Article

Social Psychological and Personality Science
2020, Vol. 11(4) 492-498

DOI: 10.1177/1948550619875150
journals.sagepub.com/home/spp
SAGE

School Deferred: When Bias Affects School Leaders

Shoshana N. Jarvis[1,2] and Jason A. Okonofua[1]

Abstract

In the classroom, Black students are disciplined more frequently and more severely for the same misbehaviors as White students. Though teachers have influence over disciplinary actions, the final decisions for exclusionary discipline (i.e., suspensions and expulsions) are principals' responsibility. We test how principals make disciplinary decisions in a preregistered experiment. Principals endorsed more severe discipline for Black students compared with White students across two time points. Further, this discipline severity was explained through Black students being more likely to be labeled a troublemaker than White students. Future efforts should focus on principals in order to mitigate the negative impacts of the school-to-prison pipeline.

Keywords

stereotyping, school discipline, racial disparities, principals

School suspension rates have nearly tripled during the span of the past 40 years and have become one of the leading issues that plague school outcomes in the United States (Children's Defense Fund, 1974; U.S. Department of Education Office for Civil Rights, 2016). Suspensions have large and lasting impacts for students because with each one, students are removed from the learning environment. One study found that students who were suspended fell behind two grade levels in their reading across 2 years—in addition to how far behind they already were (Arcia, 2006). The more students fall behind, the less likely they are to graduate from high school or enter college and the more likely they are to become incarcerated, a growing concern called the "school-to-prison pipeline" (Fabelo et al., 2011; Rocque & Paternoster, 2011). Further, recent research shows that suspensions are not only consequential for the suspended student but also cost taxpayers millions of dollars in the long run (Rumberger & Losen, 2017).

Adding complexity to this societal issue, there are large racial disparities in suspension rates. Recent national statistics show that Black students are 3.6 times more likely to be suspended than White students. In fact, 1.1 million of the 2.8 million students suspended in the 2013–2014 school year were Black (U.S. Department of Education Office for Civil Rights, 2016). As a consequence, Black students are disproportionately present in the school-to-prison pipeline.

What might cause such large disparities in discipline? Most relevant research is observational or archival in nature. Many studies have analyzed school-level data to uncover factors related to suspensions and why the rates are disproportionate by race. This body of work found that school characteristics (e.g., size and demographic population) and nonbehavioral student characteristics (e.g., race) are better predictors of suspensions than a student's behavior (Skiba et al., 2014; Skiba, Peterson, & Williams, 1997; Wu, Pink, Crain, & Moles, 1982). Further, a student's race is a better predictor of discipline than other student characteristics (e.g., socioeconomic status; Skiba, Michael, Nardo, & Peterson, 2002).

Psychological research provides theory to help explain how race contributes to inequality in outcomes. Stereotypes can shape judgment and decision-making and are most likely to do so when a person must make subjective interpretations of ambiguous information (Corning & Bucchianeri, 2010; see Eberhardt, 2019 for review), as seen with disciplinary decisions that lead to suspensions (Okonofua, Walton, & Eberhardt, 2016). Archival research shows that racial disparities are not observed for clear or objective reasons for suspensions (e.g., weapon or drug possession). Rather, Black students are more likely than White students to be suspended for misbehaviors that involve more subjective interpretation such as "disrespect" or "insubordination" (Skiba et al., 2002). When 70% of all expulsions in a large U.S. state were for disruption or other subjective misbehaviors (Rausch & Skiba, 2004), this signals that bias may have a role in who is suspended.

[1] Department of Psychology, University of California, Berkeley, CA, USA
[2] Haas School of Business, University of California, Berkeley, CA, USA

Corresponding Author:
Shoshana N. Jarvis, Haas School of Business, University of California, Berkeley, 2220 Piedmont Ave, Berkeley, CA 94720, USA.
Email: sjarvis@berkeley.edu

Recent research suggests racial disparities in discipline arise over time. Okonofua and Eberhardt (2015) asked a large and diverse sample of teachers to provide information about how they would respond to a series of misbehaviors. Teachers were randomly assigned to read that the misbehavior was by a White or Black student. Teachers endorsed disciplining the student more severely if the student was Black as compared to White. The observed racial disparities were largely explained by teachers' increased likelihood to label the Black student as a troublemaker. This initial experiment highlights the process by which stereotypes can shape the beginning of the disciplinary process. Might a similar psychological mechanism affect other points in the disciplinary process?

A teacher's interpretation of student behavior is pivotal in the initiation of discipline. Therefore, it is useful to understand what impacts their disciplinary decisions. However, a teacher's interpretation is not the only perspective that matters, and teachers do not make final decisions for which students get suspended or not. When a teacher determines that a student misbehaved, the teacher can refer the student to a principal for disciplinary action. Principals or assistant principals ultimately decide whether or not a student will be suspended from school (Dunbar & Villarruel, 2002; Mukuria, 2002). A principal's perspective differs from that of a teacher in two key ways. First, principals tend not to see a student's behavior firsthand. Rather, they must rely on a secondhand account of the incident. Their interpretation may be even more susceptible to effects of stereotypes due to this increased distance and ambiguity. Second, principals bear the responsibility of considering the safety of the entire school rather than the confines of a particular classroom environment. It is crucial to consider the perspective of principals in the disciplinary process in order to gain a comprehensive understanding of how racial disparities in suspension rates occur. The present research investigates the principal's perspective in the disciplinary process.

So what role might principals play in the large racial disparities in suspensions? The present research is the first to employ an experimental design to address this question. Despite their alternative perspective, principals' disciplinary decisions might be shaped by stereotypes in a similar way to teachers. Most people are exposed to the same negative stereotypes of Black people (Weisbuch, Pauker, & Ambady, 2009). Thus, principals may also tend to view a misbehaving Black student as a troublemaker and endorse greater discipline severity. If so, the effects of stereotypes might compound at multiple levels of the disciplinary process.

The current research investigates whether the racial stereotyping processes that affects teachers' roles in the disciplinary process might also apply to principals' roles in the process by testing how race impacts disciplinary decisions among assistant principals across two misbehaviors. Assistant principals read teacher referrals about a male student whose race was manipulated to be either White or Black by their name. Past research confirms this procedure signals race in a subtle yet effective manner (Okonofua & Eberhardt, 2015). They rated how severely they would discipline the student across two misbehaviors and responded to questions about the characteristics and future of the student. This between-subjects experimental design allows for control over the content of the misbehaviors and for causal claims between the race of the student and subsequent discipline.

For this study, there were five preregistered hypotheses. (1) Black students would be disciplined more severely after a second misbehavior compared to White students for the same misbehaviors. Further, the escalation in response between the two misbehaviors would be sharper for Black students than for White students. (2) Due to the labeling process observed in past studies, Black students, as compared to White students, would be more likely to be considered a troublemaker after two misbehaviors and (3) have their behavior more likely to be perceived to be a pattern. (4) This would manifest itself in harsher punishments for Black students compared to White students in days of detentions. (5) Finally, replicating past research, racial differences in discipline administered would be mediated by the extent to which the principal perceived the student to be a troublemaker and that their misbehavior was indicative of a pattern of misbehavior. The procedure for this study was preregistered (https://osf.io/tc47c/), and all materials and syntax are publicly available (https://osf.io/evtze/). We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Method

Participants and Design

A sample of public middle and high school assistant principals (91) was recruited from a large school district in a Southeastern state. In this district, assistant principals make disciplinary decisions, so the recruited sample was limited to assistant principals, though this distinction may vary between districts. Across 21 middle schools and 18 high schools, this district is responsible for overseeing almost 50,000 students and has a 56% White, 19% Black, 16% Hispanic, 5% Asian, 4% multiracial, and <1% Native American student population. Three participants were excluded due to preregistered exclusion criteria: taking longer than 20 min to complete ($N = 1$), not being a principal or assistant principal ($N = 1$), and not passing the attention check ($N = 1$), and three principals did not respond to any questions leaving a final sample of 85 principals.[1] On average, this sample had 7 years ($SD = 5$ years) of experience as an administrator. Race and gender demographic information of the principals were not collected to protect their anonymity, and school demographics were not able to be connected to principals. The population of principals the sample was drawn from was 55% women and 45% men and was 73% White, 24% Black, 2% Latinx, and 1% Asian.

The design of the study was a 2 (race: Black vs. White student) between-subjects and 2 (time of the misbehavior: first vs. second) within-subjects mixed design. Principals were recruited during the school year while they were interacting

with students. Our recruitment aims were to collect data from all of the assistant principals in this school district as part of a collaboration with the school district. Principals were given no incentive to participate, and they all completed the survey in the same room on the same afternoon during a weekday. We collected data from 79% of the principals in the district, and data collection ended when all who were willing to consent participated in the study.

Procedure

The procedure was adapted from previous studies with teachers (Okonofua & Eberhardt, 2015). The principals were shown an image of a school that was described as being particularly average and were asked to imagine that they were a principal at that school on a regular day. A narrative about a student was presented. The male student either had a typically Black name (e.g., Darnell) or typically White name (e.g., Greg; Greenwald, McGhee, & Schwartz, 1998). The participant read a referral by a teacher about the student misbehaving in a classroom.

Darnell/Greg is constantly disrupting the class environment by strolling around the classroom at random intervals, getting tissues from the tissue box multiple times during a 50 minute class, throwing items away constantly; in general, Darnell/Greg circulates around the room and up and down the rows to see what other students are doing, have eyes on him, and disrupt the flow of the lecture or activity the class was participating in.

They then rated the extent to which the behavior was severe, hindered the teacher's ability to maintain order in the classroom, how irritated they felt, and how severely they would discipline the student. The narrative continued with the student misbehaving again 3 days later and was referred to the principal.

> Darnell/Greg is sleeping in class. You tell him to pick his head up and get to work. He only picks his head up. He chooses to rest it on his hand and continue to sleep. So you ask him one more time and again, Darnell/Greg refuses to do work. You ask him to leave class and go to the office to tell them that he won't do his work and chose to sleep instead. He refuses to do this as well.

The principal gave the same ratings as with the previous misbehavior incident as well as the following dependent variables: (1) the likelihood the student was a troublemaker, (2) the likelihood the misbehavior was indicative of a pattern, (3) the extent the principal would worry that this would become a pattern across students, (4) how many days they would send the student to detention, and (5) the likelihood the student would be suspended in the future. Misbehaviors were designed from collected school records and featured two incidents subjective in nature—insubordination and classroom disruption. All participants saw both misbehaviors, the order of which was randomly counterbalanced.

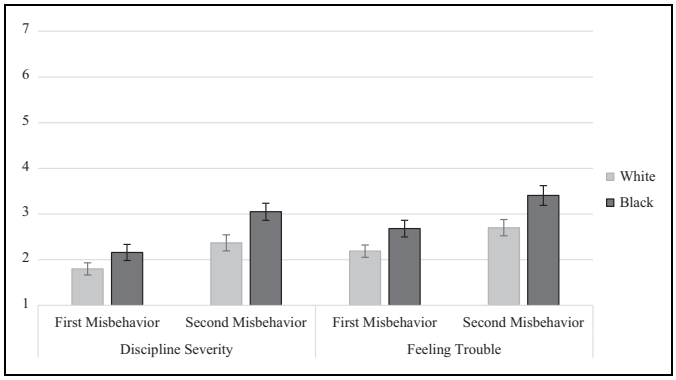

Figure 1. Mean ratings of how severely they felt students should be disciplined (left) and how troubled teachers felt by students' misbehavior (right) as a function of times misbehaved and race. Error bars represent standard errors.

Results

As a manipulation check, principals rated how likely they thought the student they had read about was Black. Principals who read about a student named Darnell were more likely to think the student was Black than the principals who read about a student named Greg, $t(82) = 2.80$, $p = .006$. Due to the overrepresentation of Black people in low-income neighborhoods, principals also rated the likelihood the student as from a low-income neighborhood. This was to test whether principals extrapolated class from the racial information and to isolate the effects of race from those of perceived social class. There were no differences in perceptions of being from a low-income neighborhood by the race of the student, $t(81) = 0.28$, $p = .778$.

A 2 (race: Black vs. White) × 2 (time: Time 1 vs. Time 2) mixed analysis of variance (ANOVA) was conducted on the discipline severity ratings.[2] There was a main effect of race, $F(1, 82) = 7.71$, $p = .007$, $d = 0.61$, indicating Black students were given more severe discipline compared to their White classmates (see Figure 1). There was also a main effect of time, $F(1, 82) = 23.87$, $p < .001$, $d = 0.60$, indicating, from the first incident to the second, discipline severity increased for all students. We did not observe the predicted Race × Time interaction, $F(1, 82) = 1.19$, $p = .278$. Principals rated Black students more severely than White students, and discipline increased across misbehaviors for all students. While the pattern of results was directionally the same as the hypothesized interaction, this did not reach statistical significance.

The extent to which principals rated the behavior as severe, hindered the teacher's ability to maintain order in the classroom, and how irritated they felt were all highly correlated for each misbehavior (Time 1: rs > .46; Time 2: rs > .62). Consistent with past studies (Okonofua & Eberhardt, 2015), these three responses were averaged at each time point to create an aggregated score of the principal's affective response called "feeling troubled" (Time 1: $\alpha = .74$; Time 2: $\alpha = .89$).

Consistent with discipline severity, a 2 (race: Black vs. White) × 2 (time: time 1 vs. time 2) mixed ANOVA was

conducted on the feeling troubled ratings. There was a main effect of race, $F(1, 80) = 9.56$, $p = .003$, $d = 0.62$, and a main effect of time, $F(1, 80) = 13.15$, $p = .001$, $d = 0.47$. There was no significant Race × Time interaction, $F(1, 80) = 0.39$, $p = .536$. Principals rated Black students more harshly at both time points, and every student had an escalation in endorsed discipline severity between the two time points.

How did these misbehaviors impact how principals viewed the students? When asked how likely the student was a troublemaker, principals were more likely to rate Black students ($M = 1.93$, $SD = 0.88$) than White students ($M = 1.54$, $SD = 0.60$) to be a troublemaker, $t(82) = 2.38$, $p = .020$, $d = 0.52$. Principals rated the extent to which they thought the behaviors of the students was reflective of a pattern. There was no difference between Black students ($M = 2.84$, $SD = 1.17$) and White students ($M = 2.48$, $SD = 1.15$) in how likely their behavior was interpreted as being a pattern, $t(83) = 1.43$, $p = .156$, $d = 0.31$. However, the pattern of the results was in the expected direction with Black students being more likely than White students to have their behavior be interpreted as being reflective of a pattern. In an exploratory analysis, when asked the extent to which they worried that the student's misbehavior would become a pattern across other students in the classroom, there was a marginal effect indicating Black students' misbehavior ($M = 1.93$, $SD = 0.86$) were more likely to be seen as a catalyst compared to White students' misbehavior ($M = 1.64$, $SD = 0.63$), $t(80) = 1.73$, $p = .088$, $d = 0.39$.

Next, principals responded to how they would discipline the students. When asked how many days of detention the students should receive after the two misbehaviors, Black students ($M = 2.33$, $SD = 0.87$) were on average given more days of detention than White students ($M = 1.88$, $SD = 0.60$), $t(82) = 2.74$, $p = .007$, $d = 0.61$. Principals were not more likely to assign Black students any detention compared to their White classmates, $\chi^2(1, N = 84) = 0.86$, $p = .422$, but the pattern of results was consistent with the hypothesis (White: $N_{no\ detention} = 10$, $N_{detention} = 31$; Black: $N_{no\ detention} = 7$, $N_{detention} = 36$). Principals indicated how likely they thought they would need to suspend the student in the future. The pattern of results was in the expected direction though marginal with Black students ($M = 1.63$, $SD = 0.69$) more likely to be suspended in the future compared with their White classmates ($M = 1.38$, $SD = 0.49$), $t(81) = 1.91$, $p = .060$, $d = 0.42$.

We hypothesized that the extent to which principals thought the student was a troublemaker and the extent to which their behavior was indicative of a pattern would act as a mediating variable between the link of the race of the student and subsequent disciplinary decisions. Mediation models with 10,000 bootstrapped samples using the Lavaan package in R Version 0.6-3 (Rosseel, 2012) were run to test these hypotheses. Because there was not a significant effect of race on behavioral pattern ratings, mediation models testing this pathway were not run.

First, we tested whether the extent to which principals thought a student was a troublemaker mediated the relationship between the race of the student and how severely principals disciplined students after the second misbehavior. The predicted indirect path through labeling the student as a troublemaker was significant, $b = 0.37$, $SE = 0.16$, $p = .023$, 95% CI [0.05, 0.69], rendering the direct effect nonsignificant, $b = 0.31$, $SE = 0.21$, $p = .131$, 95% CI [−0.09, 0.72], see Table 1. Next, we tested whether the extent to which principals thought a student was a troublemaker mediated the relationship between the race of the student and how many days of detention the principal gave the student. The predicted indirect path through labeling the student as a troublemaker was significant, $b = 0.15$, $SE = 0.07$, $p = .041$, 95% CI [0.01, 0.29], while the direct effect was attenuated, $b = 0.33$, $SE = 0.15$, $p = .034$, 95% CI [0.02, 0.63], see Table 1. Lastly, we tested whether the extent to which principals thought a student was a troublemaker mediated the relationship between the race of the student and how likely the student would be suspended in the future. The predicted indirect path through labeling the student as a troublemaker was significant, $b = 0.15$, $SE = 0.07$, $p = .027$, 95% CI [0.02, 0.28], rendering the direct effect nonsignificant, $b = 0.12$, $SE = 0.12$, $p = .322$, 95% CI [−0.12, 0.36], see Table 1. Black students were more likely to be seen as troublemakers, and these higher ratings for Black students predicted harsher discipline after the second misbehavior, more days of detention, and a greater likelihood that they would be suspended in the future.

Table 1. Path Analysis Estimates for the Effect of Race on Discipline (i.e., Discipline Severity, Days of Detention, and Likelihood of Suspension) via the Likelihood the Student is a Troublemaker.

Effect	Estimate	*SE*	*p*	95% CI
Discipline severity				
Race → Troublemaker (*a*)	.39	.16	.016	[0.07, 0.71]
Troublemaker → Discipline severity (*b*)	.94	.13	<.001	[0.68, 1.20]
Total effect (*c*)	.68	.25	.007	[0.19, 1.17]
Direct effect (*c′*)	.31	.21	.131	[−0.09, 0.72]
Indirect effect (*ab*)	.37	.16	.023	[0.05, 0.69]
Detention				
Race → Troublemaker (*a*)	.41	.17	.014	[0.08, 0.73]
Troublemaker → Detention (*b*)	.37	.10	<.001	[0.17, 0.56]
Total effect (*c*)	.48	.16	.003	[0.16, 0.79]
Direct effect (*c′*)	.33	.15	.034	[0.02, 0.63]
Indirect effect (*ab*)	.15	.07	.041	[0.01, 0.29]
Suspension				
Race → Troublemaker (*a*)	.42	.16	.010	[0.10, 0.74]
Troublemaker → Suspension (*b*)	.35	.08	<.001	[0.20, 0.51]
Total effect (*c*)	.27	.13	.041	[0.01, 0.53]
Direct effect (*c′*)	.12	.12	.332	[−0.12, 0.36]
Indirect effect (*ab*)	.15	.07	.027	[0.02, 0.28]

Discussion

After viewing the same misbehavior by either a White or Black student, principals viewed misbehavior more negatively and

endorsed more severe discipline for Black students as compared with their White classmates. (1) We predicted principals would endorse more severe discipline for Black students after a second misbehavior compared with White students. We found support for this hypothesis—principals endorsed more severe discipline for Black students after the second misbehavior compared with White students. Because misbehaviors were held constant, we can conclude it is due to the student's perceived race and not aspects of the misbehavior. We predicted the escalation in response between the two misbehaviors would be sharper for Black students than for White students. There was no evidence of an escalation in response among principals, likely because they endorsed more severe discipline for Black students at both time points unlike studies with teachers (Okonofua & Eberhardt, 2015). This suggests that the process for discipline severity may not be the same for principals as it is for teachers.

(2) Supporting our predictions, Black students were more likely to be considered troublemakers compared to White students. (3) However, there were no significant differences between Black and White students for how likely their misbehaviors were considered to be a pattern. (4) We predicted that Black students would receive harsher punishments than White students. Black students were given more days of detention compared to White students, and in exploratory analyses, Black students were marginally more likely to be expected to be suspended in the future compared to White students.

(5) We predicted that the racial differences in discipline would be mediated by the extent to which the principal perceived the student to be a troublemaker and that their misbehavior was indicative of a pattern of misbehavior. Mediation pathways for the pattern of misbehavior item as a mediator were not tested because the direct pathway was not significant. However, the extent to which principals thought the student was a troublemaker was a significant mediator for all three disciplinary outcomes. That is, the mediated pathway explained a significant portion of the variance in the causal relationship between the race of the student and subsequent discipline, and in two of three cases, rendered the nonmediated path nonsignificant.

One of our key findings is that the results observed for principals were different than what past research observed with teachers (Okonofua & Eberhardt, 2015). While past research shows teachers endorse disciplining Black students more severely than White students only after repeated misbehavior, principals endorsed more severe discipline for a Black student, as compared to a White student, after a single misbehavior. And for principals, the difference in discipline severity persisted after further misbehavior. Differences in discipline could be due to of the differences in the roles that teachers and principals play in the school environment. Principals may assume that if a student is sent to their office, the behavior has reached a particular threshold of unacceptability or difficulty to manage. Thus, the decision point in which principals display discrimination is in the determination of whose behavior is so disruptive that they cannot return to the classroom. Appraisals for how likely the student's behavior is a pattern may matter less simply because as a matter of reaching the principal's office they have displayed a semblance of a pattern. While teachers wait before escalating their responses, principals may be less likely to afford such leniency to Black students.

Principals play a key role for setting the procedural and cultural norms for their school. Rather than the purview over an approximately 20-person class, a principal's decision can affect the entire school. The policies principals enact impact the extent to which students are disciplined—and consequently future outcomes—for the entire student population. A priority should be to keep students in school as much as possible and to work with the students to keep them in school. Principals have the power to make this happen.

With use of an experimental design, this study advances previous theory about principal's disciplinary decisions by controlling the misbehaviors so that they remain consistent across the students. Thus, any racial differences observed were due to a student's race rather than objective interpretations of misbehavior. By using a between and within mixed design, we increased our power to detect effects by gauging principals' responses to multiple incidents. We were also able to rigorously test racial effects within one microenvironment by testing all of the principals from one school district in which there are racial disciplinary disparities. In this district, Black students are 3.5 times more likely than White students to be suspended.[3]

While we were fortunate to be able to test the majority of principals from an entire school district, this necessitates replications to test generalizability across the country. More research is necessary to discover what aspects of a district will make it more likely to exhibit racial disparities in discipline, such as racial composition, location, willingness to express racial bias, and size. Smaller districts, unlike the one in this study, have as few as one administrator. Differences could be observed by the number of administrators and how discrete their responsibilities are. Restricting the sample to one school district limited the sample size and only provided sufficient power to detect relatively large effects. It is possible that principals discipline Black students more severely over time, and future research with a larger sample size is necessary to test this effect. The reported data only include assistant principals from middle and high schools yet racial disparities are pervasive across grade levels (Office for Civil Rights, 2016). Suspensions in middle and high school are especially consequential for students (Rocque & Paternoster, 2011). Thus, the labeling process found in the present work is indeed noteworthy. Yet, additional research should test how labeling processes impact elementary school-level principals.

The present study is one of a few to highlight a labeling process as a key explanation for how race disparities in discipline can occur. This process was replicated with principals and represents a key psychological process that propagates racial disparities in discipline in schools. Future interventions should consider targeting the labeling process in principals—in addition to teachers (see Okonofua, Paunesku, & Walton, 2016) —as a means to mitigate the racial disparities in discipline.

Research with principals is necessary to establish a better understanding of how the school-to-prison pipeline might be perpetuated across over time. Principals serve as important gatekeepers, providing or limiting access to school, and have a great influence over policies and procedures for large groups of students. Studying where and how principals perpetuate racial disparities will be an important step in creating steps to eliminate the school-to-prison pipeline.

Acknowledgments

The authors would like to acknowledge Sepi Hejazi Moghadam, Head of Special Projects, University Relations at Google for his support and suggestions for this work. The authors would like to thank the school district for partnering on this research, Rodolfo Mendoza-Denton, Stephen Antonoplis, Felicia Zerwas, and Amanda Perez for their comments on previous drafts of this article, and the EDENS lab and Eden Adaora Okonofua.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The first author is supported by a National Science Foundation Graduate Research Fellowship. This work is supported by Google's Computer Science Education Research team, The Tides Foundation, and the Character Lab.

ORCID iD

Shoshana N. Jarvis https://orcid.org/0000-0001-6980-3555

Notes

1. A preregistered exclusion criterion was to exclude participants for being suspicious of this study being used to study racial bias. The district was planning to begin more restorative justice programs in their school. As a result, before the study, principals were concerned about being identified as biased, and demographic information was not collected to assuage this fear. Despite this precaution, 20% (18 assistant principals) indicated this study was being used to test bias. Therefore, participants were not excluded according to this criterion. This investigation reflects a conservative test as it is likely participants underreported discipline to avoid seeming biased.
2. In accordance with our preregistration, we excluded responses that were 2.5 standard deviations above or below the mean for each dependent variable. We calculated *z* scores for all responses and filtered out extreme responses individually for each analysis. For within-subjects variables, if a participant's response was an outlier for one time point, both responses were excluded from the analysis. Thus, degrees of freedom may vary across dependent variables. See the Open Science Framework for analyses without exclusions https://osf.io/9v7b4/.
3. This number was calculated using data downloaded from the Department of Education Office for Civil Rights for the 2013–2014 school year.

References

Arcia, E. (2006). Achievement and enrollment status of suspended students: Outcomes in a large, multicultural school district. *Education and Urban Society*, *38*, 359–369. doi:10.1177/0013124506286947

Children's Defense Fund. (1974). *Children out of school in America*. Washington, DC: Washington Research Project.

Corning, A. F., & Bucchianeri, M. M. (2010). Perceiving racism in ambiguous situations: Who relies on easy-to-use information? *The Journal of Social Psychology*, *150*, 258–277. doi:10.1080/00224540903365398

Dunbar, C., Jr., & Villarruel, F. A. (2002). Urban school leaders and the implementation of zero-tolerance policies: An examination of its implications. *Peabody Journal of Education*, *77*, 82–104. doi:10.1207/S15327930PJE7701_5

Eberhardt, J. L. (2019). *Biased: Uncovering the hidden prejudice that shapes what we see, think, and do*. New York: Viking.

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. Retrieved from Council of State Governments Justice Center website: https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=266653

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464. doi:10.1037/0022-3514.74.6.1464

Mukuria, G. (2002). Disciplinary challenges: How do principals address this dilemma? *Urban Education*, *37*, 432–452. doi:10.1177/00485902037003007

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes race and the disciplining of young students. *Psychological Science*, *26*, 617–624. doi:10.1177/0956797615570365

Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences*, *113*, 5221–5226. doi:10.1073/pnas.1523698113

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*, 381–398. https://doi.org/10.1177/1745691616635592

Rausch, M. K., & Skiba, R. (2004). Disproportionality in school discipline among minority students in Indiana: Description and analysis. Children left behind policy briefs. Supplementary analysis 2-A. Retrieved from https://eric.ed.gov/?id=ED488897

Rocque, M., & Paternoster, R. (2011). Understanding the antecedents of the "School-to-Jail" link: The relationship between race and school discipline. *Journal of Criminal Law and Criminology*, *101*, 633–666.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36.

Rumberger, R. W., & Losen, D. J. (2017). The hidden costs of California's harsh school discipline: And the localized economic benefits from suspending fewer high school students. Retrieved from https://eric.ed.gov/?id=ED573326

Skiba, R. J., Chung, C. G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal*, *51*, 640–670. doi:10.3102/0002831214541670

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, *34*, 317–342. doi: 10.1023/A:1021320817372

Skiba, R. J., Peterson, R. L., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and Treatment of Children*, *20*, 295–315.

U.S. Department of Education Office for Civil Rights. (2016). *2013-14 Civil rights data collection: A first look*. https://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science*, *326*, 1711–1714. doi:10.1126/science.1178358

Wu, S. C., Pink, W., Crain, R., & Moles, O. (1982). Student suspension: A critical reappraisal. *The Urban Review*, *14*, 245–303. doi: 10.1007/BF02171974

Author Biographies

Shoshana Jarvis is a graduate student in the Department of Psychology at University of California, Berkeley.

Jason Okonofua is an assistant professor in the Department of Psychology at University of California, Berkeley.

Handling Editor: Elizabeth Paluck