



Case Report

The good and bad of a reputation: Race and punishment in K-12 schools

Amanda D. Perez, Jason A. Okonofua^{*}

Department of Psychology, University of California, Berkeley; 2121 Berkeley Way West, Berkeley, CA 94702, United States of America

ARTICLE INFO

Editor: Michael Kraus

Keywords:

Bias
Discipline
Stereotypes
Education

ABSTRACT

Large racial disparities plague discipline in schools across the United States which contributes to racial disparities in life outcomes such as education attainment and incarceration. The present research investigates the role a student's reputation – as shared from one teacher to another one – plays in the discipline context. Teachers ($N = 192$) read about two incidents of misbehavior and reported the severity of discipline the student should receive and the likelihood that they would label the student as a “troublemaker.” They were randomly assigned to read about a Black or White student and to hear from a fellow teacher that the student had a good or bad reputation. Analyses revealed a three-way interaction such that a good reputation buffers against an escalation in discipline severity for a White, but not Black, student. A White student with a bad, as compared to good reputation, received a meaningful escalation in discipline, was more likely to be labeled a troublemaker, and was deemed more likely to get suspended in the future. Meanwhile, reputation was somewhat inconsequential for a Black student. The current research advances theory on the implication of racial bias in context and informs policy for how information is shared among teachers.

1. Introduction

The removal of students from their learning environments as a form of discipline has become a national concern. This issue has recently been recognized in national reports and efforts issued by the U.S. Departments of Education and Justice (Office of Civil Rights, 2016), the Council of State Governments Justice Center (Morgan, Salomon, Plotkin, & Cohen, 2014) and the Discipline Disparities Research to Practice Collaborative (Carter, Fine, & Russell, 2014). Past research has documented how suspensions, the most common form of severe discipline that removes students from school, can decrease the likelihood of high school completion (Balfanz, Byrnes, & Fox, 2015) and increase the likelihood of incarceration (Rocque & Paternoster, 2011; Shollenberger, 2015). Concerns about disciplining students can cause teachers to feel disheartened by the profession (Johnson, Yarrow, Rochkind, & Ott, 2009) and teacher attrition has become its own national concern and contributor to inequity in education outcomes (Garcia & Weiss, 2019; c.f., Nguyen, Pham, Crouch, & Springer, 2020). Ultimately, suspension can cost tax-payers millions of dollars in lost opportunities for these youth to contribute to the economy (Rumberger & Losen, 2017).

While research has begun to uncover a variety of structural (e.g., resources) and policy (e.g., zero-tolerance) reasons for racial disparities in discipline rates (see Okonofua, Perez, & Darling-Hammond, 2020;

Okonofua, Walton, & Eberhardt, 2016), it has focused on teacher-student interactions and relationships as if they existed in a vacuum. For example, in a series of experiments, researchers asked teachers to imagine themselves, first-hand, experiencing two unrelated misbehaviors by a student over the course of a week. Teachers showed an escalation in discipline severity they endorsed from one misbehavior to the next. Further, the extent to which teachers viewed the student as a troublemaker predicted how severely they ultimately wanted to discipline the student and expect the student to get suspended (Okonofua & Eberhardt, 2015). While this research provides valuable insight about the process by which discipline can escalate in the direction of exclusionary discipline, it remains unclear whether this process is impacted by second-hand information about a student's behavior in other situations or contexts, the student's reputation. The present research focuses on teachers' mindsets in the greater context of a school. Namely, how might a student's reputation – the narrative teachers have about particular students—affect the way a teacher perceives the student and responds to misbehavior?

2. What's in a reputation

Reputations matter. Previous research has defined a reputation as the “set of judgements a community makes about the personal qualities of

^{*} Corresponding author.

E-mail addresses: adpc@berkeley.edu (A.D. Perez), okonofua@berkeley.edu (J.A. Okonofua).

<https://doi.org/10.1016/j.jesp.2022.104287>

Received 9 May 2021; Received in revised form 15 December 2021; Accepted 11 January 2022

Available online 20 January 2022

0022-1031/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

one of its members" (Emler, 1990). A good reputation is considered an asset that assures access to valuable resources (Tinsley, O'Connor, & Sullivan, 2002). In the school context, that community can be the faculty and those valuable resources may include social capital like a teacher's good graces. If so, when misbehavior arises, a good reputation may buffer against otherwise severe disciplinary action. Meanwhile, a bad reputation can be deadly (e.g., people deemed witches were sentenced to death in early American history). While the consequences tend to not be as dire in modern times, the consequences remain negative and there remains a low bar for a bad reputation to be formed (Skowronski & Carlston, 1987; Ellemers, Pagliaro, & Barreto, 2013). Yet, to date, experimental research has not investigated how reputations function in the context of school discipline.

Experimental research has primarily focused on teachers' own, firsthand, experience with a student's misbehavior in their classroom (Gilliam et al., 2016; Okonofua & Eberhardt, 2015). Classrooms, however, do not exist in isolation. It is not uncommon for teachers to talk to each other and communicate experiences they have had with students. A meta-analysis on teacher collaboration shows that communication among teachers can contribute to positive outcomes for students, teachers, and the school at large (Vangrieken, Dochy, Raes, & Kyndt, 2015). For students, teacher collaboration (e.g., sharing information about experiences with students) has been found to improve understanding and achievement (see Egodawatte, McDougall, & Stoilescu, 2011; Reeves, Pun, & Chung, 2017). However, these benefits depend on the nature of the teacher collaboration or the valence of information shared (Vangrieken et al., 2015). Discussions among teachers could include the sharing of positive reputations which might lead a teacher to further challenge a student to reach their potential or give that student the benefit of the doubt (e.g., Curry, 2008; Lomos, Hofman, & Bosker, 2011). However, it can also involve sharing negative reputations which might increase the likelihood that teacher will view the student as a troublemaker and increase how severely teachers want a student to be disciplined (Okonofua et al., 2016). Research has yet to experimentally investigate how knowledge of a good or bad reputation can shape teachers' responses to misbehavior. In the present research, we investigate the role of reputation in responses to misbehavior.

3. Does race matter?

Black students face a heightened risk of school discipline and the associated negative life outcomes. For example, while Black boys represent 7.9% of public-school students, they represent over 25% of students suspended from school (Office of Civil Rights, 2016). In fact, Black boys are more likely to be suspended than White boys, Black girls, and White girls combined (18%, 5%, 10%, 2%, respectively). Despite efforts to mitigate the disparity in discipline (e.g., Goyer et al., 2019; Gregory, Allen, Mikami, Hafen, & Pianta, 2015; Gregory, Huang, Anyon, Greer, & Downing, 2018; Okonofua et al., 2020), the disparity has remained pervasive for many years (see Losen, Hodson, Keith II, Morrison, & Belway, 2015; Skiba, Horner, Chung, Rausch, May, & Tobin, 2011). What processes might contribute to the persistence of this racial inequality in discipline rates?

Experimental research has begun to uncover psychological processes that can contribute to Black students being at heightened risk of severe discipline. Namely, a student's race can shape a teacher's accounts of a series of misbehaviors and their disciplinary responses to them. Okonofua and Eberhardt (2015) asked teachers to read about a series of encounters with a Black or White student misbehaving in their class and found the student's race caused a different trajectory of disciplinary responses. Teachers showed a sharper escalation in how severely they wanted to discipline the student from the first misbehavior to the next one, if the student was Black as compared to White – the "Black escalation effect". Without information about how the student behaves in other classrooms, teachers were more likely to assume the Black student was a troublemaker and to expect the Black student to get suspended in

the future. What if teachers had heard the student had a good or bad reputation based on a fellow teacher's account of the student's behavior in their classroom?

The fixedness of the "troublemaker" label implies that the student is consistently prone to a pattern of misbehavior that extends from one isolated incident to the next (Okonofua et al., 2016; Okonofua & Eberhardt, 2015). When teachers were exposed to the same misbehavior, they were more likely to view a Black, as opposed to White, student through this fixedness lens and in turn endorsed more severe discipline. However, it is not clear that fixedness is a preconceived notion or one that develops due to experience with the Black student's misbehavior. A misbehaving Black student may face a sharper escalation in discipline, because by default, it is assumed that his misbehavior is indicative of a fixed trait, one that is not context specific (e.g., present in any classroom). This could imply that seeking a non-punitive remedy would be fruitless (Okonofua, Paunesku, & Walton, 2016). Meanwhile, by default, a White student's misbehavior may be viewed as situational (e.g., just having a bad day), not a consistent problem, and thus severe discipline may seem less warranted. If so, evidence of the fixedness (e.g., a bad reputation) before a teacher's own experience with a student's misbehavior would matter for a White student but less so for a Black student. On one hand, a teacher may show a sharper escalation in discipline severity for a White student when, contrary to the default attribution, a bad reputation means his misbehavior is less likely to be situational. Likewise, a good reputation could buffer against an escalation in discipline for a White student by scaffolding the default assumption that his misbehavior is situational. On the other hand, for a Black student, a bad reputation may merely be interpreted as stereotype consistent and thus teachers respond with the same escalation in discipline severity. And what about a good reputation for a Black student? The fixedness of anti-Black stereotypes may cause a good reputation to have little effect on the escalation of discipline for a Black person. For example, Bertrand and Mullainathan (2004) sent fictitious resumes to help wanted ads in two major cities. The resumes were randomly assigned to be low or high quality and randomly assigned to be perceived as sent by a Black or White person by way of stereotypical Black and White names. While a higher quality resume significantly increased the likelihood that a White person would receive a call-back, the quality of the resume mattered less for a Black applicant. A good record, did not benefit a Black applicant. Thus, the race disparity grows with increased reputational information. This effect has been found to persist despite racial bias becoming more covert or implicit in nature (for review see: Quillian, Pager, Hexel, & Midtbøen, 2017). Might a similar pattern of effects exist in the discipline context, such that a good reputation does not buffer against escalated discipline for a Black child?

The current research is the first to test these possibilities in a 2 (student race: Black versus White) X 2 (reputation: good versus bad) X 2 (first infraction versus second infraction) factorial design. The first two conditions are between-subjects, while the last is within-subjects. We predicted a three-way interaction such that a good, as compared to bad, reputation buffers against an escalation in discipline from one misbehavior to the next for a White, but not Black, student. We also explored effects of reputation on troublemaker label and suspension prediction. We showed practicing teachers information from a fellow teacher about a student's reputation (good or bad) and then two incidents of misbehavior by that student (Black or White). Similar to previous research, our primary outcomes were teachers' reports of how severely the student should be disciplined after each incident (discipline severity), how likely they were to call the student a troublemaker (troublemaker-labeling), and to what extent they predicted the student will get suspended in the future (suspension-prediction). All outcomes were rated on a 1–5 scale with higher values reflecting higher levels of the specific question. Other secondary and exploratory (detention-endorsement and feeling troubled) outcomes are reported in the Supplemental Materials (Table S1–S2).

The focus of this research is to determine if there is a main effect of

reputation if the student is White but not Black. Thus, the experiment is designed with good versus bad reputation stimuli to provide results for that research question. This design does not include a condition where no reputation is communicated and thus this experiment cannot replicate previous research that shows teachers want more severe discipline for a Black student as compared to a White student (Okonofua & Eberhardt, 2015). Rather, we seek to advance theory from that work to determine whether a good reputation can buffer against troublemaker labeling and discipline severity and to explore how the impact of a reputation may differ according to a student's race.

4. Method

4.1. Participants

We recruited $N = 192$ practicing teachers via two sampling methods: utilizing an online pool of teachers ($n = 64$) collected through collaboration with an existing research teacher network as well as a sample of teachers ($n = 128$) in a West Coast high school. The former sample was paid \$10 gift cards, and the latter sample did not receive compensation. Across sampling methods, all participation was optional. There were no significant differences between the samples for the analyses and therefore they were analyzed together as one sample (See Table S2). Our sample consisted of 54% Female, 38% Male, and 8% unknown with 50% White, 16% Asian, 9% Latine 5% Black, 11% Other, and 9% unknown participants ($M_{age} = 41.79$, $SD_{age} = 12.6$; $M_{year\text{teaching}} = 11.84$, $SD_{year\text{teaching}} = 8.55$).

Power analyses using Gpower (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that this sample size has a power above 0.80 to detect the effect size found in previous related work (Okonofua & Eberhardt, 2015; Cohen's d for independent group differences = 0.89 & Cohen's d for within group differences = 2.10) on discipline outcomes. All measures, manipulations, and exclusions in the study are disclosed. No data was collected after analyses were run.

4.2. Procedure and measures

After completing consent forms, teachers were shown a picture of a school and were told to imagine themselves as a teacher at the school as well as a typical day as a teacher in the pictured school. They were told that the school had students from a middle-income neighborhood and that the student to teacher ratio was 22:1. Finally, before moving on to the main portion of the study, they were told that they were about to read actual teaching reports about typical incidents involving a misbehaving student. They were told that their job was to read each story carefully and answer the questions as though they were the actual teacher in the class.

Subsequently, they were shown one of the following four scenarios to read and answer questions for: A Black student (Darnell) with a bad reputation, a Black student (Darnell) with a good reputation, a White student (Greg) with a bad reputation, or a White student (Greg) with a good reputation. These names were drawn from previous research on common or stereotypically Black and White names (Greenwald, McGhee, & Schwartz, 1998; Levitt & Dubner, 2005) and research with this discipline procedure that showed these names subtly cue race (Okonofua & Eberhardt, 2015). However, unlike past research, participants first engage with different framings of the misbehavior narrative according to their randomly assigned condition. For the bad-reputation condition they read the following vignette to manipulate reputation:

It is your first day of school and you are sitting in your classroom before the school day begins. You are reading through your class roster and notice you have a boy named [Darnell/Greg] in your class. You remember hearing about [Darnell/Greg] from his math teacher last year. She said that he would come to class unprepared and did not always get along with other students. She told you a story about a time when [Darnell/Greg]

repeatedly did not follow her instructions. In all, she said he had a negative demeanor and received a poor grade in the class.

For the good-reputation condition they read the following vignette to manipulate reputation:

It is your first day of school and you are sitting in your classroom before the school day begins. You are reading through your class roster and notice you have a boy named [Darnell/Greg] in your class. You remember hearing about [Darnell/Greg] from his math teacher last year. She said that he was helpful in class and seemed to really enjoy learning. She told you a story about a time when [Darnell/Greg] helped another student to understand a lesson. In all, she said he had a positive demeanor and received an excellent grade in the class.

Next, teachers read about an infraction done by the student:

A few weeks have passed and it is midterm day. [Darnell/Greg] comes to class late. You ask him for his tardy pass. He doesn't respond. You ask him again, and he slams it on your desk. Then, while the class is taking the test, [Darnell/Greg] makes a lot of noise stomping to his desk.

After reading this first infraction they answered the following questions: *How severe was [Darnell's/Greg's] behavior? To what extent is [Darnell/Greg] hindering you from maintaining order in the class? How irritating is [Darnell/Greg]? and How severely should [Darnell/Greg] be disciplined?* The first three questions were combined as a composite called "Feeling Troubled" as was done in previous research (Okonofua et al., 2020; Okonofua & Eberhardt, 2015). This composite scale had a Cronbach alpha of 0.74 and 0.80 for infraction 1 and 2 respectively.

Once they answered these questions they saw the following vignette:

3 days later, [Darnell/Greg] misbehaves again... Today, [Darnell/Greg] is upset because you "bother" him when he "wants to sit quietly and do nothing". And he says that you should just leave him alone. So you give him reading assignments and just busy work. But [Darnell/Greg] doesn't do anything you give him. And he calls you "crazy".

After reading this second and final infraction they answered the same questions as before in addition to the following questions: *How likely is it that you would say that [Darnell/Greg] is a troublemaker? To what extent would you be willing to recommend that your principal give [Darnell/Greg] detention for 2 days? How easy is it to see [Darnell/Greg] getting suspended further down the road?*

Finally, participants completed manipulation checks, demographic information, suspicion probes, and survey questions unrelated to the present research (i.e., big five inventory, social dominance orientation, and system justification) prior to being debriefed and thanked for their participation. There were no failures for the manipulation check or suspicion probes so there were no data exclusions based on these criteria.

5. Results

5.1. Analysis plan

All analyses were run in the statistical programming software R and the accompanying R Studio Graphical User Interface. Discipling-severity was measured at two timepoints and was analyzed using the linear mixed effects regression, *aov()*, function in the *stats* package (R Core Team, 2021). All other dependent variables were measured only after the second infraction and therefore the linear model *lm()* function was used. Due to our interest in comparing Greg with a bad reputation to Greg with a good reputation, as well as Darnell with a bad reputation to Darnell with a good reputation, we tested and report these contrasts. We also ran contrast analyses on all combinations of cells and applied Tukey's correction to adjust for the number of comparisons being run.

All unreported contrasts are reported in the Supplemental Materials. All data needed to evaluate the conclusions in the paper can be found on OSF, https://osf.io/4pf5b/?view_only=b581d15444b447038c42181f5a7e0180.

5.2. Discipline severity

Do a student's reputation and race and time interact? Yes, anova analyses revealed a three-way interaction, $F(1,178) = 6.02, p = 0.015$. There were no other main effects or interactions (see Table 1).

As seen in previous literature, we saw the Black escalation effect (Okonofua & Eberhardt, 2015) such that there was a significant increase in discipline severity for the Black student regardless of reputation type (See Fig. 1). For the Black student with a bad reputation, discipline severity rose from infraction 1 ($M = 2.44, SD = 0.76$) to infraction 2 ($M = 2.55, SD = 0.96$), $b = 0.40, SE = 0.11, z = 3.54, p = 0.009, d = 0.45$. For the Black student with a good reputation, discipline severity rose from infraction 1 ($M = 2.35, SD = 0.65$) to infraction 2 ($M = 2.50, SD = 0.69$), $b = 0.50, SE = 0.12, z = 4.33, p < 0.001, d = 0.67$.

As predicted, there was no significant difference between infraction 1 ($M = 1.81, SD = 0.85$) and infraction 2 ($M = 2.00, SD = 0.86$) discipline severity for the White student with a good reputation, $b = 0.19, SE = 0.11, z = 1.67, p = 0.68, d = 0.22$ (See Fig. 2). Yet, there was a significant difference for the White student with a bad reputation such that discipline severity rose from infraction 1 ($M = 1.69, SD = 0.72$) to infraction 2 ($M = 2.36, SD = 0.96$), $b = 0.67, SE = 0.12, z = 5.51, p < 0.001, d = 0.79$. There were no other meaningful significant contrasts, but see Table S3 in the Supplemental Materials for all remaining contrasts.

5.3. Troublemaker

There was a main effect of reputation on teachers' likelihood to label the student as a troublemaker, $b = 0.23, SE = 0.07, t = 3.03, p = 0.003$. There was no main effect of student race, $b = -0.04, SE = 0.07, t = -0.60, p = 0.55$, or interaction between race and reputation conditions, $b = -0.09, SE = 0.07, t = -1.28, p = 0.20$. Contrast analyses reveal, similar to discipline severity, we found no difference in troublemaker-labeling for the Black student with a good ($M = 1.89, SD = 0.88$) or bad reputation ($M = 2.15, SD = 0.96$; $b = -0.26, SE = 0.21, t = -1.25, p = 0.60, d = 0.28$). The White student, on the other hand, was heavily penalized and was more likely to be labeled a troublemaker if he had a bad reputation ($M = 2.43, SD = 1.21$) as compared to a new one ($M = 1.79, SD = 0.93, b = -0.64, SE = 0.21, t = -3.02, p = 0.02, d = 0.60$) (See Fig. 3). There were no other significant contrasts, but see Table S4 in the Supplemental Materials for all remaining contrasts.

5.4. Future suspend

There was a main effect of reputation on teachers' likelihood to label the student as a troublemaker, $b = 0.26, SE = 0.08, t = 3.33, p = 0.001$. There was a trending effect of student race, $b = -0.14, SE = 0.08, t = -1.82, p = 0.07$ and no significant interaction between race and reputation conditions, $b = -0.11, SE = 0.08, t = -1.42, p = 0.158$. Contrast

Table 1

ANOVA results using discipline severity as the outcome variable.

Variable	SS	MS	F	p
Reputation	1.16	1.16	1.06	0.31
Race	0.08	0.08	0.07	0.79
Infraction Time	17.15	17.15	55.78	< 0.001**
Reputation*Race	0.00	0.003	0.002	0.96
Reputation*Infraction Time	0.77	0.77	2.51	0.12
Race*Infraction Time	0.02	0.02	0.07	0.80
Race*Infraction Time*Reputation	1.85	1.85	6.02	0.02*

$p < 0.05^* p < 0.001^{**}$.

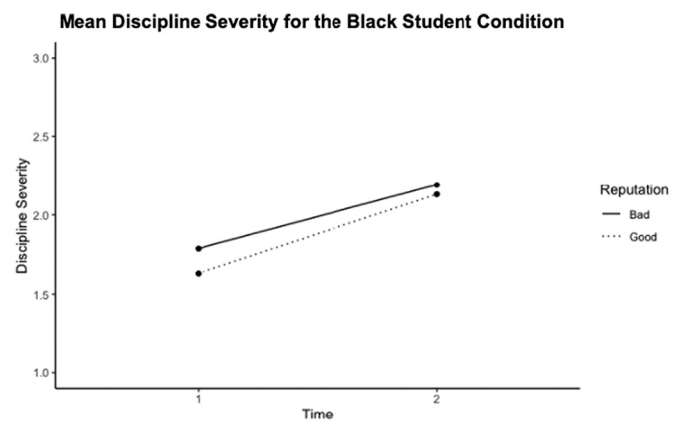


Fig. 1. Line chart showing the condition level differences for the mean discipline severity based on student reputation for the Black student condition.

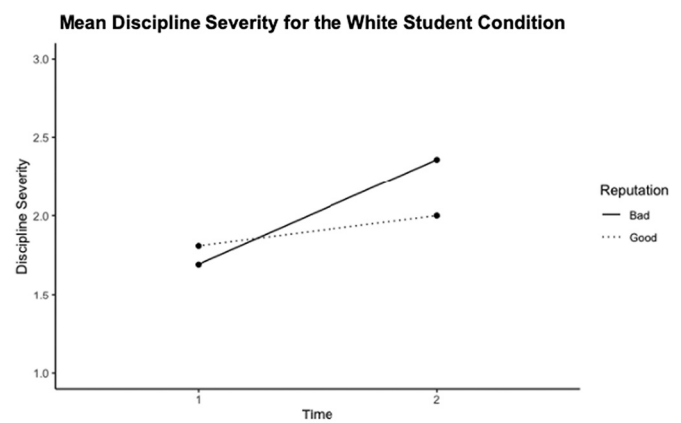


Fig. 2. Line chart showing the condition level differences for the mean discipline severity based on student reputation for the White student condition.

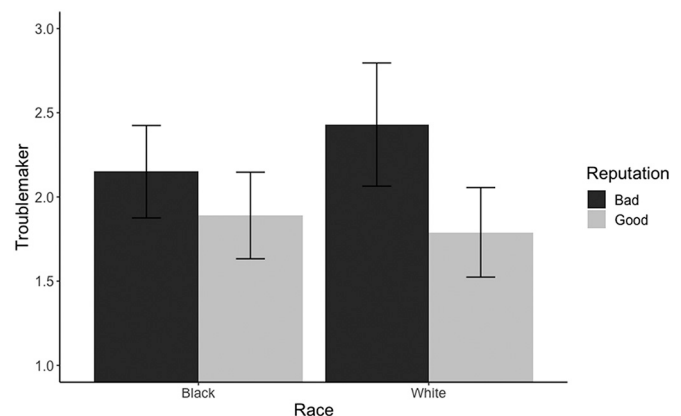


Fig. 3. Bar chart showing mean differences on troublemaker ratings based on student race and student reputation. Error bars represent 95% Confidence Intervals.

analyses reveal, similar to discipline severity and troublemaker labeling, we found no difference in likelihood to suspend the Black student in the future based on good ($M = 2.20, SD = 1.04$) or bad reputation ($M = 2.50, SD = 1.07$; $b = -0.30, SE = 0.22, t = -1.36, p = 0.53, d = 0.28$). The White student, on the other hand, had a higher likelihood of future suspension in the bad reputation condition ($M = 3.00, SD = 1.10$) in comparison to the good reputation condition ($M = 2.27, SD = 1.08, b =$

-0.73 , $SE = 0.22$, $t = -3.34$, $p = 0.005$, $d = 0.68$) (See Fig. 4). There were no other meaningful significant contrasts, but see Table S5 in the Supplemental Materials for all remaining contrasts.

5.5. Fixedness process (post-hoc exploratory analysis)

Due to the significant finding on suspension-prediction primary outcome, we explored if the same troublemaker-labeling mediation process evidenced in previous research (Okonofua & Eberhardt, 2015) applied to the difference in reputation for a White student in the present research. This analysis would help understand the consistency of the psychological process that determines discipline. In the previous research, a Black student's Blackness (as marked by a continuous variable of how likely the student was to be Black) predicted troublemaker-labeling which in turn predicted suspension-prediction. Here, we tested if for a White student, reputation replaces Blackness such that reputation predicts troublemaker-labeling which in turn predicts suspension-prediction. The causal path between troublemaker-labeling and suspension-prediction is ambiguous from a methodological standpoint since the outcomes were measures during the same psychological moment following the experimental manipulations (e.g., alternatively suspension-prediction could mediate reputation effect on troublemaker). Thus, these mediation findings should be interpreted with caution. Mediation analyses were conducted with data solely from the White student condition entered into the R-coding mediation macro called *MedTextR*. The predicted mediation (indirect path through labeling a student as a troublemaker) was significant and partially mediated, $b = 0.26$, $SE = 0.26$, 95% confidence interval (CI) = [0.08, 0.57] (see Fig. 5). We focus on suspension-prediction, instead of discipline-severity, because unlike discipline-severity, the suspension-prediction outcome came after the troublemaker-labeling outcome in the procedure. However, mediation for discipline-severity is also reported in Supplemental Materials (Fig. S1).

6. Discussion

Previous research has shown that teachers show a sharper escalation in discipline severity and higher likelihood to see the student being suspended in the future, when a Black student, as compared to White, is involved in the incidents of misbehavior (Okonofua & Eberhardt, 2015). The present work expands on these findings to test the impact of learning more information about the student from a fellow teacher at the school, a common occurrence in K-12 education (Vangrieken et al., 2015). It also advances theory on the psychological mechanism involved in the discipline process. Namely, the process is not necessarily about the race of the student, but rather about what a teacher may infer from race:

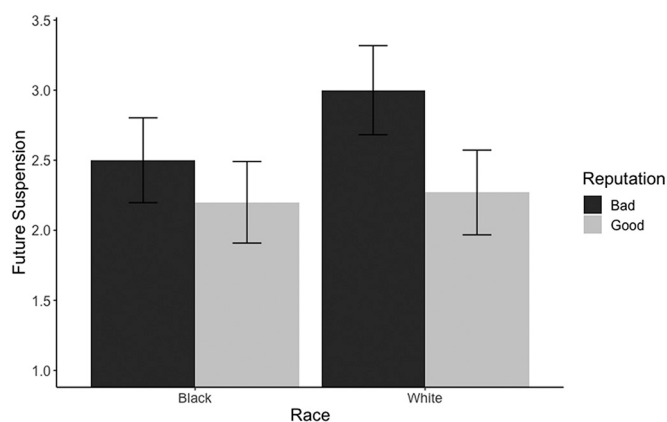


Fig. 4. Bar chart showing mean differences in likelihood of future suspension based on student race and student reputation. Error bars represent 95% Confidence Intervals.

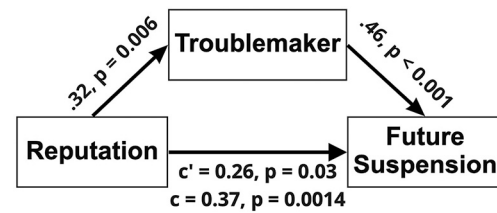


Fig. 5. Mediation diagram showing the partial mediation of troublemaker on the reputation to suspension relationship. "Troublemaker" and "Future Suspension" outcomes were measured during the same psychological moment after the reputation manipulation and thus that causal path is ambiguous and should be interpreted with caution. "Good Reputation" is the reference level for the Reputation variable.

reputation.

There are clear practical implications for these findings. Teachers regularly communicate with one another about their students. The information can exacerbate racial inequity in discipline decisions because the student's reputation elicits a different interpretation based on the student's race. These findings highlight how Black students are negatively impacted by stereotypes such that a good reputation is only a buffer for the White, but not Black student. If the student is Black, a good or bad reputation does not affect a teacher's responses to his misbehavior. Meanwhile, a fellow teacher's report of a White student's good or bad reputation predicts discipline responses for the student's subsequent misbehavior. Past research suggests that a White student's misbehaviors are more likely to be seen as isolated events such that there is no escalation in severity of discipline teachers endorse or heighten prediction of future suspensions (for example, see Okonofua & Eberhardt, 2015). The present research uncovers a more nuanced process such that this may only be the case if the White student has no reputation – like in previous research – or has a good reputation. Meanwhile, if that White student has a bad reputation, teachers respond to subsequent misbehavior in a similar manner as if the student was Black, escalation in discipline severity and relatively high prediction of future suspension. If only White students benefit from having a good reputation, racial disparities in discipline widen. Also, it is possible that White, as compared to Black, students are less likely to have a bad reputation or have it communicated by fellow teachers given they are underrepresented in the percentage of students suspended from school (Office of Civil Rights, 2016). Future archival or observation research should explore this possibility. These findings suggest that education policy and teacher professional development could better mitigate racial disparities in discipline with targeted focus on how teachers share information and how that information should be interpreted given the realities of racial bias (see Okonofua et al., 2020).

This research also has theoretical significance. Previous experimental research on racial disparities in discipline has focused on teachers reaching their own conclusions based on a firsthand account of a student's behavior in the classroom (Okonofua & Eberhardt, 2015). However, the research did not investigate the contribution of secondhand accounts of student behavior. Other research has focused on the role of principals and the secondhand accounts of misbehavior they receive in office referral forms (Jarvis & Okonofua, 2020; Rausch & Skiba, 2004; Skiba, Michael, Nardo, & Peterson, 2002). Across both lines of research, the student's race has been found to predict interpretation of the misbehavior. The present research draws on both lines of research to investigate and evidence that secondhand accounts of a student's reputation affects how a teacher interprets and responds to subsequent misbehavior by the student. Research on stereotyping should experimentally test the various sources of misbehavior information that can be affected by the stereotypes.

This research also advances theory on the nature of stereotypes. Previous research shows that a student's blackness increases the

likelihood he will be viewed as a troublemaker which in turn increases the likelihood the educator (teacher or principal) will expect the student to be suspended in the future (Jarvis & Okonofua, 2020; Okonofua & Eberhardt, 2015). The present research shows how the troublemaker-labeling process can apply to a White student as well. In place of a negative racial stereotype, a bad reputation can increase the likelihood that a White student will be viewed as a troublemaker and in turn a teacher's expectation that the student will get suspended.

There were a few limitations to the present research that can benefit from future research. Similar to previous research, the present research did not show an effect of student race on how troubled the teacher felt (e.g., feeling hindered from doing their job) by student misbehaviors (Okonofua et al., 2020). Future research should investigate why teachers and principals show racial bias in their discipline decisions with or without also showing a difference in how troubled they feel by the student's misbehavior. The present research tested a new discipline outcome, detention, but did not find any significant effect of student race or reputation on it. However, for the most part, archival research has shown that racial disparities in discipline exist on more punitive or formal disciplinary actions such as corporal punishment, expulsions, and referrals to law enforcement (see Office of Civil Rights, 2016). Future experimental research should explore how the psychological processes reported in this paper might contribute to these other discipline outcomes.

The present research shows that a student's reputation – as communicated from a fellow teacher – can affect how a teacher responds to a White, but not Black, student's misbehavior. This difference occurred despite the student's misbehavior being held constant. Taken together with other recent research, stereotyping can affect interpretation of both firsthand and secondhand accounts of misbehavior, all of which can ultimately contribute to the large racial disparities in school discipline examined across the United States.

Author notes

Both authors designed the research; Okonofua recruited participants; Perez performed research, collected data, and analyzed data; Both authors wrote the paper. The authors declare no conflicts of interest with respect to authorship or publication of this article. University of California at Berkeley's Institutional Review Board approved all research reported in this article. Perez is supported by a National Science Foundation predoctoral fellowship.

Acknowledgements

We thank Equity, Diversity, and Empathy Navigation Sciences (EDENS) Lab, R. Mendoza-Denton, and E. Adaora Okonofua for their feedback and support. We also thank the teachers and school districts for their participation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jesp.2022.104287>.

References

- Balfanz, R., Byrnes, V., & Fox, J. H. (2015). Sent home and put off track. *Closing the school discipline gap: Equitable remedies for excessive exclusion*, 17–30.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94(4), 991–1013.
- Carter, P., Fine, M., & Russell, S. (2014). *Discipline disparities: A research-to-practice collaborative*. The Equity Project at Indiana University.
- Core Team, R. (2021). R: A language and environment for statistical computing. In *R foundation for statistical computing*. Vienna, Austria <https://www.R-project.org/>.
- Curry, M. (2008). Critical friends groups: The possibilities and limitations embedded in teacher professional communities aimed at instructional improvement and school reform. *Teachers College Record*, 110(4), 733–774.
- Egodawatte, G., McDougall, D., & Stoilescu, D. (2011). The effects of teacher collaboration in grade 9 applied mathematics. *Educational Research for Policy and Practice*, 10, 189–209. <https://doi.org/10.1007/s10671-011-9104-y>
- Ellemers, N., Pagliaro, S., & Barreto, M. (2013). Morality and behavioural regulation in groups: A social identity approach. *European Review of Social Psychology*, 24(1), 160–193.
- Emler, N. (1990). A social psychology of reputation. *European Review of Social Psychology*, 1(1), 171–193.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Garcia, E., & Weiss, E. (2019). The teacher shortage is real, large and growing, and worse than we thought. In *The first report in "the perfect storm in the teacher labor market" series*. Economic Policy: Institute.
- Goyer, J. P., Cohen, G. L., Cook, J. E., Master, A., Apfel, N., Lee, W., ... Walton, G. M. (2019). Targeted identity-safety interventions cause lasting reductions in discipline citations among negatively stereotyped boys. *Journal of Personality and Social Psychology*, 117(2), 229.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2015). The promise of a teacher professional development program in reducing racial disparity in classroom exclusionary discipline. *Closing the school discipline gap: Equitable remedies for excessive exclusion*, 168.
- Gregory, A., Huang, F. L., Anyon, Y., Greer, E., & Downing, B. (2018). An examination of restorative interventions and racial equity in out-of-school suspensions. *School Psychology Review*, 47(2), 167–182.
- Jarvis, S. N., & Okonofua, J. A. (2020). School deferred: When Bias affects school leaders. *Social Psychological and Personality Science*, 11(4), 492–498. <https://doi.org/10.1177/1948550619875150>
- Johnson, J., Yarrow, A., Rochkind, J., & Ott, A. (2009). *Teaching for a living: How teachers see the profession today*. New York: Public Agenda/Education Week.
- Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York, NY: Harper Perennial.
- Lomos, C., Hofman, R. H., & Bosker, R. J. (2011). The relationship between departments as professional learning communities and student achievement in secondary schools. *Teaching and Teacher Education*, 27, 722–731. <https://doi.org/10.1016/j.tate.2010.12.003>
- Losen, D. J., Hodson, C. L., Keith, M. A., II, Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?*
- 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*. New York, NY: Council of State Governments Justice Center.
- Nguyen, T. D., Pham, L. D., Crouch, M., & Springer, M. G. (2020). The correlates of teacher turnover: An updated and expanded meta-analysis of the literature. *Educational Research Review*, 31, 100355.
- Office of 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>.
- Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes race and the disciplining of young students. *Psychological Science*, 26(5), 617–624. <https://doi.org/10.1177/0956797615570365>
- Okonofua, J. A., Perez, A. D., & Darling-Hammond, S. (2020). When policy and psychology meet: Mitigating the consequences of bias in schools. *Science Advances*, 6(42), eaba9479.
- 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.
- Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, 114(41), 10870–10875.
- 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. Center for Evaluation and Education Policy. Indiana University.
- Reeves, P. M., Pun, W. H., & Chung, K. S. (2017). Influence of teacher collaboration on job satisfaction and student achievement. *Teaching and Teacher Education*, 67, 227–236.
- Roque, M., & Paternoster, R. (2011). Understanding the antecedents of the "school-to-jail" link: The relationship between race and school discipline. *The Journal of Criminal Law and Criminology*, 101(2), 633–666.
- 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* (Civil Rights Project-Proyecto Derechos Civiles).
- Shollenberger, T. L. (2015). Racial disparities in school suspension and subsequent outcomes. In *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 31–44).
- Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A National Investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, 40(1), 85–107.
- Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, 34(4), 317–342. <https://doi.org/10.1023/A:1021320817372>

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of personality and social psychology*, 52(4), 689.

Tinsley, C. H., O'Connor, K. M., & Sullivan, B. A. (2002). Tough guys finish last: The perils of a distributive reputation. *Organizational Behavior and Human Decision Processes*, 88(2), 621–642.

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17–40.