including involvement in the criminal justice system. Using federal data covering over 32 million students at nearly 96,000 schools, our research demonstrates that the disciplinary gap between black and white students across five types of disciplinary actions is associated with county-level rates of racial bias. Our work emphasizes the need for policy targeting racial disparities in education and psychological bias.

are biased, they may more readily perceive transgressions from black students than from white students. The consequences of such interactions may be especially likely to lead to disparate outcomes in high-bias regions when there is ample opportunity for these sorts of intergroup interactions to take place (i.e., intergroup contact is frequent). Second, the norms and structural factors that characterize regions higher in bias may constrain even those individuals who are not biased themselves into engaging in or suborning actions that negatively impact students of color (15, 16). For example, biased administrators or local voters may use their sociopolitical power to support policies that disproportionately punish students of color, such as zero-tolerance policies or implementation of random drug sweeps (17). Additionally, biases assessed at the regional level might reflect affordances of the local environment (e.g., confederate statues, biased media) that undergird these biases and prime behaviors that contribute to disciplinary disparities (18), especially insofar as these affordances reflect the attitudes of the sociopolitically powerful. Overall, these and other reasons, and the likely possibility that they work in concert to inform behavior (19), substantiate the possibility that there will be a relationship between regional bias and disciplinary outcomes.

Most previous research has focused on out-of-school suspensions—likely because they are the most frequently used and are regularly found to be associated with negative outcomes (20, 21). However, other disciplinary outcomes, although used less often, are also likely damaging to students (22). For instance, school arrests have been associated with increased risk of engaging in antisocial behavior (23) and of dropping out (8). In addition, although alternative forms of discipline (e.g., in-school suspension) are intended to insulate students from the negative consequences of exclusionary discipline, the criteria by which students are assigned the former kind of discipline often remain vulnerable to bias (24). As such, examining the presence and basis of disparities in the application of a wide range of disciplinary actions is warranted.

The present analyses combine regionally coded implicit and explicit racial bias measures from approximately 1.6 million respondents who visited Project Implicit (25) with the most recent available data from the Civil Rights Data Collection (CRDC) conducted by the US Department of Education, a mandated census of disciplinary action in all US public kindergarten through grade 12 schools. The CRDC allowed well-powered examinations of five different disciplinary metrics: in-school suspensions, out-of-school suspensions, law enforcement referrals, school-related arrests, and expulsions. We designed the analyses to determine the extent to which regional estimates of pro-white/anti-black implicit and explicit bias are associated with black–white outcomes in disciplinary gaps, and controlled for a range of covariates addressing regional characteristics of the local environment and its population. Additionally, we used analytical techniques to correct for some of the ways in which naïve estimates from nonrandomly sampled Project Implicit users would result in misleading or inaccurate county-level estimates of bias. Our analysis also used regional estimates of sexuality bias from Project Implicit to determine whether demonstrated patterns are distinct to racial bias or an epiphenomenal associate of bias measures in general.

Results

We initially preregistered an analytic plan for this work using the 2013–2014 CRDC data. We departed from our original preregistration in a number of ways. Most of these departures were necessitated based on unforeseen realities of the data. Some differences were created after helpful suggestions during peer review or to remain methodologically congruent with similar works. Consequently, we opted to replicate our analysis using the more recent 2015–2016 data, which became available subsequent

to the preregistration. These analyses are the focus of the main text. Additional analyses, including of the 2013–2014 data, can be found in *SI Appendix*. The preregistration plan, earlier versions of this paper (before analysis of the new data), code, data, and additional details can be found on the project's Open Science Framework (OSF) page at https://osf.io/pu79a/ (26).

Project Implicit Estimates. Before conducting the main analysis, we used multilevel regression with poststratification to make the estimated level of bias for each county more representative of the population (27). After this procedure, the unstandardized county-level estimates show a pro-white bias on both implicit (mean = 0.39; SD = 0.02) and explicit measures (mean = 0.78; SD = 0.17), where on both scales 0 = no bias, and positive numbers indicate a pro-white bias.

Implicit–Explicit Correlation. The poststratified estimates for county-level implicit and explicit bias are highly correlated ($r = 0.79$). This correlation is due to the high level of correlation in the raw county means ($r = 0.26$) and because the state-level predictors for the poststratification multilevel models are similarly associated with implicit and explicit bias (*SI Appendix*, Fig. S1). Consequently, we model the contributions of explicit and implicit bias separately to address issues of collinearity, although *SI Appendix*, Fig. S2 displays estimates from models where both bias measurements are entered as predictors simultaneously, with generally similar findings.

Disciplinary Action Frequency. Table 1 shows the percentage of students of each race reported having received each of the actions under consideration. Additionally, Fig. 1 shows the relative risk ratio for out-of-school suspensions across counties. Interactive maps displaying the relative risk ratios and raw counts of black and white students who are enrolled and receive each type of disciplinary outcome by county can be found at https://osf.io/pu79a/.

Associations Across Counties. We draw inferences based on the posterior distribution of model parameters. Results are reported using the mean of the posterior and 95% highest density interval (HDI). The HDI corresponds to the range of values that are most probable, given the data. Most critically, if the intervals exclude zero, then the data are consistent with a belief that the association is positive or negative, depending on the direction. Parameter estimates for all fixed effects for each model reported in this paper, along with several models testing alternative specifications can be found at https://osf.io/pu79a/.

Fig. 2 shows the estimate of primary interest for each of the models. The estimates displayed are the coefficients for the interaction between race and the two bias measurements. Given that black Americans are the baseline group, negative values for this coefficient indicate that as one moves into counties with higher levels of bias, the gap between the probability of a black student being disciplined and the probability of a white student being disciplined grows. The models illustrated in Fig. 2 indicate that explicit bias is consistently, positively associated with disparities. Across the five disciplinary metrics and the two data

Table 1. Percentage of students of each race receiving each type of disciplinary action

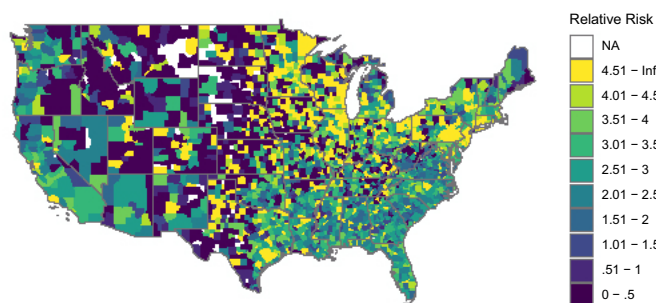| Metric	Black	White
School arrests	0.28%	0.08%
Expulsions	0.51%	0.18%
Law enforcement referral	0.91%	0.34%
In-school suspension	11.22%	4.23%
Out-of-school suspension	13.46%	3.5%

Fig. 1. Relative risk ratio for out-of-school suspensions for each county in the continental United States. Relative risk ratio is computed as the percentage of black students suspended to the percentage of white students suspended, as reported in the 2015–2016 CRDC. Higher values indicate more black students suspended, relative to white students. Counties with the value NA either have no schools have zero black students enrolled or have zero white students enrolled. Interactive maps displaying the relative risk ratios and raw counts of black and white students who are enrolled and receive each type of disciplinary outcome by county can be found on T.R.'s website and at https://osf.io/pu79a/.

collection years, only once is zero included as a probable value (the 2013–2014 estimate for expulsions $est = -0.07$, $[-0.15, 0]$; see *SI Appendix*, Fig. S2).

The estimates for implicit bias tend to be weaker in magnitude, and the HDI includes zero as a probable value for expulsions in the 2015–2016 data. This pattern generally replicates what we find with the 2013–2014 CRDC data, although for that dataset, all estimates for implicit bias include zero within the HDI, with the exception of out-of-school suspensions ($est = -0.04$, $[-0.07, -0.02]$; see *SI Appendix*, Fig. S2). This provides evidence for more robust associations between explicit bias and disciplinary disparities than between implicit bias and disciplinary disparities.

To better illustrate these relationships, Fig. 3 shows the predicted probability of discipline for black and white students for each disciplinary metric as a function of each bias type. For example, for out-of-school suspensions, in a hypothetical county at the mean level of explicit bias (i.e., 0 standardized), the model predicts that 9.6% [9.3%, 9.8%] of black students will be suspended, while the percentage of white students suspended is just 3.3% [3.3%, 3.4%]. If we move to a county one SD above the mean of explicit bias, the predicted percentage of black students suspended increases to 9.8% [9.4%, 10.1%], while the percentage of white students suspended decreases to 3.1% [3%, 3.2%].

Across the country, the average number of black and white students enrolled in a county is 2,484 and 7,862, respectively. Assuming this average-sized county is also at the mean level of explicit bias, the model predicts that 237 black students and 262 white students would receive an out-of-school suspension. If this county were one SD above the mean of explicit bias, the predicted number of suspended black students would grow to 242, while the number of suspended white students would shrink to 246.

Exploratory Analysis: Moderators. We performed a series of analyses to understand what local circumstances covary with either increases or decreases in the relationship between racial biases and disparities in discipline. In these analyses, we selected a subset of our covariates as potential moderators of the effects: county-level black–white segregation, the proportion of the county population that is white, and the county-level racial gap in socioeconomic status. These variables correspond to the two general categories of plausible mechanisms we highlighted in our introduction: direct contact with biased individuals and biased policies or other structural forces. If relationships between racial biases and disciplinary disparities were manifested through some form of direct group contact, then one would expect that areas

characterized by low levels of segregation would show stronger associations than areas with higher levels of segregation. If these relationships are manifest through policies, norms, or environmental affordances, biased individuals would need to hold sociopolitical power to enact regulations or instantiate local norms that would lead to disparate outcomes. To the extent that sociopolitical power is associated with numerical superiority and wealth, one would expect that the association between disciplinary disparities and bias would be stronger in counties with a high proportion of whites and larger racial gaps in socioeconomic status (full details of these analyses are found in *SI Appendix*).

We evaluate these hypotheses through the three-way interaction parameters between a given moderating variable, bias (i.e., level of explicit or implicit) and student race. None of the moderating variables examined consistently yielded interaction effects across disciplinary metrics and years. Just one possible moderating variable–disciplinary metric pairing yielded a consistent interaction across the two data collection years—the proportion of the population that is white and out-of-school suspensions. In other words, in regions with a larger proportion of the population that is white, the association between explicit bias and disparities in out-of-school suspensions is stronger than in regions where whites make up a smaller share of the population. We consider this tentative and partial support for a sociopolitical power hypothesis (see *SI Appendix*, Figs. S4 and S5).

Additional Analyses: Sexuality Bias as Predictor. To test whether the relationships observed above were specific to estimates of racial bias, we ran the same set of analyses with sexuality bias predicting racial disparities in discipline.

Project Implicit Estimates. The unstandardized county-level estimates of bias adjusted with poststratification evidence a pro-straight bias in both implicit (mean = 0.37; SD = 0.05) and explicit measures (mean = 1.53; SD = 0.48), where on both scales 0 = no bias, and positive numbers indicate a pro-straight bias. As with the racial bias estimates, implicit and explicit bias are highly correlated at the county-level ($r = 0.76$) and at the individual level ($r = 0.4$).

Associations Across Counties. *SI Appendix*, Fig. S3 shows the estimate of primary interest for each of the models. The estimates displayed are the coefficients for the interaction between race and each of the two bias measurements. This figure illustrates that in general, county-level explicit and implicit biases in favor of straight individuals are not as consistently associated with racial disciplinary disparities. Just four of the associations between bias and disciplinary outcomes fail to include zero as
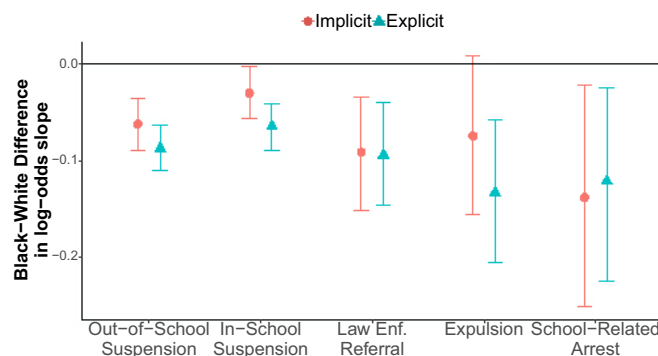

Fig. 2. Association between each metric and county-level estimates of explicit and implicit racial bias. Negative values indicate that there is a positive association between bias and the disciplinary gap. The point is the mean of the posterior, and error bars represent 95% bayesian HDIs.
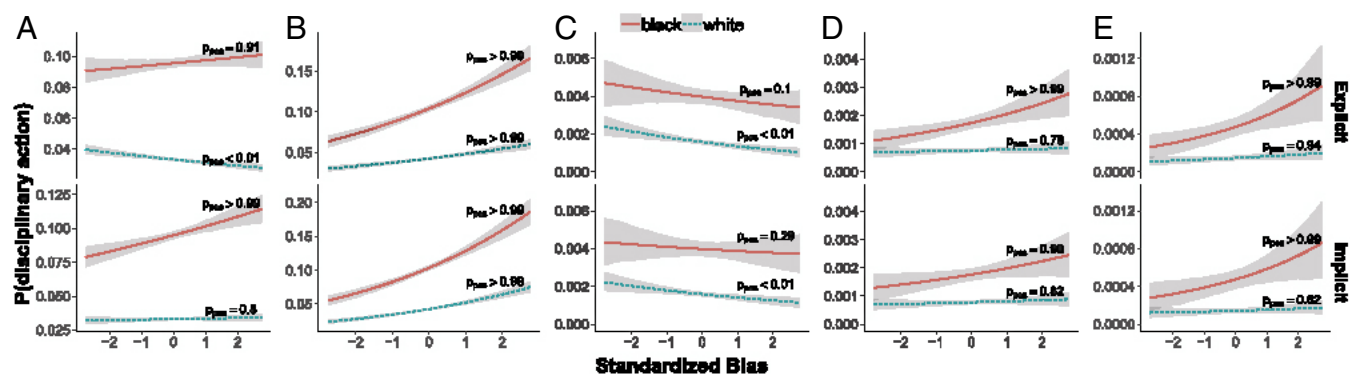
Fig. 3. Association between explicit bias (top row) or implicit bias (bottom row) and the probability of receiving out-of-school suspensions (*A*), in-school suspensions (*B*), law enforcement referrals (*C*), expulsions (*D*), and in-school arrests (*E*) for black and white students. Lines represent the mean of the posterior. Bands indicate 95% uncertainty intervals; p_{pos} represents the probability that the association is positive.

a probable value (explicit bias/in-school suspension, implicit bias/out-of-school suspension, implicit bias/in-school suspension, and implicit bias/expulsion), and the majority of parameters are estimated to be very small. These associations are even weaker in the 2013–2014 data collection, with just two associations failing to exclude zero (implicit bias/out-of-school suspension and implicit bias/law enforcement referrals).

Discussion

These analyses across two separate data collections and five types of disciplinary actions are fully consistent with county-level estimates of racial bias, particularly explicit bias, being associated with racial disciplinary disparities. Specifically, counties with higher rates of explicit biases that favor whites had greater black–white disciplinary disparities across all five outcomes examined. The role of implicit bias is less pronounced. The relationship between implicit bias and disciplinary disparities is also often associated with disciplinary disparities, but the association here is weaker in magnitude and occasionally includes zero as a probable value. In the 2015–2016 data collection, zero is a probable value for the association between implicit bias and expulsions, and in the 2013–2014 data collection, zero is a probable value for all disciplinary actions except for out-of-school suspensions. It should be noted that our analyses cover the vast majority of school-aged students in the United States, and our models include a large set of covariates, suggesting that the relationships between bias and discipline are not due to confounds that can often co-occur with racial disparities, such as socioeconomic status or population demographics. (28, 29).

Our exploratory analysis of possible moderators of these effects suggested that the association between racial bias and disciplinary disparities would be strongest in counties with a large proportion of the population that is white. This is partially consistent with hypotheses whereby disciplinary disparities are realized through the sociopolitical power of white residents who are able to dictate legislation, policies, or norms that contribute to these disparate outcomes, although additional work confirming these tentative hypotheses is needed.

Of course, this research is not without limitations. The strong association between implicit and explicit bias makes drawing comparisons between them somewhat difficult. In general, the evidence appears to suggest that explicit bias is the more consequential of the two, but we hesitate to dismiss the role of implicit bias. Additionally, although we used a poststratification scheme to make our estimates more representative with respect to county age distributions, we cannot account for other ways in which Project Implicit data are not representative of the general population.

The correlational nature of the analyses also presents challenges for interpretation, as it is impossible to definitively establish the causal relationship between bias and disciplinary disparities. The conclusion that explicit biases predict disciplinary disparities is consonant with a great deal of research on disciplinary disparities (30). However, it is possible that living in a region in which black students are disciplined to a greater extent than white students exacerbates and/or reinforces the explicit racial biases of community members or that the relationship between explicit racial biases and disciplinary disparities is bidirectional. It is also possible that some other variable is driving both of these associations (e.g., absence of positive portrayals of African Americans in the media could lead to increased biases in the community and lead teachers to be quicker to discipline black students). Our analyses trade off the ability to ask these more detailed questions about mechanism with the strengths of statistical power and population coverage offered by the large datasets we used here.

Nevertheless, our work compliments other research indicating that racial dynamics are an important source for the observed differences in disciplinary rates between black and white students. For instance, students, caregivers, and administrators perceive suspensions and the disproportionate use of them as at least partially racially motivated (31–33). Additionally, other work has shown that even after controlling for a range of other factors, race remains associated with the likelihood of receiving disciplinary actions (24, 29). Further, experimental evidence using vignettes shows that disciplinary decision-making for teachers differs depending on the race of the student (9). The present research adds to this work by showing associations between disciplinary actions and measurements of racial bias.

Our work also compliments existing studies examining the degree to which community-level implicit and explicit racial biases are associated with racial disparities in key areas, such as health and policing (12–14), by extending this type of inquiry to educational outcomes. To properly assess the meaning of these findings, it is imperative that future work focus specifically on what it means to exist in a community that is estimated to have high or low levels of implicit or explicit bias.

As we have highlighted, students who are subject to the disciplinary actions examined here are at substantially higher risk for negative life outcomes (8). In dispensing discipline differentially across racial groups, educational agencies are also differentially allocating life prospects. Although we cannot make causal claims, the association between racial biases and disciplinary outcomes is worrisome, especially when considered in concert with other literature on race and school discipline.

Riddle and Sinclair

We offer the research presented here to prompt additional scrutiny with respect to how and why educational agencies in the United States differentially administer disciplinary actions, especially when those actions are known to have dire consequences for student welfare. Our work suggests that the mechanisms responsible for these disparities likely exist, at least partly, in the larger community. Through understanding and reducing these disciplinary disparities specifically and the biases that exist in the community more broadly, there exists an avenue for education to maximize its promise as the great equalizer it has the potential to be.

Materials and Methods

Analytic Approach. Analytic details can be found in *SI Appendix*.

Data Sources. Below, we describe the data sources for the 2015–2016 analysis. The data sources for 2013–2014 can be found in *SI Appendix*.

Disciplinary actions. To assess rates of discipline, we used data from the CRDC conducted by the US Department of Education. The dataset comes from the 2015–2016 academic year and has data on "all [local educational agencies] and schools, including long-term secure juvenile justice facilities, charter schools, alternative schools, and schools serving students with disabilities" (ref. 34, p. 6). The CRDC data represent 96,360 institutions enrolling approximately 50.6 million students, of which approximately 24.7 million are white and 7.8 million are black. Previous works using CRDC data have identified a number of districts whose data are in error and have excluded juvenile justice facilities, as these institutions constitute dramatically different educational environments, where the meaning of disciplinary actions may be quite different (35). We followed similar practices, excluding all juvenile justice facilities. Additionally, we excluded data for a specific disciplinary metric for any schools which reported disciplining more students than it reported enrolling for any race for that metric (e.g., a school reported expelling five Asian American students when they reported enrolling less than that number). We also excluded any school for all disciplinary actions if they had an overreporting error for three or more metrics. After these exclusions are applied, the final sample used for modeling consists of 95,827 institutions, enrolling 50.5 million students, of which 24.7 million are white and 7.8 million are black. From these data, we focus on the number of black and white students who were subjected to each of the disciplinary actions described below. It is important to note that black Americans are not the only group subject to disciplinary disparities. In particular, Native Americans, Latinx individuals, and individuals of more than one race are all disciplined at rates greater than white Americans (36). We regret that the scope of the paper is necessarily limited and cannot address similar questions for these groups.

Types of disciplinary actions. We report here on five types of disciplinary actions: in-school suspensions, out-of-school suspensions, law enforcement referrals, school-related arrests, and expulsions.

"Expulsion" is defined as when a student is prohibited from returning to the educational institution for the remainder of the school year or longer. The institution may or may not set up alternative educational services for the student. "Law enforcement referrals" are actions where a student is "reported to any law enforcement agency or official for an incident including a school police unit, for an incident that occurs on school grounds, during school-related events, or while taking school transportation, regardless of whether official action is taken. Citations, tickets, and court referrals are considered referrals to law enforcement" (ref. 37, p. 51). "School-related arrests" refer to "an arrest of a student for any activity conducted on school grounds, during off-campus school activities (including while taking school transportation)" (ref. 37, p. 51). "Out-of-school suspensions" are actions where the student "is temporarily removed from his or her regular school for at least half a day (but less than the remainder of the school year)" (ref. 37, p. 51). Finally, "in-school suspensions" are actions where the student is "temporarily removed from his or her regular classroom(s) for at least half a day..., but remains under the direct supervision of school personnel. Direct supervision means school personnel are physically in the same location as the student under their supervision." (ref. 37, p. 51). We also note that we did not explore instances where a student receives more than one disciplinary action of the same type. Indeed, this an analysis of this type is not possible for any action other than suspensions, as the other outcomes are not identified in this manner by the CRDC. We are also not able to investigate students who receive more than one type of disciplinary action. As such, our analysis cannot, in general, speak to the issue of repeated disciplinary actions.

Racial bias. We used measurements of implicit and explicit bias available from data collected through Project Implicit (25). For a full description of the implicit and explicit bias measures available in these data, refer to (12, 25). We used the Implicit Association Test (IAT) D-score as a measure of implicit bias and the difference between reported warmth toward whites and warmth toward black (both measured from 0 = very cold to 10 = very warm) as a measure of explicit bias, both of which are consistent with previous research on this topic (12). Additionally, we used only respondents who had geographic information that would allow us to place them in a US county, identified as white, and visited the site anytime after it went live in 2002 through the end of 2016. This consisted of approximately 1.6 million total respondents from 3,110 counties; 1.46 million respondents provided data for the IAT, and 1.27 million provided explicit bias ratings.

Sexuality bias. We used measurements of implicit and explicit sexuality bias available from data collected through Project Implicit. We opted for these measures as robustness checks because there were enough observations to closely mimic the racial bias analyses, and they have been previously used for similar purposes (13). We used the IAT D-score as a measure of implicit bias. Our warmth score represented the average of the difference between straight men and gay men and the difference between straight women and lesbian women. We used only respondents who had geographic information that would allow us to place them in a US county and visited the site anytime between when it went live in 2002 through the end of 2016. This consisted of approximately 997,000 total respondents from 3,069 counties. Of these respondents, 890,000 respondents provided data for the IAT, and 948,000 respondents provided explicit bias ratings.

Covariates. Each county-level variable used as a covariate in the final model and the corresponding state-level variable used as a predictor in the poststratification scheme (described below) were taken from the same source. Population size and proportions, socioeconomic indicators, mobility, and segregation indices were all taken from the American Community Survey (ACS, https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml) 5-y estimates for the time period ending in 2016. Urban–rural indicators were taken from the 2010 US Census, and crime rates were taken from the Federal Bureau of Investigation Uniform Crime Reporting program (https://www.icpsr.umich.edu/icpsrweb/NACJD/discover-data.jsp), as made available through the National Archive of Criminal Justice Data for each year from 2012–2014 and 2016 (we were unable to locate data for 2015). Each of these variables is described below.

Population Size and Proportions. We obtained the total population, the proportion of the population that is white, the proportion of the population that is black, and the ratio of black-to-white people in the population from ACS table B02001 (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml).

Socioeconomic Indicators. For the state-level socioeconomic predictors for the poststratified estimates, we obtained estimates for the percentage of population with a bachelor's degree or higher, the percentage of the population aged 16 y or over in the labor force that is unemployed, the median household income, and the percentage of families and people whose income in the last year was below the poverty line from the ACS table DP03 (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml). For the county-level data, we used the same set of variables described above and also included estimates of the difference between blacks and whites in each county on these socioeconomic indicators, using multiple imputation to fill in observations for some counties where estimates were unavailable.

Urban–Rural Indicator. We obtained estimates of housing density per square mile of land area from Census table GCT-PH1 (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml).

Mobility. We obtained estimates of population mobility by summing the percentage of black Americans who moved from a different county, state, or country into the county of interest (county-level covariate) or who moved from a different state or country into the state of interest (state-level covariate). We took these metrics from the ACS table S0701 (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml).

Crime. We computed estimates of the number of violent crimes per person by taking the number of crimes reported divided by the population size for each year and averaging the resulting proportions across the 4 y of data (https://www.icpsr.umich.edu/icpsrweb/NACJD/discover-data.jsp).

Segregation. We computed a dissimilarity index as described by ref. 38. This metric reflects the proportion of a racial group within the

county that would have to move in order for all census tracts to have group distributions that matched the overall distribution of the county. These computations were done using data from the ACS table B02001 (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml).

1. Kinsler J (2011) Understanding the black–white school discipline gap. *Econ Educ Rev* 30:1370–1383.
2. Okonofua JA, Walton GM, Eberhardt JL (2016) A vicious cycle: A social–psychological account of extreme racial disparities in school discipline. *Perspect Psychol Sci* 11: 381–398.
3. Aud S, KewalRamani A, Frohlich L (2011) *America's Youth: Transitions to Adulthood* (National Center for Education Statistics, Washington, DC), NCES 2012-026.
4. Yeager DS, Purdie-Vaughns V, Hooper SY, Cohen GL (2017) Loss of institutional trust among racial and ethnic minority adolescents: A consequence of procedural injustice and a cause of life-span outcomes. *Child Dev* 88:658–676.
5. Gregory JF (1995) The crime of punishment: Racial and gender disparities in the use of corporal punishment in US public schools. *J Negro Educ* 64:454–462.
6. Skiba RJ, et al. (2011) Race is not neutral: A national investigation of african american and latino disproportionality in school discipline. *Sch Psychol Rev* 40:85–107.
7. Pager D, Western B, Sugie N (2009) Sequencing disadvantage: Barriers to employment facing young black and white men with criminal records. *Ann Am Acad Polit Soc Sci* 623:195–213.
8. Hirschfield P (2009) Another way out: The impact of juvenile arrests on high school dropout. *Soc Educ* 82:368–393.
9. Okonofua JA, Eberhardt JL (2015) Two strikes: Race and the disciplining of young students. *Psychol Sci* 26:617–624.
10. Gregory A, Weinstein RS (2008) The discipline gap and african americans: Defiance or cooperation in the high school classroom. *J Sch Psychol* 46:455–475.
11. Rudman LA (2004) Sources of implicit attitudes. *Curr Dir Psychol Sci* 13:79–82.
12. Leitner JB, Hehman E, Ayduk O, Mendoza-Denton R (2016) Blacks' death rate due to circulatory diseases is positively related to whites' explicit racial bias: A nationwide investigation using project implicit. *Psychol Sci* 27:1299–1311.
13. Orchard J, Price J (2017) County-level racial prejudice and the black-white gap in infant health outcomes. *Soc Sci Med* 181:191–198.
14. Hehman E, Flake JK, Calanchini J (2017) Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Soc Psychol Personal Sci* 9:393–401.
15. Jetten J, Spears R, Manstead ASR (1996) Intergroup norms and intergroup discrimination: Distinctive self-categorization and social identity effects. *J Personal Soc Psychol* 71:1222–1233.
16. Pettigrew TF (1958) Personality and sociocultural factors in intergroup attitudes: A cross-national comparison. *J Conflict Resolut* 2:29–42.
17. Verdugo RR (2002) Race-ethnicity, social class, and zero-tolerance policies: The cultural and structural wars. *Educ Urban Soc* 35:50–75.
18. Keith Payne B, Vuletich HA, Lundberg KB (2017) The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychol Inquiry* 28:233–248.
19. Rose Markus H, Kitayama S (2010) Cultures and selves: A cycle of mutual constitution. *Perspect Psychol Sci* 5:420–430.
20. Evans Cuellar A, Markowitz S (2015) School suspension and the school-to-prison pipeline. *Int Rev Law Econ* 43:98–106.
21. Matjasko JL (2011) How effective are severe disciplinary policies? School policies and offending from adolescence into young adulthood. *J Sch Psychol* 49:555–572.
22. Karega Rausch M, Skiba RJ, Simmons AB (2005) The academic cost of discipline: The relationship between suspension/expulsion and school achievement. *Annual Meeting of the American Educational Research Association*. Available at http://www.indiana.edu/~atlantic/wp-content/uploads/2014/03/Academic-Cost-of-School-Discipline.pdf. Accessed March 20, 2019.
23. Hemphill SA, Toumbourou JW, Herrenkohl TI, McMorris BJ, Catalano RF (2006) The effect of school suspensions and arrests on subsequent adolescent antisocial behavior in Australia and the United States. *J Adolesc Health* 39:736–744.
24. Anyon Y, et al. (2014) The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Child Youth Serv Rev* 44:379–386.
25. Xu K, Nosek B, Greenwald A (2014) Psychology data from the race implicit association test on the project implicit demo website. *J Open Psychol Data* 2:e3.
26. Riddle T (2019) Data from "Disparities in educational discipline." Open Science Framework. Available at https://osf.io/pu79a/. Deposited December 18, 2017.
27. Park DK, Gelman A, Bafumi J (2004) Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Polit Anal* 12:375–385.
28. Williams DR (1999) Race, socioeconomic status, and health the added effects of racism and discrimination. *Ann New York Acad Sci* 896:173–188.
29. Skiba RJ, Michael RS, Carroll Nardo A, Peterson RL (2002) The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Rev* 34:317–342.
30. McIntosh K, Girvan EJ, Horner RH, Smolkowski K (2014) Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *J Appl Res Child Inform Policy Child Risk* 5:4.
31. Haight W, Gibson PA, Kayama M, Marshall JM, Wilson R (2014) An ecological-systems inquiry into racial disproportionalities in out-of-school suspensions from youth, caregiver and educator perspectives. *Child Youth Serv Rev* 46:128–138.
32. Gibson PA, Haight W (2013) Caregivers' moral narratives of their African American children's out-of-school suspensions: Implications for effective family–school collaborations. *Soc Work* 58:263–272.
33. Ruck MD, Wortley S (2002) Racial and ethnic minority high school students' perceptions of school disciplinary practices: A look at some canadian findings. *J Youth Adolesc* 31:185–195.
34. US Department of Education (2018) *2015–2016 Civil Rights Data Collection Public-Use Data File User's Manual* (US Department of Education Office for Civil Rights, Washington, DC).
35. Losen DJ, et al. (2015) *Are We Closing the School Discipline Gap?* Available at https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/are-we-closing-the-school-discipline-gap/AreWeClosingTheSchoolDisciplineGap_FINAL221.pdf. Accessed February 27, 2019.
36. United States Government Accountability Office (2018) *Discipline Disparities for Black Students, Boys, and Students with Disabilities. Technical Report* (United States Government Accountability Office, Washington, DC).
37. National Center for Education Statistics (2014) *2013–2014 Civil Rights Data Collection Survey* (National Center for Education Statistics, US Department of Education, Washington, DC).
38. Massey DS, Denton NA (1988) The dimensions of residential segregation. *Soc Forces* 67:281–315.