



Racial/Ethnic Discrimination and Heterogeneity Across Schools in the U.S. Public Education System: A Correspondence Audit of Principals

S. Michael Gaddis
NWEA

Charles Crabtree
Dartmouth College

John B. Holbein
University of Virginia

Steven Pfaff
University of Washington

Although numerous studies document different forms of discrimination in the U.S. public education system, very few provide plausibly causal estimates. Thus, it is unclear to what extent public school principals discriminate against racial and ethnic minorities. Moreover, no studies test for heterogeneity in racial/ethnic discrimination by individual-level resource needs and school-level resource strain – potentially important moderators in the education context. Using a correspondence audit, we examine bias against Black, Hispanic, and Chinese American families in interactions with 52,792 public K-12 principals in 33 states. Our research provides causal evidence that Hispanic and Chinese American families face significant discrimination in initial interactions with principals, regardless of individual-level resource needs. Black families, however, only face discrimination when they have high resource needs. Additionally, principals in schools with greater resource strain discriminate more against Chinese American families. This research uncovers complexities of racial/ethnic discrimination in the K-12 context because we examine multiple racial/ethnic groups and test for heterogeneity across individual- and school-level variables. These findings highlight the need for researchers conducting future correspondence audits to expand the scope of their research to provide a more comprehensive analysis of racial/ethnic discrimination in the U.S.

VERSION: February 2024

Suggested citation: Gaddis, S. Michael, Charles Crabtree, John B. Holbein, and Steven Pfaff. (2024). Racial/Ethnic Discrimination and Heterogeneity Across Schools in the U.S. Public Education System: A Correspondence Audit of Principals. (EdWorkingPaper: 24-911). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/xxm8-ral3>

**Racial/Ethnic Discrimination and Heterogeneity Across Schools
in the U.S. Public Education System: A Correspondence Audit of Principals**

S. Michael Gaddis
Senior Research Scientist
NWEA
michael.gaddis@nwea.org

Charles Crabtree
Assistant Professor
Department of Government
Dartmouth College

John B. Holbein
Associate Professor
Frank Batten School of Leadership and Public Policy
University of Virginia

Steven Pfaff
Professor
Department of Sociology
University of Washington

February 2024 Draft

Keywords: discrimination, public schools, principals, race/ethnicity, correspondence audit, field experiment

Acknowledgments: We thank David P. Baker, Quintin Beazer, William Berry, Nicholas Carnes, Katie E. Corcoran, Penny Edgell, Matt Golder, Sona Golder, Paul Goldman, D. Sunshine Hillygus, Yusaku Horiuchi, Holger L. Kern, Jae Yeon Kim, Helen Ladd, Edvard Nergård Larsen, Cecilia Mo, Donald Moynihan, Walter Mebane, Ross Matsueda, Ying Shi, Gi-Wook Shin, Ian Solomon, Lucy Sorensen, Kiyoteru Tsutsui, Danielle Vance-McMullen, Jacob Vigdor, Carly Wayne, and Ariel White for their assistance and feedback during various stages of this project. We also thank audiences at Dartmouth College, Pennsylvania State University, the University of Michigan, the University of Washington, West Virginia University, Brigham Young University, the University of Virginia, Duke University, the NYU-CESS 10th Annual Experimental Political Science Conference, and the Harvard Experimental Political Science Graduate Student Conference for insightful comments. We also thank the Human Subjects Committees at Pennsylvania State University (#STUDY00002386), the University of Michigan (#HUM00118066), and Duke University (#B0361) for reviewing and approving the experiments in our study.

Racial/Ethnic Discrimination and Heterogeneity Across Schools in the U.S. Public Education System: A Correspondence Audit of Principals

ABSTRACT

Although numerous studies document different forms of discrimination in the U.S. public education system, very few provide plausibly causal estimates. Thus, it is unclear to what extent public school principals discriminate against racial and ethnic minorities. Moreover, no studies test for heterogeneity in racial/ethnic discrimination by individual-level resource needs and school-level resource strain – potentially important moderators in the education context. Using a correspondence audit, we examine bias against Black, Hispanic, and Chinese American families in interactions with 52,792 public K-12 principals in 33 states. Our research provides causal evidence that Hispanic and Chinese American families face significant discrimination in initial interactions with principals, regardless of individual-level resource needs. Black families, however, only face discrimination when they have high resource needs. Additionally, principals in schools with greater resource strain discriminate more against Chinese American families. This research uncovers complexities of racial/ethnic discrimination in the K-12 context because we examine multiple racial/ethnic groups and test for heterogeneity across individual- and school-level variables. These findings highlight the need for researchers conducting future correspondence audits to expand the scope of their research to provide a more comprehensive analysis of racial/ethnic discrimination in the U.S.

INTRODUCTION

How prevalent and pervasive is racial/ethnic discrimination in the public education system today? The broader context of such discrimination has shifted in the last 70 years. While 20th-century racial/ethnic discrimination in public education can be traced back to the state-sanctioned Jim Crow-era “separate but equal” segregation doctrine, *Brown vs. Board* altered the course of public education (Kluger, 2011). Although some modern laws and policies have aimed to combat discrimination, several legal scholars argue that the overall response often has been tepid and is often unclear (Crenshaw, 1988; Freeman, 1977; 1989), and social scientists affirm that discrimination in various domains is far from a problem of the past (e.g., Gaddis et al., 2023; Quillian et al., 2017). In public education, evidence suggests numerous racial/ethnic disparities persist (Farkas, 2003; Mickelson, 2003; O’Connor, Lewis, and Mueller, 2007), leading scholars to investigate discrimination as a potential explanation.

Unfortunately, much of the research on discrimination in education has distinct limitations. The evidence on racial/ethnic discrimination in public K-12 education mostly comes from surveys or interviews examining the effects of students’ perceptions of experiencing discrimination (Benner et al., 2018; Dotterer, McHale, and Crouter, 2009; Priest et al., 2013; Smalls et al., 2007). While these studies are essential, they fail to capture the scope of racial/ethnic discrimination across a wide variety of schools and racial/ethnic groups in a causal framework. An exception to this pattern comes from two recent studies using correspondence audits to provide causal evidence on racial/ethnic discrimination in K-12 schools (Oberfield and Incantalupo, 2021; Rivera and Tilcsik, 2023). However, these two studies have some crucial limitations. First, they provide an aggregate view of racial/ethnic discrimination against only one racial/ethnic minority group (i.e., Black families). Second, they rely on either small or

geographically limited samples. Third, these studies fail to examine potential heterogeneity across schools, which is essential for researchers and policymakers to understand where the problem is the largest and, thus, where to target potential interventions.

In this paper, we use a large-scale correspondence audit to examine discrimination against Black, Hispanic, and Chinese Americans in interactions with 52,792 public K-12 principals in 33 states. We test for heterogeneity by individual-level resource needs and school-level resource strain. Our research provides the first causal evidence that Hispanic and Chinese American families face significant discrimination in initial interactions with public school principals. The levels of discrimination that Hispanic and Chinese American families experience do not significantly vary by signals of individual-level resource needs. Black families, however, only face discrimination when they signal high resource needs. Additionally, principals in schools with greater resource strain discriminate more against Chinese American families than principals in schools with less resource strain. The levels of discrimination that Black and Hispanic families experience do not significantly vary by school-level resource strain. Our larger, more comprehensive study provides new additional insight into the “what,” “where,” and “when” of racial/ethnic discrimination in the K-12 education context while simultaneously setting up new research to investigate the potential mechanisms of and solutions to this discrimination (Gaddis, 2019).

BACKGROUND

Studying Racial/Ethnic Discrimination in Schools

We first note that when we use the term “racial/ethnic discrimination,” we mean differential treatment of individuals who are all-else-equal except for their race/ethnicity

(National Research Council, 2004). As prominent social scientists who led the charge in using audit studies to examine discrimination, Pager and Shepherd (2008: 182) note:

Racial discrimination refers to unequal treatment of persons or groups on the basis of their race or ethnicity ... A key feature of any definition of discrimination is its focus on behavior. Discrimination is distinct from racial prejudice (attitudes), racial stereotypes (beliefs), and racism (ideologies) that may also be associated with racial disadvantage. Discrimination may be motivated by prejudice, stereotypes, or racism, but the definition of discrimination does not presume [an] underlying cause.

Discrimination in any form is challenging to study due to the methodological complexities of witnessing and recording such an action (Gaddis, 2018a; 2019). Rather than eliminating discrimination, anti-discrimination policies have mainly changed the tone and means of discrimination (Gaddis, 2019). This shift has produced a multi-faceted, although somewhat haphazard, methodological approach to studying the problem of discrimination. Simultaneously, the amount of research on discrimination has increased drastically over the past few decades (Crabtree et al., 2021).

Empirical examinations vary widely in precisely what is reported about discrimination due to their locus of action, the perspective of discrimination, and the type of data and method (Gaddis, 2024). The three potential loci of action are at the (1) individual level of those experiencing discrimination, (2) individual level of those enacting discrimination, or (3) from a higher group or organizational level. The perspectives of discrimination are perceived (e.g., a self-report from an individual experiencing discrimination), presumed (e.g., a statistical residual after accounting for all known important covariates), intended (e.g., a recorded outcome from a partially covert experiment), or observed (e.g., a recorded outcome from a fully covert experiment). Finally, the type of data and method can include administrative, surveys, interviews, time diaries, and experiments.

The depth and breadth of evidence of racial/ethnic discrimination vary across contexts due to researchers' differential levels of access to individual respondents, ability to conduct experiments in different contexts, and availability of administrative data, among other issues. In the K-12 education context, evidence of racial/ethnic discrimination often comes from surveys or interviews about perceived discrimination from students experiencing discrimination. These studies generally find that perceived racial discrimination has a detrimental effect on academic, behavioral, and psychological outcomes (Benner et al., 2018; Dotterer, McHale, and Crouter, 2009; Priest et al., 2013; Smalls et al., 2007). However, relying on these types of studies to understand discrimination in education is limiting in many dimensions. First, studies using surveys or interviews about perceived discrimination do not provide evidence about the prevalence of discrimination because not all forms of discrimination are accurately noticed and reported by the targets of discrimination. Second, these studies rarely record information on *specific instances* of discrimination – including the actor and scenario – that would be useful in unpacking potential mechanisms and moderators of discrimination. Third, these studies often do not examine racial/ethnic groups beyond Black students, and often capture limited experiences at a few schools or in a single city. Finally, and perhaps most importantly, surveys or interviews about perceived discrimination from students experiencing discrimination do not capture the type of *observed* evidence that would provide causal evidence of racial/ethnic discrimination.

Beyond the bulk of survey and interview research on perceived discrimination, other work sometimes examines intended discrimination from individuals enacting discrimination through partially covert experiments. This line of work typically uses vignettes or document review experiments with respondents – often teachers or counselors – to examine racial/ethnic discrimination. These types of studies capture intended racial discrimination on disciplinary

recommendations, course placement recommendations, grading, and writing evaluations (Francis, de Oliveira, and Dimmitt, 2019; Gilliam et al., 2016; Malouff and Thorsteinsson, 2016; Okonofua and Eberhardt, 2015; Owens, 2022; Quinn, 2020). While these studies shift the locus of action from individuals experiencing to individuals enacting discrimination, they still occur in research settings and record actions that do not have specific consequences or stakes for the respondents, and the link between the outcomes observed and real-world outcomes is unclear and often contested. Moreover, these survey outcomes may still be subject to social desirability or cognitive-behavioral disconnect bias (Gaddis, 2022). Finally, these studies also rarely examine racial/ethnic groups beyond Black students and often have relatively small sample sizes, limiting the ability of researchers to make claims about discrimination with precision and their capacity to examine treatment effect heterogeneity.

Valuable, but Sparse: Racial/Ethnic Discrimination Correspondence Audits in Schools

Although useful, the studies discussed above fail to capture *observed* evidence of racial/ethnic discrimination – something *only* fully covert field experiments can do. Correspondence audits are the primary type of covert field experiments researchers use to observe the existence of racial/ethnic discrimination (Gaddis, 2018b). Correspondence audits are a specific type of field experiment in which researchers manipulate a signaled characteristic (e.g., race, gender, sexual orientation) to examine discrimination based on that characteristic. Using correspondence audits, researchers can covertly examine actual behavior rather than answers about attitudes, opinions, or self-reported behavior on surveys that may be influenced by social desirability bias. Scholars have used correspondence audits to examine discrimination based on many individual characteristics, including race/ethnicity, religion, class, gender, age, sexual orientation, and others (Gaddis, 2018a). This research confirms the persistence of racial/ethnic

discrimination over time in the labor market, housing, higher education, and other contexts (Gaddis et al., 2021). Moreover, there is no substantial evidence that racial/ethnic discrimination has dissipated over time (Quillian et al., 2017).

Despite an overall strong body of evidence on racial/ethnic discrimination correspondence audits, few to date document the existence and scope of any discrimination in the public school context. To our knowledge, only three published studies use a correspondence audit to examine discrimination within the U.S. K-12 education context. First, Pfaff and colleagues (2021) email public school principals with embedded signals of parents' religiosity. The authors find that principals are significantly less likely to respond to parents' emails that signal they are Muslim or atheist. In a second study, Oberfield and Incantalupo (2021) send emails to public and charter school principals with signals of race/ethnicity (i.e., Black or White) and student ability. They find that discrimination against Black families is concentrated among White principals and against Black students with no signal of ability. In the final study, Rivera and Tilcsik (2023) also email public school principals posing as parents and disclosing whether their child has a disability. The authors find that principals are significantly less likely to respond to parents' emails indicating their child has a disability, particularly if the parent is Black.

While these last two studies are vital to understanding some specific forms of discrimination in the K-12 education context, they do not provide a comprehensive aggregate view of racial/ethnic discrimination against multiple groups. Moreover, these studies provide no details of potential heterogeneity in racial/ethnic discrimination in the K-12 education context – neither by individual nor contextual factors. We expand on the Rivera and Tilcsik (2023) study by (1) examining racial/ethnic discrimination in 33 states (compared to their four), (2) using four racial/ethnic groups – White, Black, Hispanic, and Chinese American (compared to their two;

i.e., Black and White), and (3) testing for heterogeneity in discrimination based on randomly manipulated individual-level resource needs and school-level resource strain. Our larger, more comprehensive study provides new additional insight into the “what,” “where,” and “when” of racial/ethnic discrimination in the K-12 education context while simultaneously setting up new research to investigate the potential mechanisms of and solutions to this discrimination (Gaddis, 2019).

The Central Roles of Principals and Resources

Public school principals, in particular, are the front-line officials entrusted with opening avenues into schooling. Some scholars argue that over time, principals’ roles have shifted from a teacher of teachers to the central conduit through which a school’s culture and connections flow (Brown, 2005). Principals occupy a mediating position between teachers and parents, teachers and district officials, and schools and the public. They are obliged to ensure competence, access, and equity in public education. However, rapid demographic, social, and cultural change may be challenging for public school officials, straining their capacity to ensure equitable and fair treatment. Principals remain much less culturally and ethnically diverse than the students they serve (Cooper, 2009; Evans, 2007). For example, whereas more than 80% of K-12 public school principals are White, less than 3% identify as “other,” a category that includes Asian-Americans (Goldring, Gray, and Bitterman, 2013). This mismatch raises questions about the ability of public schools to deal fairly with students of varied backgrounds, particularly those from groups regarded as newcomers or outside the American mainstream (Cooper, 2009; Evans, 2007).

Principals, however, are more than faceless bureaucrats directing schools at a macro level. They are individuals with their own biases. Research shows that not only is the selection process of principals rife with bias favoring White men but that, once on the job, White male

principals themselves exhibit both racial and gender bias (Bartanen and Grissom, 2023; Gullo and Beachum, 2020; Harris, 2022; Jarvis and Okonofua, 2020; Young and Fox, 2002). These biases appear in recommendations for student discipline, teacher hiring, assistant principal promotion, and more (ibid). Thus, we might expect that some principals, in the aggregate, would exhibit bias against students and families from certain racial/ethnic groups.

Principals also face incredible stress stemming from responsibilities and demands from parents, teachers, and district and state officials, perceived limitations in their authority and discretion to handle issues, and increased pressure of modern accountability systems (Mitani, 2018; Wang, Hauseman, and Pollock, 2021; West, Peck, and Reitzug, 2010; West et al., 2014). For example, principals whose schools are deemed failing by accountability measures can face stigma and sanctions, including losing their jobs (Holbein 2016; Holbein and Hassell 2019). Research suggests that principals' responses to these systems vary by how well their students perform and the socioeconomic background of the student body they serve (Carnoy, Elmore, and Siskin, 2003; Finnigan and Stewart, 2009; Ladd and Zelli, 2002).

This pressure is particularly severe for principals in schools with greater numbers of higher-needs students (e.g., FRL students and ESL students). These schools have high resource strain; they are stretched thin from providing extra resources for a high percentage of their student body. Principals of these schools have a more difficult time retaining good teachers (Baker and Cooper, 2005; Grissom, 2011), serve families who are less involved with school (Gutman & McLoyd, 2000; Lee & Bowen, 2006), and oversee teachers who have low expectations for their students and the school (Boser, Wilhelm, & Hanna, 2014; Diamond, Randolph, & Spillane, 2004). These principals often feel like they occupy a “middle-

management” role in which they can do little to shape their school’s performance (West et al., 2014).

Principals in schools with high resource strain may be desperate to reduce stress and do anything they can to right the ship. Although principals cannot deny access to families that move to their school’s catchment area, they may be less inclined to assist families who might *potentially* enroll their child in a principal’s school, especially if they believe that child would contribute to the resource strain. Initial communication with a prospective family might be one of the only times a principal has any influence on shaping their school’s student body. Thus, racial/ethnic discrimination during principals’ initial interactions with families may be targeted toward families with stated or perceived high resource needs, or among principals whose schools have high resource strain, or both.

The Importance of an Email

In this research, we examine public school principals’ responses to an email query from parents of different races/ethnicities. On the surface, ignoring someone’s email might seem a minor issue. However, we argue that existing work on this form of discrimination and the general impact of discrimination on parents suggests that this may have significant downstream effects. Furthermore, a non-response from a principal may indicate a more hostile racial/ethnic environment in the school.

The direct effect of a principal’s discrimination against parents and the subsequent indirect effect on children is important on its own. When racially/ethnically motivated, ignoring someone else is one critical form of what scholars refer to as everyday racism (Essed, 1991), papercut discrimination (Block et al., 2021), or microaggressions (Sue et al., 2007). Racial/ethnic minorities are acutely aware of these types of subtle discrimination and often attribute them to

their race/ethnicity (Wang, Leu, and Shoda, 2011). Additionally, discrimination can lead parents to alter the racial-ethnic socialization of their children (Umaña-Taylor and Hill, 2020; Wang et al., 2020). Experience with discrimination increases the likelihood that parents prepare their children for bias and promote mistrust of other racial/ethnic groups (Hughes et al., 2006; Woo et al., 2020). Research on preparation for bias and promotion of mistrust suggests that these forms of racial-ethnic socialization have adverse effects on academic achievement, depressive symptoms, stress, and behavioral problems (Atkin, Woo, and Yeh, 2019; Daga and Raval, 2018; Liu and Lau, 2013). Thus, when parents experience discrimination in the public school context, they may prepare their children for the same experience in ways that have long-term negative consequences.

Furthermore, evidence of discrimination by school principals may be indicative of a broader negative school climate. Principals are at the head of school leadership and set the tone for the entire school, including the racial/ethnic climate, attention to inequalities, and the value placed on diversity (Grissom, Egalite, and Lindsay 2021). The school climate generally is shaped from the top down. It impacts numerous aspects of teachers' and students' lives, including physical and mental health, substance abuse, absenteeism, learning motivation, academic success, and occupational satisfaction (Thapa et al., 2013). The racial climate of a school, in particular, can influence academic achievement, disciplinary referrals, and school belonging (Griffin and Allen, 2006; Mattison and Aber, 2007; Voight et al., 2015). Thus, the behavior we record may foreshadow additional negative attitudes and behaviors that racial/ethnic minorities face in these schools.

RESEARCH QUESTIONS AND CONTRIBUTIONS

In this research, we address three research questions. First, to what extent do K-12 public school principals discriminate on the basis of race/ethnicity against families seeking to gain more information about and enroll their child in a school? Second, does this racial/ethnic discrimination, if any, vary when a family signals greater individual-level resource needs? Third, does this racial/ethnic discrimination, if any, vary when the principal's school has greater school-level resource strain?

Our research joins only two prior studies examining racial/ethnic discrimination by public school principals using a field experiment. Our study is similar in that it adds to a desperately underdeveloped line of educational inequality research capturing observed racial/ethnic discrimination among the individuals enacting discrimination. Moreover, we build upon the prior two field experiments by examining racial/ethnic discrimination (1) across a much broader population of schools and states, (2) against more racial/ethnic groups, and (3) by important individual-level and school-level resource characteristics to test for heterogeneity in discrimination based on the resource triage thesis.

DATA AND METHODS

This paper reports the results of a novel correspondence audit conducted before the COVID-19 pandemic¹ with a second similar experiment as a robustness check (included in Appendix A). With this large-scale field experiment, we tested for discrimination against Black and Hispanic families as well as against the largest Asian subgroup in the United States—

¹ The COVID-19 pandemic may have exacerbated, attenuated, or done little to the effects that we observe. For instance, the COVID-sparked rise in anti-Asian sentiments may have led to more discrimination against Asians (see Lu et al., 2021). However, this is ultimately an empirical question; we cannot know for sure without replicating our work in the future.

Chinese Americans. We provide additional methodological details and supplementary analyses in the Appendix.

In our experiment, we emailed the principals of 52,792 public K-12 schools in 33 U.S. states. These states were AL, AR, CA, CO, DE, FL, GA, IA, ID, IL, IN, LA, MA, MI, MN, MO, MS, NC, ND, NE, NH, NJ, NM, NY, OH, RI, SC, TN, TX, VA, VT, WA, and WI. We included all states for which we could acquire principals' email addresses either by contacting state Departments of Education or by downloading contact information from the websites of those institutions. We restricted our sampling frame within these states to principals of regular, operational, non-charter, public K-12 schools.

We wrote each email ostensibly from a parent who claimed he was moving into the area, looking at local real estate, and wanted to “meet with [the principal] or a member of [their] staff to talk about [their] school.” In our emails, we manipulated the racial/ethnic identity of the parent sending the email and their child via names included with each message. We randomly assigned the racial/ethnic identity of the family (i.e., White, Chinese American, Black, or Hispanic), with a roughly equal number of principals receiving correspondence from each of our conditions. We sent only one email to each principal to avoid experiment discovery (Larsen, 2020; Vuolo, Uggen, and Lageson, 2018). We used a scripted email design to automate and standardize sending our messages (Crabtree, 2018).

The body of each email identified either a son or a daughter as the prospective student; we randomized the child's gender. The children's names we used in this experiment were Molly (White daughter), Connor (White son), Shanice (Black daughter), DeShawn (Black son), Ying (Chinese American daughter), Wei (Chinese American son), Isabella (Hispanic daughter), and Diego (Hispanic son). We paired each first name with a corresponding racial/ethnic last name

using data from the U.S. Census on the most commonly occurring surnames: Erickson (White: 96.4%), Washington (Black: 89.9%), Wang (Chinese American: 94.5%), and Vazquez (Hispanic: 94.5%) (U.S. Census Bureau, 2012). Our effect estimates are statistically and substantively indistinguishable regardless of whether the child was male or female.² Independent of race/ethnicity, male and female children received responses at the same rate ($p = 0.37$).

At the end of each email, we signed a male parent's name as the prospective student's father; all parents' names were male. The fathers' names used in this experiment were Jake (White parent), Fong (Chinese American parent), Tyrone (Black parent), and José (Hispanic parent). We chose these first names for the children and parents because prior research shows that people commonly perceive these names at high rates as belonging to an individual from the racial/ethnic group we wanted to signal (Crabtree et al., 2023; Gaddis, 2017a; Gaddis, 2017b; Gaddis, Kreisberg, and Crabtree, 2023). Additionally, demographic birth data indicate that these names overwhelmingly belong to individuals from the same signaled racial/ethnic groups (ibid).

We signaled individual-level resource needs with two different signals. First, for each of the racial/ethnic treatments, we included an additional treatment that signaled individual-level resource needs through a statement in the email that the family previously qualified for free lunch. Approximately 50% of our emails for each of the four racial/ethnic groups included this signal. Second, for the Chinese American and Hispanic treatments, we included an additional treatment that signaled individual-level resource needs through a statement in the email that the student previously needed assistance as an English as a Second Language (ESL) student.

Approximately 50% of our emails for the Chinese American treatment and the Hispanic

² For example, male Chinese American $\beta = -10.7$ percentage points, 95% CI = [-12.5, -9.0]; female Chinese American $\beta = -9.7$ percentage points, 95% CI = [-11.5, -8.0]; difference between male and female Chinese American $\beta = -1.0$ percentage point, 95% CI = [-2.5, 1.0].

treatment included this signal. We show the descriptive statistics for our treatment conditions in Table 1.

<Insert Table 1 here>

Given the parameters of this study, it is difficult, if not impossible, to measure school-level resource strain directly. We envision resource strain as how thinly stretched a school is to provide for its student body. Although we cannot measure resources, we can measure the population that likely *needs* more resources. Thus, we identified two obtainable measures that we believe are sufficient proxies for school-level resource strain. We captured these measures using school-level data from the National Center for Education Statistics (NCES) and Census-tract data from the American Community Survey (ACS) 3-year. First, we matched school identifiers in NCES data to record the percentage of students in a school receiving free- or reduced-price lunch. Second, we matched addresses to Census tracts in ACS data to record the percentage of residents below the poverty line.

We intentionally selected a large sample of public school principals. We wanted to ensure that our estimates were sufficiently powered to avoid errors often arising in underpowered research designs (Gelman and Carlin, 2014). A large sample size also allows us to examine heterogeneity in individual-level resource needs and school-level resource strain. This large sample size is justified because previous research has shown that estimating treatment effect heterogeneity requires much more power than many researchers assume, and thus, subgroup tests are often underpowered (Coppock, Leeper, and Mullinix, 2018). As Andrew Gelman (2018) shows, tests of treatment effect heterogeneity require around 16 times as many observations to obtain sufficient statistical power as main effects.

Similar to other correspondence audit studies of public officials, bureaucrats, and gatekeepers, our outcome of interest is whether the recipient responded to the email request. We re-coded response data into a binary: a positive email response became a “1,” and a negative or no email response became a “0.” We choose not to examine message response quality (e.g., response tone) because recent research shows that doing so can induce post-treatment bias (Coppock, 2019; Montgomery, Nyhan, and Torres, 2018). In short, conditioning a second-order outcome on a first-order outcome essentially breaks the randomization of the experimental design. Thus, scholars strongly recommend against conditioning on post-treatment variables when analyzing experimental data (Coppock, 2019). Additionally, we examined a random sample of 500 replies and found no meaningful substantive differences in reply content or quality.

To determine the effect of our race/ethnicity treatments, we estimate a set of ordinary least-square (OLS) models that regress our binary outcome measure for whether we received a positive principal response (our dependent variable) on our race/ethnicity treatment indicators (our primary independent variables). We use linear regression over logistic regression for interpretability and because recent research has shown that linear regression is preferable for many reasons (Gomila, 2021; Hellvik, 2009); our conclusions remain unchanged, however, if we use logistic regression. In our preferred models, we also include all additional non-race/ethnicity treatment indicators (i.e., gender, FRL, and ESL signals) and state fixed effects. To account for possible heterogeneity in our error term, we use robust standard errors. The reference category across all models is the “White” racial condition.

From these models, we calculate adjusted predictions at the means for the response rate of each of the four racial/ethnic profiles and, when appropriate, by individual-level resource

needs and school-level resource strain. We then translate these into discrimination ratios, defined as the response rate for White parents divided by the response rate of the three other racial/ethnic profiles separately. Each racial/ethnic response rate is the number of positive responses divided by the number of total emails sent for that profile. This is represented by the following equation:

$$DR_{wm} = \frac{\frac{r_w}{n_w}}{\frac{r_m}{n_m}}$$

Where DR_{wm} is the discrimination ratio for Whites (w) and racial/ethnic group m , r_w is the number of positive responses for White parents, n_w is the number of total emails sent for White parents, r_m is the number of positive responses for racial/ethnic group m parents, and n_m is the number of total emails sent for racial/ethnic group m parents. After we calculate each discrimination ratio, we transform this measure using the natural logarithm of the discrimination ratio. This allows us to work with a dependent variable that is approximately normally distributed. When calculated this way, the discrimination ratio is a relative, rather than an absolute, measure that allows for straightforward interpretation across different groups. The discrimination ratio is effectively a risk ratio or relative risk measure and is increasingly used in similar types of research (Gaddis et al., 2021). Moreover, by calculating discrimination ratios, we can directly compare our results to prior research to better understand the size and severity of the discrimination we find in our study. We provide additional details on the discrimination ratio metric in Appendix C.

RESULTS

We start by estimating the predicted discrimination ratio for each of the three non-White racial/ethnic groups (i.e., Black, Chinese American, and Hispanic). Figure 1 shows the discrimination ratios for each of these groups across four types of models. Model 1 includes only race/ethnicity treatment indicators; model 2 includes all treatment covariates; model 3 includes all treatment covariates and block fixed effects; and model 4 includes all treatment covariates and state fixed effects.

[Figure 1 about here]

The results show that the model specification has virtually no impact on our results because this is a well-designed experiment with a large sample size. The discrimination ratios for White/Black parents cross the gold line at 1, indicating no statistically significant evidence of discrimination in the aggregate, averaging across other experimental treatments. The discrimination ratios for White/Chinese American parents are centered at 1.22, indicating that White parents receive approximately 22% more responses than Chinese American parents. The discrimination ratios for White/Hispanic parents are centered at 1.06, indicating that White parents receive approximately 6% more responses than Hispanic parents. To put these results in perspective, the discrimination against Chinese American parents is less than the discrimination Black job-seekers face in the labor market ($DR = \sim 1.30$) and more than the discrimination Black housing-seekers face in the rental housing market ($DR = \sim 1.18$) according to a recent meta-analysis (Gaddis et al., 2021). The discrimination against Hispanic parents is about the same as Hispanic housing-seekers face in the rental housing market (ibid). These effects for Chinese American parents are large. By any comparison point – whether relative to racial discrimination found in previous audit studies from a variety of contexts, other discrimination ratios estimated

in our study, or the base response rate – these effects are substantively meaningful. In short, our results show that public school principals exhibit high rates of discrimination against Chinese American parents and, to a slightly lesser extent, Hispanic parents.

[Figure 2 about here]

[Figure 3 about here]

To examine potential heterogeneity in discrimination, we first examine whether these discrimination ratios vary by individual-level resource needs, as shown in Figures 2 and 3. Figure 2 shows three discrimination ratios for each racial/ethnic group: (1) when both the racial/ethnic group and the White comparison group do not signal FRL, (2) when the racial/ethnic group signals FRL and the White comparison group does not signal FRL, and (3) when both the racial/ethnic group and the White comparison group signal FRL. Figure 3 shows two discrimination ratios for Chinese American and Hispanic families: (1) when the racial/ethnic group does not signal ESL and (2) when the racial/ethnic group signals ESL.

The results from Figure 2 show no clear pattern of heterogeneity in discrimination by individual-level resource needs across racial/ethnic groups. There is no significant heterogeneity in the discrimination ratios by individual-level FRL status for Chinese American and Hispanic families. The discrimination ratio for Black families is significant in one case – when Black families signal FRL and White families do not. The discrimination ratio is modest ($DR = 1.05$) and significant. The results from Figure 3 show no significant heterogeneity in the discrimination ratios by individual-level ESL status for Chinese American and Hispanic families. Thus, we find evidence that discrimination varies by individual-level resource needs in only one of the five cases (i.e., Black families signaling FRL).

[Figure 4 about here]

[Figure 5 about here]

Next, we examine whether there is heterogeneity in discrimination by school-level resource strain. Figures 4 and 5 show four discrimination ratios for each racial/ethnic group by each quartile of the percentage of students in a school receiving free- or reduced-price lunch (Figure 4) and each quartile of the percentage of residents below the poverty line (Figure 5). There is no significant heterogeneity in the discrimination ratios by school-level percent FRL for Black and Hispanic families. However, discrimination is significantly greater for Chinese American families in schools in the highest two quartiles of percent FRL compared to schools in the lowest quartile of percent FRL ($p < 0.05$). We find a similar pattern when we examine heterogeneity in discrimination ratios by the percentage of residents below the poverty line. Again, there is no significant heterogeneity for Black and Hispanic families. However, discrimination is significantly greater for Chinese American families in schools in the highest two quartiles of percent below the poverty line compared to schools in the lowest quartile of percent below the poverty line ($p < 0.05$).

Our final examination focuses on combining individual-level resource needs and school-level resource strain (see Appendix D for the additional analyses). We find two critical differences when we examine heterogeneity in discrimination by the individual-level FRL signal and school-level percent FRL or percent of residents below the poverty line. First, the discrimination ratio for Black families in Figure 2 is only significant when the school has high resource strain (i.e., the highest quartile of FRL students). Second, the discrimination ratios for Chinese American families are significantly larger when they signal FRL and are in the highest quartile of FRL students or residents below the poverty line. These results are somewhat suggestive that high individual-level resource needs and high school-level resource strain might

combine to increase racial/ethnic discrimination in the case of Black and Chinese American families.

DISCUSSION

In this paper, we reported the results from a correspondence audit examining the behavior of public school principals. We build on prior literature by collecting a large sample size covering most U.S. states, examining discrimination against three racial/ethnic groups, and testing for heterogeneity in racial/ethnic discrimination by individual-level resource needs and school-level resource strain. We find considerable discrimination against Chinese American families, some discrimination against Hispanic families, and no significant discrimination against Black families in the aggregate. Except for one case – when Black families signal FRL and White families do not – we find that discrimination by principals does not significantly vary by individual-level resource needs. We also find that discrimination against Chinese American families was greater in schools with higher resource strain. Our research provides rare causal evidence of racial/ethnic discrimination against several groups using robust evidence from a field experiment.

There are at least three critical takeaways from our research that can help guide future research. First, our findings are the only to show that principals' racial/ethnic discrimination affects multiple racial/ethnic groups. Although the levels of discrimination vary across groups, no racial/ethnic group we study escapes unscathed. Additionally, while Chinese American families faced the highest levels of discrimination, Black families faced no significant discrimination except in specific circumstances. The finding is in line with Rivera and Tilcsik (2023), who found mostly targeted instances of discrimination. By studying multiple

racial/ethnic groups, we uncovered important differences in discrimination by racial/ethnic groups that did not necessarily align the way previous research might have predicted (i.e., Black people almost always face higher levels of discrimination than other racial/ethnic groups, according to nearly all field experiments on discrimination). Thus, one of the critical implications for future research is that scholars should try to include multiple racial/ethnic groups in all types of correspondence audits examining racial/ethnic discrimination. Such research would provide a broader scope and more complete analysis of racial/ethnic discrimination in the U.S. and perhaps uncover new pathways to test for mechanisms.

Could these aggregate-level discrimination rates against different racial/ethnic groups simply reflect differences in principals' race/ethnicity? Unfortunately, we were unable to collect or record data on the race/ethnicity of each principal we contacted. However, we believe it is unlikely that our findings are simply the result of principals' race/ethnicity for two reasons. First, few principals are not White; these rates are lower than national population demographics for race/ethnicity. Second, there are similar rates of principals who are Black compared to Hispanic principals. Thus, if principal race/ethnicity alone were driving these results, we might expect similar rates of discrimination for Black and Hispanic families. Nonetheless, future research could examine this issue in more detail using survey experiments in which researchers can easily collect data on race/ethnicity of the respondents.

Our second critical takeaway stems from two key points regarding heterogeneity in racial/ethnic discrimination by principals. For Black families, discrimination is moderated by signals of individual-level resource needs. For Chinese American families, discrimination is moderated by the school-level resource strain, although discrimination is still present and large regardless of school-level resource strain. These findings are important and the first to explore

discrimination heterogeneity in the public school context. Still, this research raises many new questions about precisely why this heterogeneity exists.

One potential explanation is that principals target discrimination toward families with stated or perceived high resource needs and in schools with high resource strain because principals fear scenarios that are likely to increase the burden on existing resources. This ‘resource triage’ thesis suggests principals’ decisions are influenced by considerations of future strain and stress. Such a phenomenon would be similar to how educational accountability systems induce some educators to engage in ‘educational triage’ to conserve and focus resources on students – those closest to passing standardized tests – who are most likely to reduce the stress on staff within schools (Booher-Jennings, 2005; Lauen and Gaddis, 2012; 2016). Other research finds that the pressure and stress of educational accountability may lead to efforts to push under-performing students out of school, at least temporarily (Holbein and Ladd, 2017; Jacob, 2005). As with educational triage, principals might engage in discrimination-based resource triage to conserve and focus resources in an effort to reduce stress.

Scholars conducting field experiments on racial/discrimination should always consider ways to examine heterogeneity by both randomizing additional treatments that represent potential mechanisms or moderators and using existing differences in organizational or geographic characteristics to their advantage. In the education context, more research is needed to provide additional data points on why principals engage in discrimination and how it might be related to resources. Future research should also pursue at least two new but related avenues: (1) other forms of heterogeneity in racial/ethnic discrimination by principals (i.e., non-resource-based characteristics), and (2) why principals choose not to respond to certain racial/ethnic

groups. Scholars could examine the first topic with additional field experiments, while they could explore the latter with survey experiments.

Our final takeaway stems from the finding that discrimination against Chinese American families is substantial. In schools with the highest levels of resource strain, Chinese American families face discrimination as large as the largest discrimination ratios *any* racial/ethnic group faces in *any* context (Gaddis et al., 2021). Moreover, these discrimination ratios dwarf what other groups experience in the public services and higher education contexts, the two domains of discrimination studies closest to ours. Few field experiments in the U.S. have examined discrimination against any Asian American subgroups. Our research suggests this may be a drastic oversight. While scholars should conduct additional research to help understand why discrimination against Chinese Americans is so high in the public school context, society would also benefit from additional field experiments examining discrimination against Asian Americans in any context.

Our work has some important limitations. First, we focused on measuring discrimination against only Chinese Americans – and not other Asian American groups – in this paper, given the size and visibility of this subpopulation. However, discrimination against Asian Americans could vary depending on the specific nation of origin and generational status. These are distinct possibilities given the large gaps in economic and social outcomes experienced within different sub-populations of Asians and by nativity. However, future research would do well to acknowledge that Americans may struggle to identify the ethnic origins of smaller groups successfully. Second, our research measured discrimination in only one context and stage, initial interactions with principals in the public education sector, albeit an important one. Public schooling is a vitally-important domain given schools' core role in contemporary society. Future

research could examine discrimination among different actors in the public educator context at different stages. Third, the kind of discrimination we uncovered is of a subtle and everyday variety rather than behavior expressing overt hostility. Our work should not be the last to use experimental research designs to test for discrimination against racial/ethnic minorities in the public education system.

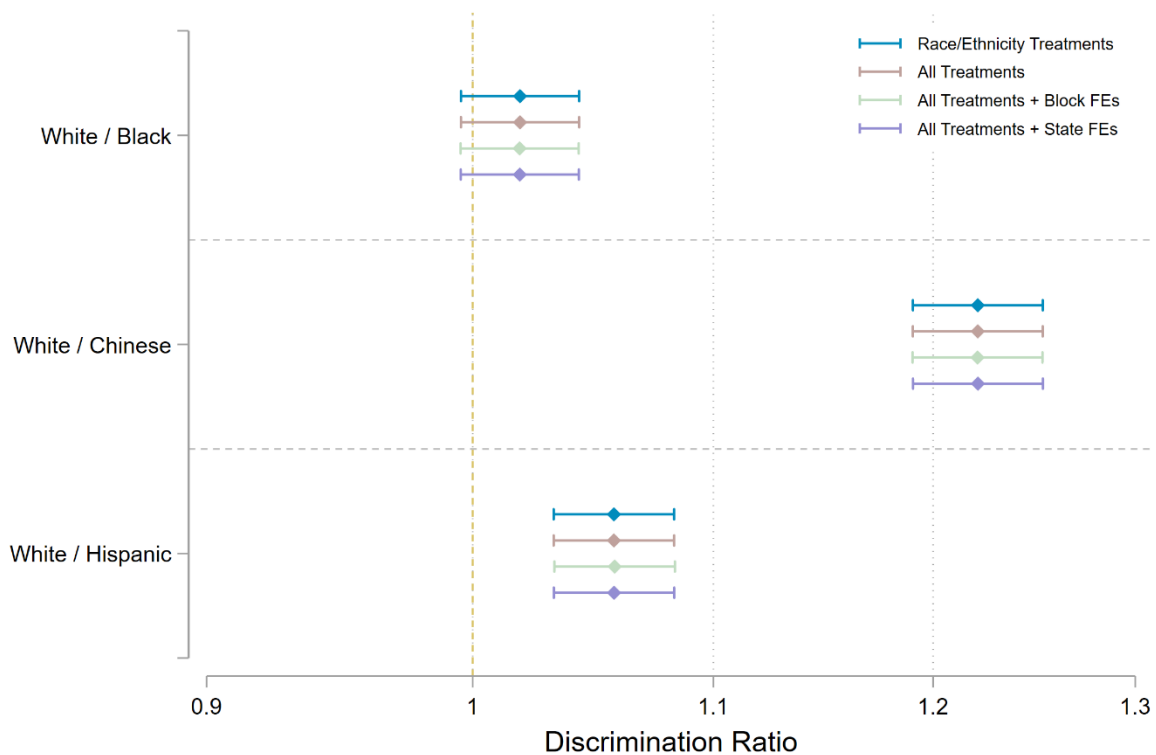
In summary, our research provides causal evidence that racial/ethnic minority families face varying level of discrimination in initial interactions with principals that is sometimes moderated by individual-level resources needs and school-level resource strain. Broadly, these findings highlight the need for researchers conducting future correspondence audits to expand the scope of their research to provide a more comprehensive analysis of racial/ethnic discrimination in the U.S. Moreover, additional research in the public school context may help us understand the mechanisms of this discrimination and, therefore, the tools we need to employ to reduce or eliminate it. As it currently stands, however, schools are not immune to the forms of racial/ethnic discrimination we see in other sectors of our modern society.

TABLES AND FIGURES

Table 1. Descriptive Statistics

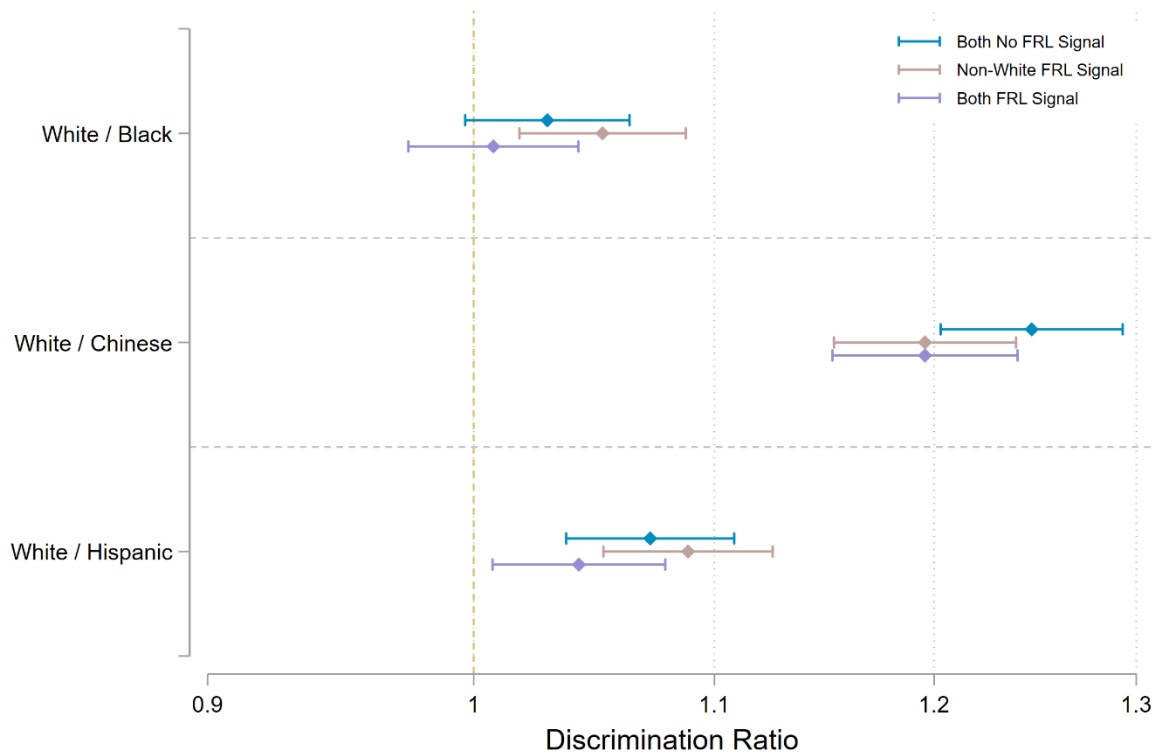
<u>Treatments</u>	<u>N</u>	<u>%</u>
White	13,199	25.0
Black	13,200	25.0
Hispanic	13,204	25.0
Chinese American	13,189	25.0
Male	26,393	50.0
Female	26,399	50.0
FRL signal	26,404	50.0
ESL signal	13,190	25.0
Total	52,792	

Figure 1. Discrimination Ratios by Model



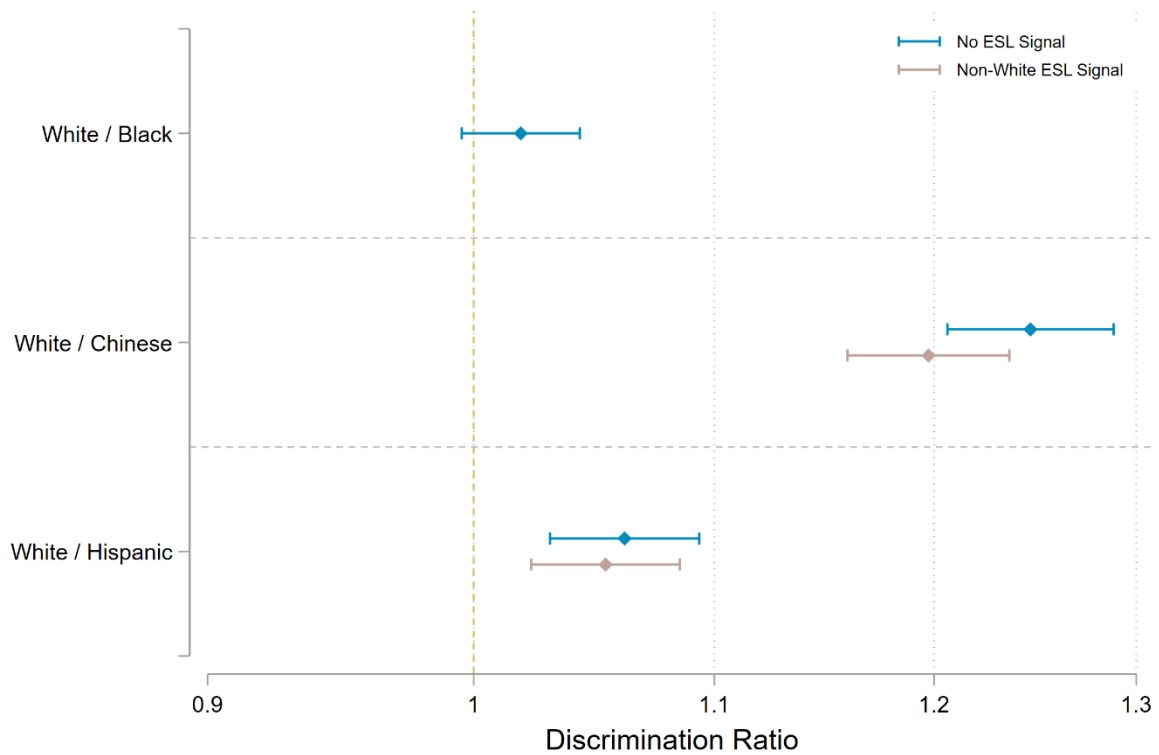
Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include (1) only race/ethnicity treatment indicators, (2) all treatment covariates, (3) all treatment covariates and block fixed effects, and (4) all treatment covariates and state fixed effects. The effects are not substantively different from each other across model specifications (within treatments). Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination.

Figure 2. Discrimination Ratios by Student FRL Status



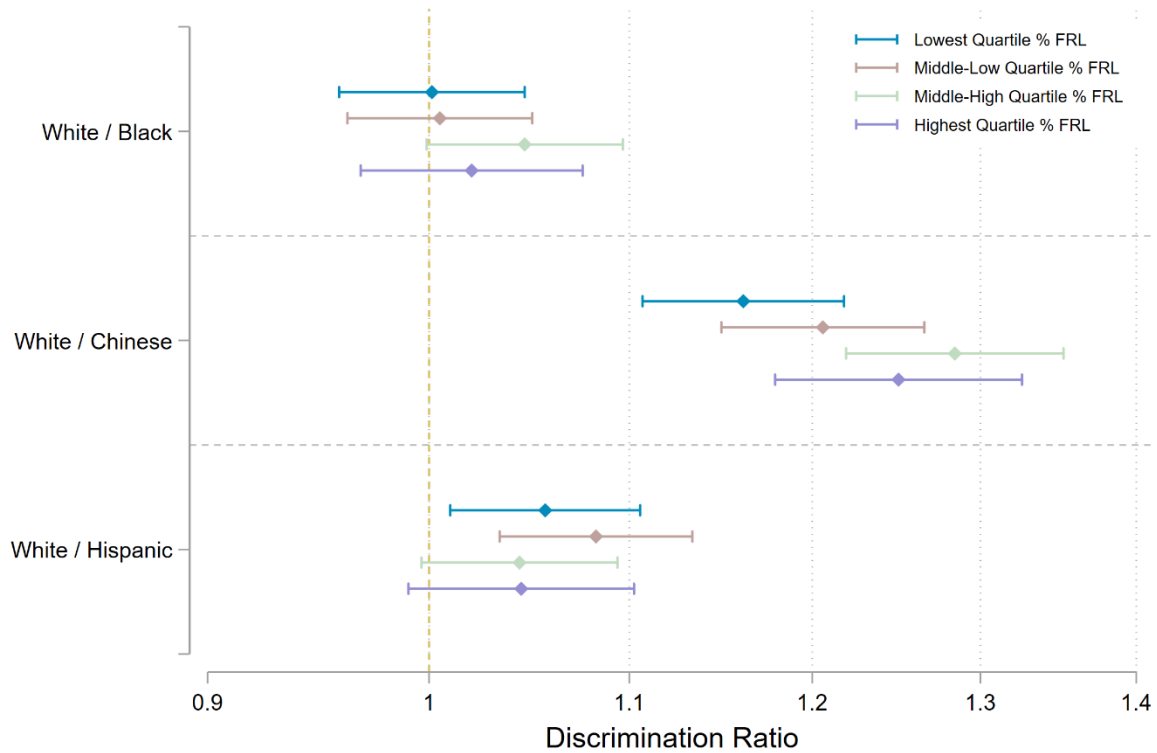
Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination.

Figure 3. Discrimination Ratios by Student ESL Status



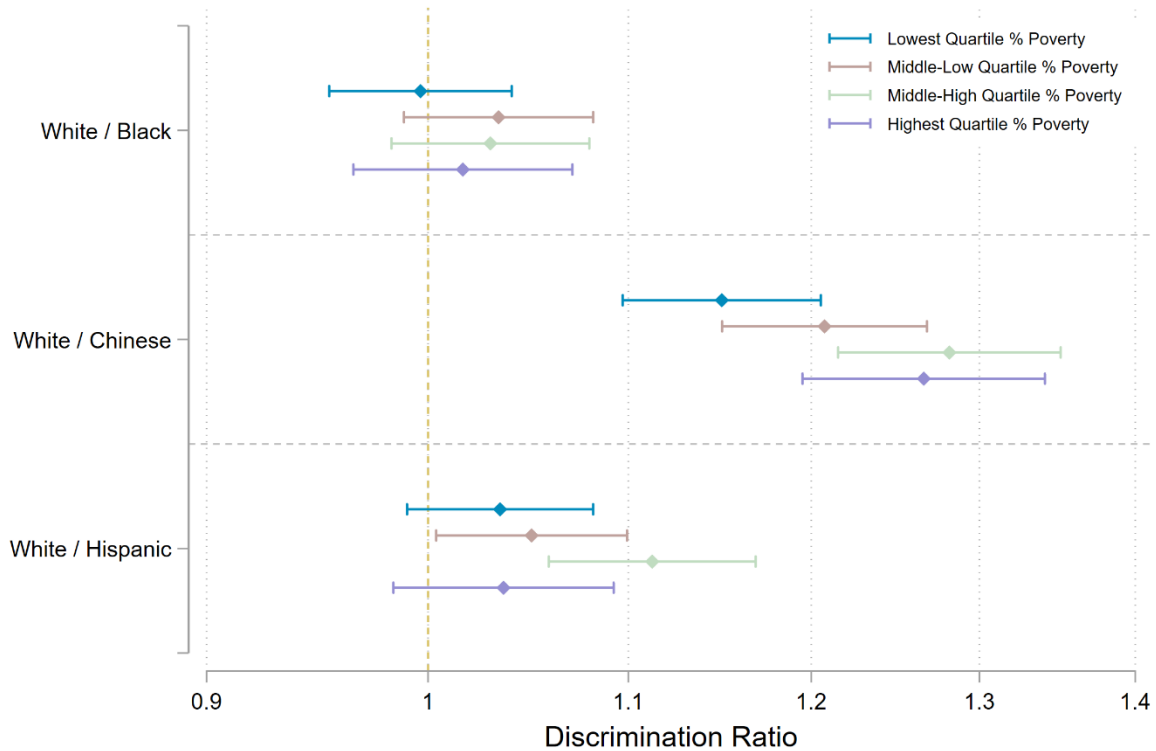
Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination.

Figure 4. Discrimination Ratios by the Percentage of FRL Students in the School



Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination. Percentage FRL students variable: Quartile 1 = $0 \leq 31.7\%$; Quartile 2 = $31.7\% \leq 53.1\%$; Quartile 3 = $53.1\% \leq 74.7\%$; Quartile 4 = $74.7\% \leq 100$.

Figure 5. Discrimination Ratios by the Percentage of Residents Below the Poverty Line in Census Tract



Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination. Percentage residents below poverty line variable: Quartile 1 = $0 \leq 8.7\%$; Quartile 2 = $8.7\% \leq 14.2\%$; Quartile 3 = $14.2\% \leq 19.7\%$; Quartile 4 = $19.7\% \leq 100$.

APPENDIX

A. ROBUSTNESS CHECK: SECOND EXPERIMENT

Independent from the first experiment, part of the current author team ran a similar correspondence audit (Experiment 2). We sent emails from parents to approximately 1,400 randomly selected public school principals in 2 U.S. states: Kentucky and North Carolina. The content of the correspondence was a simple request for information about a school's music and art courses. Like our first experiment, we randomly assigned the racial/ethnic identity of the family (i.e., White, Chinese American, Black, or Hispanic), with a roughly equal number of principals receiving correspondence from each of our conditions.

In this experiment, the body of each email identified a son as the prospective student. The sons' names used in this experiment were: Jacob (White), Yingpei (Chinese American), DeShawn (Black), and José (Hispanic). At the end of each email, we signed a female parent's name as the prospective student's mother; all parents' names were female. The mothers' names used for the racial treatments were: Emma (White), Zhi (Chinese American), Deaundra (Black), and Maria (Hispanic). We chose these names for the same signaling and demographic reasons discussed above. We paired each first name with a corresponding racial/ethnic last name using data from the U.S. Census on the most commonly occurring surnames: Mueller (White: 97.0%), Wang (Chinese American: 94.5%), Washington (Black: 89.9%), and Juarez (Hispanic: 94.7%). Again, we sent only one email to each principal.

Because the two experiments were independently designed and executed, there are six noteworthy differences between them. First, each experiment signaled different genders of parents and children. Second, each experiment signaled race/ethnicity with different specific names. Third, each experiment had a different type of request – a higher cost request (i.e., a

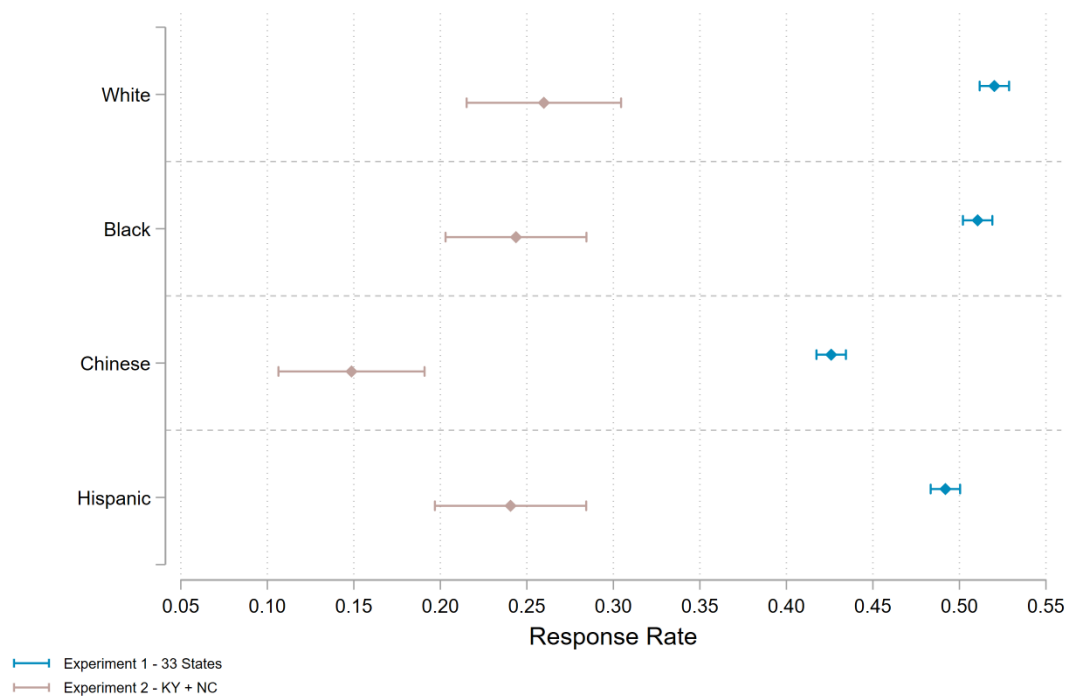
request to meet) in Experiment 1 and a lower cost request (i.e., information gathering) in Experiment 2. Fourth, each experiment used different email servers to process emails. Fifth, each experiment took place during different months and years. Finally, each experiment sampled principals from different sets of states, although we included North Carolina in both experiments.

We believe these differences are a distinct strength of our overall research. Rather than being a weakness, our two experiments' independent designs and executions add to the robustness of our findings. Ultimately, our results are remarkably similar across the two experiments, suggesting that the discrimination we document is quite robust. In Appendix Figure A1, we show the raw response rates and the discrimination ratios for each of the racial/ethnic groups in each experiment. Although the response rates are higher in Experiment 2, the overall results are similar. In short, aggregate-level discrimination against Chinese American families is much higher than against other racial/ethnic groups. However, we cannot conduct additional analyses of heterogeneity in Experiment 2 due to the much smaller sample size.

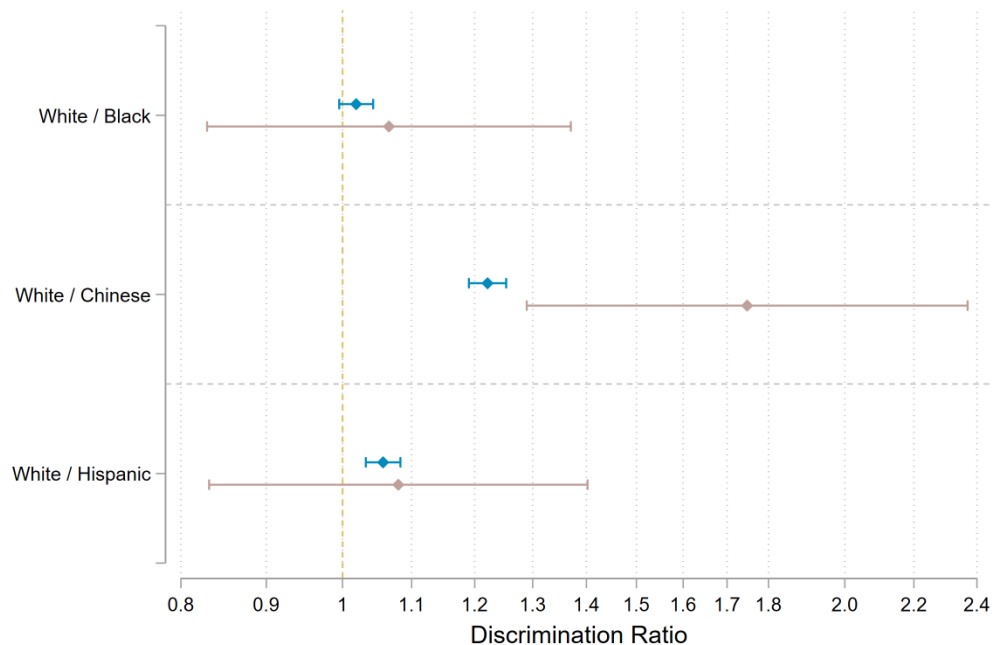
Our experiments provide a unique opportunity to replicate the results independently within a single article. This robustness check is a critical advantage to our research, given the replication crisis many social science fields have faced in recent years.

Appendix Figure A1. Response Rates and Discrimination Ratios by Experiment

A.



B.



Note: Points are raw response rates (panel A) or predicted discrimination ratios (panel B), and horizontal lines are the 95% confidence intervals. The size of the confidence intervals reflects the sample size difference between the two experiments. Experiment 1 includes 52,792 principals in 33 states. Experiment 2 includes 1,444 principals in 2 states. The underlying models are OLS regression models that include only race/ethnicity treatment indicators calculated with robust standard errors. Panel B results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination.

B. OVERALL DETAILS OF DATA DESIGN AND COLLECTION

B1. EMAIL MESSAGES

Below, we provide the exact email messages that we sent to principals in each experiment. Each message was sent from a randomly assigned White, Chinese American, Hispanic, or Black parent, as detailed in the main text.

Appendix Figure B1. Email Messages for Experiment 1

From: [Randomized White/Chinese American/Black/Hispanic Sender]
To: [Principal]
Subject: Appointment request

Dear [Principal's Name],

Hello. My family and I are moving to the area this summer. We are considering several neighborhoods and are looking at schools for our [daughter/son], we would like to meet with you or a member of your staff to talk about your school.

Sincerely yours,
Randomized White/Chinese American/Black/Hispanic Sender

Appendix Figure B2. Email Messages for Experiment 2

From: [Randomized White/Chinese American/Black/Hispanic Sender]
To: [Principal] Subject: Music & Art at Your School

Good Evening,

My name is [Randomized White/Chinese American/Black/Hispanic] and I'm a parent who's considering moving to your area in the next few months. I'm emailing you on behalf of my son [randomized son's name White/Chinese American/Black/Hispanic].

I'd like to know a little bit about your school. Could you tell me a little bit about the music and art programs your school offers? [Son's name] is very creative. I want to make sure he has the opportunity to develop his talents in these areas.

I'm emailing a few other schools in your area to see what they have to offer as well. However, I'd really appreciate hearing about your school.

I look forward to hearing from you. Thanks for your time.

Sincerely,
Randomized White/Chinese American/Black/Hispanic Sender

B2. ETHICAL CONSIDERATIONS

Our study was reviewed by Human Subjects Committees at Pennsylvania State University (STUDY00002386), the University of Michigan (HUM00118066), and Duke University (B0361). To ensure that our experiment would not allow the identification of individuals, we have anonymized the replication files.

In our experiments we had one form of deception – the name of the parent emailing principals. In order to study the effect of race/ethnicity on responses, we used aliases in the email address and email signature. This manipulation is standard practice in correspondence audits trying to examine racial/ethnic and other forms of discrimination.

The potential costs of this deception are minor compared to the potential gains of our research. The existing literature on biases against Asians in America lacks consensus about the existence, extent, and causes of discrimination. Given the political consequences of discrimination for both in- and out-group members, we think that much more well-powered experimental research needs to be done on this important subject. One possible reason for the conflicting findings in the literature is that studying discrimination is very hard and involves many important research design challenges. These issues include thorny statistical concerns such as selection, confounding, and post-treatment bias. The audit study approach we take allows us to sidestep these statistical issues and credibly measure the degree to which public school principals discriminate against Chinese American parents. We think that the gains in credibility offered by this approach outweigh both the costs imposed by slightly deceiving participants and the societal costs created by using a different, less-accurate, and less-powered research design to assess discrimination.

Our sample size was large. However, a large sample size by itself does not automatically make a study unethical. Some may argue that the cumulative effect of the time spent on our

study makes it unethical. It is not clear to us that this type of calculation is actually useful or relevant. Adding up the time spent on a study by a collective pool of research subjects is very odd, because no individual actually bears that cost. One could say that future researchers bear that cost, but that is true of all social science studies, not just ours. This type of argument is never discussed in reference to the time spent by the participants of large-scale surveys fielded in the social sciences, such as the ANES, CES, and CPS. Given that respondents to those studies spend hours, not minutes (like our study), on completing those surveys, we might be much more worried about those studies than ours. Thus, we do not think that a cumulative time calculation is the best way of quantifying an “ethics bar” for a study. Establishing the true costs and benefits of a study requires much more nuance and depth. This large sample size was intentional. We chose such a large sample to ensure that we could avoid some of the problems that come with underpowered research designs. Previous research has shown that estimating heterogeneity in treatment effects requires much more statistical power than many researchers suspect and, thus, many heterogeneity tests are woefully under-powered. Moreover, we anticipated that reviewers might want us to estimate additional (post-hoc) tests of treatment effect heterogeneity. Tests of treatment effect heterogeneity demand a much higher degree of statistical power than is often realized. For example, Andrew Gelman – prominent social scientist and methodologist – has as a rule of thumb that tests of treatment effect heterogeneity require 16 times the statistical power as main effects.

In total, our design features closely follow the ethical guidelines of the American Sociological Association and the American Political Science Association for ethical behavior in survey recruitment and in field experiments.

B3. SAMPLING FRAME

In Experiment 1, within the 33 states we examined, we dropped all schools with missing principal contact information. We also excluded schools that could not be uniquely matched to NCES (National Center for Education Statistics) data and schools with missing covariate data in the NCES or American Community Survey (ACS). Based on state and NCES data, we dropped inactive, private, charter, non-traditional, adult, and virtual schools as well as schools serving restricted populations such as schools for the blind and deaf and schools located on military bases. We also excluded schools with less than 100 students, schools that are majority American-Indian, and schools that offer pre-Kindergarten or Kindergarten as the highest grade. If several schools shared a principal we only kept one of the schools, chosen randomly. We dropped public charter schools from our sample. In Experiment 2, we included all public schools in our sampling framework regardless of their school type or data availability.

B4. USING NAMES TO SIGNAL RACE/ETHNICITY

Our primary outcome of interest is the response rate by racial/ethnic group, particularly the difference in responses between Chinese and White Americans. Thus, we refer throughout the text to evidence of “Chinese American discrimination” rather than “Asian American discrimination.” However, multiple mechanisms may drive our results. While we did not conduct a pre-test to assess perceptions of our names before we fielded our two experiments, we did conduct a post-test to help us better understand the potential mechanisms at play. In February 2022, we fielded a survey with a national (quota-based) sample of Americans via Lucid Marketplace. 262 respondents completed the survey and evaluated the names in our study. Respondents’ perceptions of our names varied somewhat by our intended signal of race/ethnicity: 65.3% guessed the intended race/ethnicity for our Black names, 73.3% for our

Hispanic names, 78.9% for our White names, and 70.6% for our Asian names. When we asked respondents what specific Asian origin our names signaled, the plurality (40.8%) said Chinese, with 18.4% stating Japanese and 14.0% stating Korean. Thus, while we intended to convey a Chinese American identity to our participants, some principals may have responded as if they were interacting with a family from another Asian ethnicity. Another way of thinking about this is that our effects are more generally driven by a mix of anti-Chinese and anti-Asian attitudes. Future work should examine this in more detail. However, this issue is not unique to our study, as all audits using names to signal race/ethnicity encounter similar issues.

C. CALCULATING THE SAMPLING VARIANCE OF THE DISCRIMINATION RATIO

Beyond the discrimination ratios, we must calculate the sampling variance of the estimate of each discrimination ratio. Using dichotomous data and a risk ratio outcome measure, the equation to calculate the variance is:

$$\text{Var}(\ln(\text{DR}_{wm})) = \frac{1}{r_w} - \frac{1}{n_w} + \frac{1}{r_m} - \frac{1}{n_m}$$

Where DR_{wm} is the discrimination ratio Whites (w) and racial/ethnic group m , r_w is the number of positive responses for White parents, n_w is the number of total emails sent for White parents, r_m is the number of positive responses for racial/ethnic group m parents, and n_m is the number of total emails sent for racial/ethnic group m parents.

D. REGRESSION AND SUPPLEMENTAL ANALYSES

Appendix Table D1. OLS Regressions Predicting Principal Response – Experiment 1

	(1) R/E Treatments	(2) All Treatments	(3) All Treatments + Block FEs	(4) All Treatments + State FEs
Black	-0.010	-0.010	-0.010	-0.010
Chinese American	-0.102***	-0.107***	-0.107***	-0.107***
Hispanic	-0.032***	-0.037***	-0.038***	-0.037***

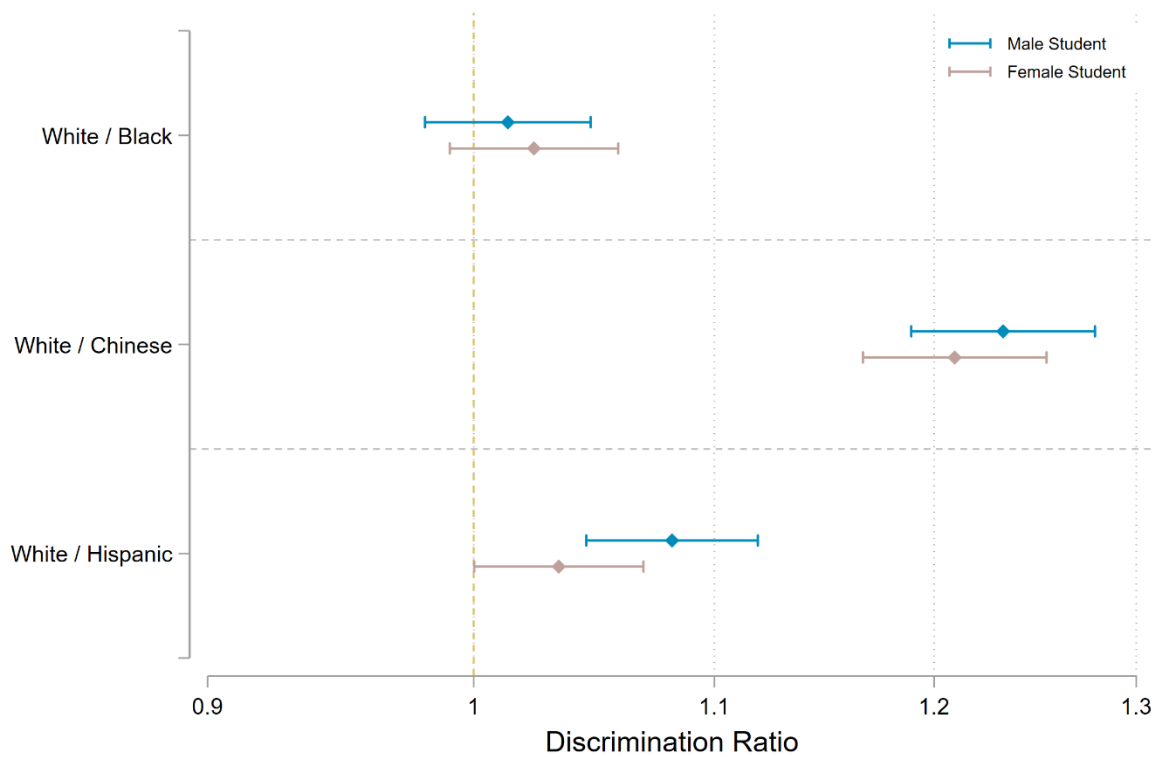
Note: + = $p < 0.10$, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

Appendix Table D2. OLS Regressions Predicting Principal Response – Experiment 2

	(1) R/E Treatments	(2) All Treatments + State FEs
Black	-0.014	-0.013
Chinese American	-0.113***	-0.113***
Hispanic	-0.014	-0.014

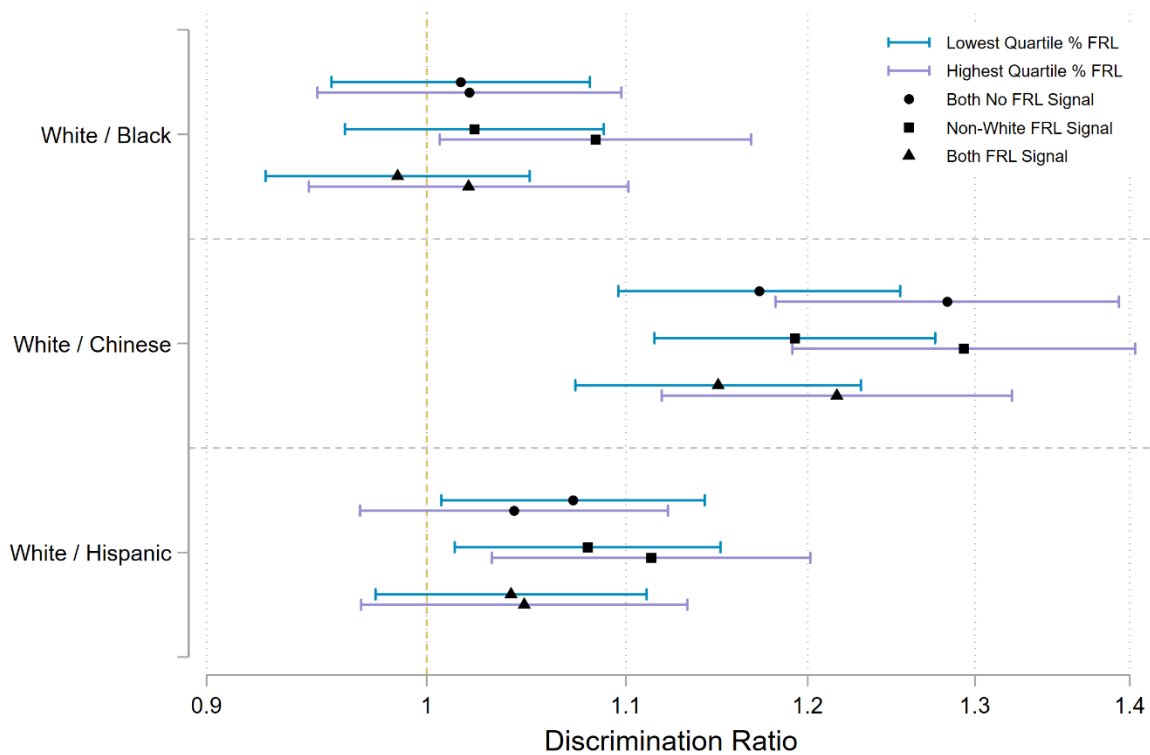
Note: + = $p < 0.10$, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

Appendix Figure D1. Discrimination Ratios by Student Gender



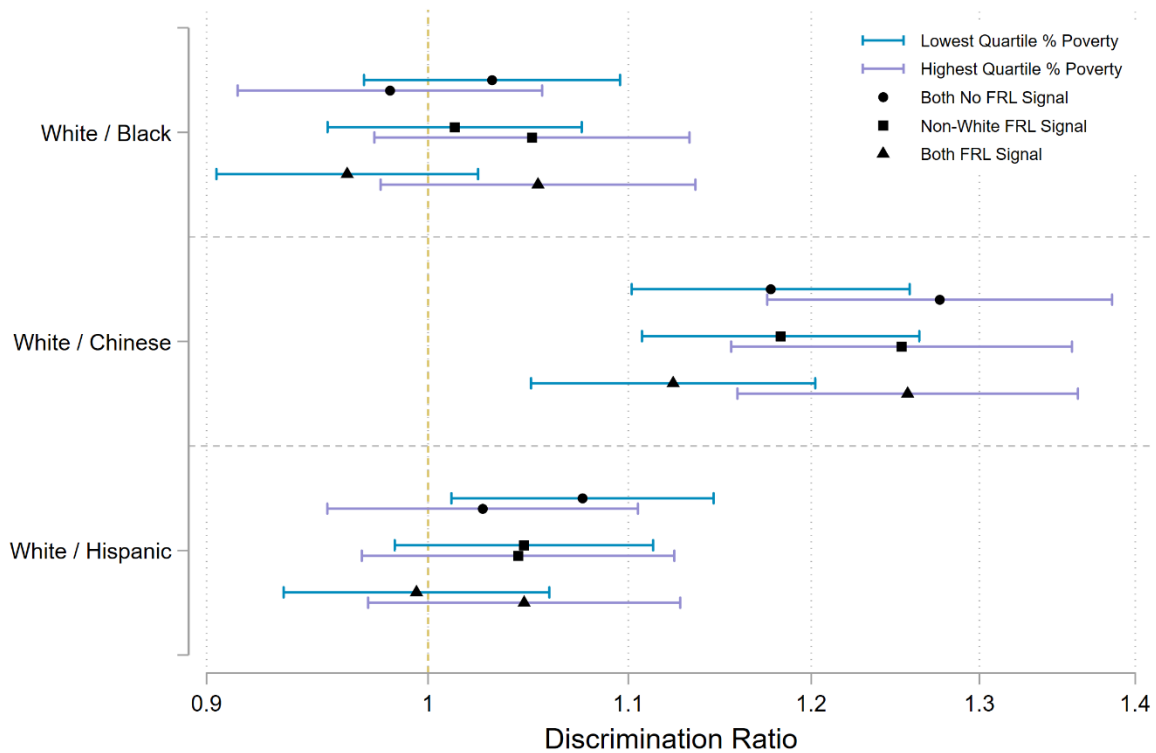
Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination.

Appendix Figure D2. Discrimination Ratios by the Percentage of FRL Students in the School and Student FRL Status



Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination. Percentage FRL students variable: Quartile 1 = $0 \leq 31.7\%$; Quartile 4 = $74.7\% \leq 100$.

Appendix Figure D3. Discrimination Ratios by the Percentage of Residents Below the Poverty Line and Student FRL Status



Note: Points are predicted discrimination ratios, and horizontal lines are the 95% confidence intervals. N=52,792 principals in 33 states. The underlying models are OLS regression models that include all treatment covariates and state fixed effects. Results shown on a natural log scale. Results that cross the gold line at 1 indicate no statistically significant evidence of discrimination. Percentage residents below poverty line variable: Quartile 1 = $0 \leq 8.7\%$; Quartile 2 = $8.7\% \leq 14.2\%$; Quartile 3 = $14.2\% \leq 19.7\%$; Quartile 4 = $19.7\% \leq 100$.

REFERENCES

- Atkin, Annabelle L., Hyung Chol Yoo, and Christine J. Yeh. 2019. "What Types of Racial Messages Protect Asian American Adolescents from Discrimination? A Latent Interaction Model." *Journal of Counseling Psychology* 66(2):247–54. doi: [10.1037/cou0000297](https://doi.org/10.1037/cou0000297).
- Baker, Bruce D., and Bruce S. Cooper. 2005. "Do Principals With Stronger Academic Backgrounds Hire Better Teachers? Policy Implications for Improving High-Poverty Schools." *Educational Administration Quarterly* 41(3):449–79. doi: [10.1177/0013161X04269609](https://doi.org/10.1177/0013161X04269609).
- Bartanen, Brendan, and Jason A. Grissom. 2023. "School Principal Race, Teacher Racial Diversity, and Student Achievement." *Journal of Human Resources* 58(2):666–712. doi: [10.3368/jhr.58.4.0218-9328R2](https://doi.org/10.3368/jhr.58.4.0218-9328R2).
- Benner, Aprile D., Yijie Wang, Yishan Shen, Alaina E. Boyle, Richelle Polk, and Yen-Pi Cheng. 2018. "Racial/Ethnic Discrimination and Well-Being during Adolescence: A Meta-Analytic Review." *American Psychologist* 73(7):855–83. doi: [10.1037/amp0000204](https://doi.org/10.1037/amp0000204).
- Block Jr., Ray, Charles Crabtree, John B. Holbein, and J. Quin Monson. 2021. "Are Americans Less Likely to Reply to Emails from Black People Relative to White People?" *Proceedings of the National Academy of Sciences* 118(52).
- Booher-Jennings, Jennifer. 2005. "Below the Bubble: 'Educational Triage' and the Texas Accountability System." *American Educational Research Journal* 42(2):231–68. doi: [10.3102/00028312042002231](https://doi.org/10.3102/00028312042002231).
- Boser, Ulrich, Megan Wilhelm, and Robert Hanna. 2014. *The Power of the Pygmalion Effect: Teachers' Expectations Strongly Predict College Completion*. Center for American Progress.
- Brown, Kathleen. 2005. "Pivotal Points: History, Development, and Promise of the Principalsip." *Sage Handbook of Educational Leadership* 109–41.
- Carnoy, Martin, Richard Elmore, and Leslie Siskin. 2003. *The New Accountability: High Schools and High-Stakes Testing*. Routledge.
- Chiu, Ming Ming, and Lawrence Khoo. 2005. "Effects of Resources, Inequality, and Privilege Bias on Achievement: Country, School, and Student Level Analyses." *American Educational Research Journal* 42(4):575–603. doi: [10.3102/00028312042004575](https://doi.org/10.3102/00028312042004575).
- Condrón, Dennis J., and Vincent J. Roscigno. 2003. "Disparities within: Unequal Spending and Achievement in an Urban School District." *Sociology of Education* 76(1):18–36. doi: [10.2307/3090259](https://doi.org/10.2307/3090259).
- Cooper, Camille Wilson. 2009. "Performing Cultural Work in Demographically Changing Schools: Implications for Expanding Transformative Leadership Frameworks." *Educational Administration Quarterly* 45(5):694–724. doi: [10.1177/0013161X09341639](https://doi.org/10.1177/0013161X09341639).
- Coppock, Alexander. 2019. "Avoiding Post-Treatment Bias in Audit Experiments." *Journal of Experimental Political Science* 6(1):1–4. doi: [10.1017/XPS.2018.9](https://doi.org/10.1017/XPS.2018.9).
- Coppock, Alexander, Thomas J. Leeper, and Kevin J. Mullinix. 2018. "Generalizability of Heterogeneous Treatment Effect Estimates across Samples." *Proceedings of the National Academy of Sciences* 115(49):12441–46. doi: [10.1073/pnas.1808083115](https://doi.org/10.1073/pnas.1808083115).
- Crabtree, Charles. 2018. "An Introduction to Conducting Email Audit Studies." Pp. 103–17 in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance, Methodos Series*, edited by S. M. Gaddis. Cham: Springer International Publishing.

- Crabtree, Charles, S. Michael Gaddis, John Holbein, and Edvard Larsen. 2021. "Changes in the Amount and Type of Social Science Research on Discrimination: 1950 to 2015."
- Crenshaw, Kimberlé Williams. 1988. "Race, Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law." *Harvard Law Review* 191(7):1331–87.
- Daga, Suchi S., and Vaishali V. Raval. 2018. "Ethnic–Racial Socialization, Model Minority Experience, and Psychological Functioning among South Asian American Emerging Adults: A Preliminary Mixed-Methods Study." *Asian American Journal of Psychology* 9(1):17–31. doi: [10.1037/aap0000108](https://doi.org/10.1037/aap0000108).
- Diamond, John B., Antonia Randolph, and James P. Spillane. 2004. "Teachers' Expectations and Sense of Responsibility for Student Learning: The Importance of Race, Class, and Organizational Habitus." *Anthropology & Education Quarterly* 35(1):75–98. doi: [10.1525/aeq.2004.35.1.75](https://doi.org/10.1525/aeq.2004.35.1.75).
- Dotterer, Aryn M., Susan M. McHale, and Ann C. Crouter. 2009. "Sociocultural Factors and School Engagement Among African American Youth: The Roles of Racial Discrimination, Racial Socialization, and Ethnic Identity." *Applied Developmental Science* 13(2):61–73. doi: [10.1080/10888690902801442](https://doi.org/10.1080/10888690902801442).
- Essed, Philomena. 1991. *Understanding Everyday Racism: An Interdisciplinary Theory*. SAGE.
- Evans, Andrea E. 2007. "School Leaders and Their Sensemaking About Race and Demographic Change." *Educational Administration Quarterly* 43(2):159–88. doi: [10.1177/0013161X06294575](https://doi.org/10.1177/0013161X06294575).
- Farkas, George. 2003. "Racial Disparities and Discrimination in Education: What Do We Know, How Do We Know It, and What Do We Need to Know?" *Teachers College Record* 105(6):1119–46. doi: [10.1111/1467-9620.00279](https://doi.org/10.1111/1467-9620.00279).
- Finnigan, Kara S., and Tricia J. Stewart. 2009. "Leading Change under Pressure: An Examination of Principal Leadership in Low-Performing Schools." *Journal of School Leadership* 19(5):586–621. doi: [10.1177/105268460901900504](https://doi.org/10.1177/105268460901900504).
- Francis, Dania V., Angela C. M. de Oliveira, and Carey Dimmitt. 2019. "Do School Counselors Exhibit Bias in Recommending Students for Advanced Coursework?" *The B.E. Journal of Economic Analysis & Policy* 19(4). doi: [10.1515/bejeap-2018-0189](https://doi.org/10.1515/bejeap-2018-0189).
- Freeman, Alan. 1989. "Antidiscrimination Law: The View From 1989 Civil Rights Symposium." *Tulane Law Review* 64(6):1407–42.
- Freeman, Alan David. 1977. "Legitimizing Racial Discrimination through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine." *Minnesota Law Review* 62(6):1049–1120.
- Gaddis, S. 2017. "How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies." *Sociological Science* 4:469–89. doi: [10.15195/v4.a19](https://doi.org/10.15195/v4.a19).
- Gaddis, S. Michael. 2017. "Racial/Ethnic Perceptions from Hispanic Names: Selecting Names to Test for Discrimination." *Socius* 3:2378023117737193. doi: [10.1177/2378023117737193](https://doi.org/10.1177/2378023117737193).
- Gaddis, S. Michael. 2018a. "An Introduction to Audit Studies in the Social Sciences." Pp. 3–44 in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance, Methodos Series*, edited by S. M. Gaddis. Cham: Springer International Publishing.
- Gaddis, S. Michael. 2018b. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Springer.
- Gaddis, S. Michael. 2019. "Understanding the 'How' and 'Why' Aspects of Racial-Ethnic Discrimination: A Multimethod Approach to Audit Studies." *Sociology of Race and Ethnicity* 5(4):443–55.

- Gaddis, S. Michael. 2022. "Cognitive-Behavioral Disconnect Bias in the Study of Racial/Ethnic Discrimination: A Method to Examine and Measure Respondents' Alignment of Predicted and Actual Behavior."
- Gaddis, S. Michael. 2024. *Understanding Racial/Ethnic Discrimination in the 21st Century*.
- Gaddis, S. Michael, A. Nicole Kreisberg, and Charles Crabtree. 2022. "Assessing Immigrant Generation and Citizenship Status from Names."
- Gaddis, S. Michael, Edvard Larsen, Charles Crabtree, and John Holbein. 2023. "Discrimination Against Black and Hispanic Americans Is Highest in Hiring and Housing Contexts: A Meta-Analysis of Correspondence Audits."
- Gelman, Andrew. 2018. "You Need 16 Times the Sample Size to Estimate an Interaction than to Estimate a Main Effect « Statistical Modeling, Causal Inference, and Social Science." Retrieved August 1, 2021 (<https://statmodeling.stat.columbia.edu/2018/03/15/need-16-times-sample-size-estimate-interaction-estimate-main-effect/>).
- Gelman, Andrew, and John Carlin. 2014. "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors." *Perspectives on Psychological Science* 9(6):641–51. doi: [10.1177/1745691614551642](https://doi.org/10.1177/1745691614551642).
- Gilliam, Walter S., Angela N. Maupin, Chin R. Reyes, Maria Accavitti, and Frederick Shic. 2016. "Do Early Educators' Implicit Biases Regarding Sex and Race Relate to Behavior Expectations and Recommendations of Preschool Expulsions and Suspensions?"
- Goldring, Rebecca, Lucinda Gray, and Amy Bitterman. 2013. *Characteristics of Public and Private Elementary and Secondary School Teachers in the United States: Results from the 2011-12 Schools and Staffing Survey. First Look. NCES 2013-314*. National Center for Education Statistics.
- Gomila, Robin. 2021. "Logistic or Linear? Estimating Causal Effects of Experimental Treatments on Binary Outcomes Using Regression Analysis." *Journal of Experimental Psychology: General* 150(4):700–709. doi: [10.1037/xge0000920](https://doi.org/10.1037/xge0000920).
- Griffin, Kimberly, and Walter Allen. 2006. "Mo' Money, Mo' Problems? High-Achieving Black High School Students' Experiences with Resources, Racial Climate, and Resilience." *The Journal of Negro Education* 75(3):478–94.
- Grissom, Jason A. 2011. "Can Good Principals Keep Teachers in Disadvantaged Schools? Linking Principal Effectiveness to Teacher Satisfaction and Turnover in Hard-to-Staff Environments." *Teachers College Record* 113(11):2552–85. doi: [10.1177/016146811111301102](https://doi.org/10.1177/016146811111301102).
- Grissom, Jason A., Anna J. Egalite, and Constance A. Lindsay. 2021. *How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research*. New York, NY: The Wallace Foundation.
- Gullo, Gina Laura, and Floyd D. Beachum. 2020. "Does Implicit Bias Matter at the Administrative Level? A Study of Principal Implicit Bias and the Racial Discipline Severity Gap." *Teachers College Record* 122(3):1–28. doi: [10.1177/016146812012200309](https://doi.org/10.1177/016146812012200309).
- Gutman, Leslie Morrison, and Vonnice C. McLoyd. 2000. "Parents' Management of Their Children's Education Within the Home, at School, and in the Community: An Examination of African-American Families Living in Poverty." *The Urban Review* 32(1):1–24. doi: [10.1023/A:1005112300726](https://doi.org/10.1023/A:1005112300726).
- Harris, Deborah Carol Fields. 2022. "Unconscious Bias Towards African American Women in the Education Workforce." Pp. 23–43 in *Addressing Issues of Systemic Racism During Turbulent Times*. IGI Global.

- Hellevik, Ottar. 2009. "Linear versus Logistic Regression When the Dependent Variable Is a Dichotomy." *Quality & Quantity* 43(1):59–74. doi: [10.1007/s11135-007-9077-3](https://doi.org/10.1007/s11135-007-9077-3).
- Holbein, John B., and Helen F. Ladd. 2017. "Accountability Pressure: Regression Discontinuity Estimates of How No Child Left Behind Influenced Student Behavior." *Economics of Education Review* 58:55–67. doi: [10.1016/j.econedurev.2017.03.005](https://doi.org/10.1016/j.econedurev.2017.03.005).
- Hughes, Diane, James Rodriguez, Emilie P. Smith, Deborah J. Johnson, Howard C. Stevenson, and Paul Spicer. 2006. "Parents' Ethnic-Racial Socialization Practices: A Review of Research and Directions for Future Study." *Developmental Psychology* 42(5):747–70. doi: [10.1037/0012-1649.42.5.747](https://doi.org/10.1037/0012-1649.42.5.747).
- Jacob, Brian A. 2005. "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." *Journal of Public Economics* 89(5):761–96. doi: [10.1016/j.jpubeco.2004.08.004](https://doi.org/10.1016/j.jpubeco.2004.08.004).
- Jarvis, Shoshana N., and Jason A. Okonofua. 2020. "School Deferred: When Bias Affects School Leaders." *Social Psychological and Personality Science* 11(4):492–98. doi: [10.1177/1948550619875150](https://doi.org/10.1177/1948550619875150).
- Kluger, Richard. 2011. *Simple Justice: The History of Brown v. Board of Education and Black America's Struggle for Equality*. Knopf Doubleday Publishing Group.
- Ladd, Helen F., and Arnaldo Zelli. 2002. "School-Based Accountability in North Carolina: The Responses of School Principals." *Educational Administration Quarterly* 38(4):494–529. doi: [10.1177/001316102237670](https://doi.org/10.1177/001316102237670).
- Larsen, Edvard N. 2020. "Induced Competition in Matched Correspondence Tests: Conceptual and Methodological Considerations." *Research in Social Stratification and Mobility* 65:100475. doi: [10.1016/j.rssm.2020.100475](https://doi.org/10.1016/j.rssm.2020.100475).
- Lauen, Douglas Lee, and S. Michael Gaddis. 2012. "Shining a Light or Fumbling in the Dark? The Effects of NCLB's Subgroup-Specific Accountability on Student Achievement." *Educational Evaluation and Policy Analysis* 34(2):185–208. doi: [10.3102/0162373711429989](https://doi.org/10.3102/0162373711429989).
- Lauen, Douglas Lee, and S. Michael Gaddis. 2016. "Accountability Pressure, Academic Standards, and Educational Triage." *Educational Evaluation and Policy Analysis* 38(1):127–47. doi: [10.3102/0162373715598577](https://doi.org/10.3102/0162373715598577).
- Lee, Jung-Sook, and Natasha K. Bowen. 2006. "Parent Involvement, Cultural Capital, and the Achievement Gap Among Elementary School Children." *American Educational Research Journal* 43(2):193–218. doi: [10.3102/00028312043002193](https://doi.org/10.3102/00028312043002193).
- Lewis-McCoy, R. L'Heureux. 2014. "Inequality in the Promised Land: Race, Resources, and Suburban Schooling." in *Inequality in the Promised Land*. Stanford University Press.
- Liu, Lisa L., and Anna S. Lau. 2013. "Teaching about Race/Ethnicity and Racism Matters: An Examination of How Perceived Ethnic Racial Socialization Processes Are Associated with Depression Symptoms." *Cultural Diversity and Ethnic Minority Psychology* 19(4):383–94. doi: [10.1037/a0033447](https://doi.org/10.1037/a0033447).
- Lu, Yao, Neeraj Kaushal, Xiaoning Huang, and S. Michael Gaddis. 2021. "Priming COVID-19 Salience Increases Prejudice and Discriminatory Intent against Asians and Hispanics." *Proceedings of the National Academy of Sciences* 118(36):e2105125118. doi: [10.1073/pnas.2105125118](https://doi.org/10.1073/pnas.2105125118).
- Malouff, John M., and Einar B. Thorsteinsson. 2016. "Bias in Grading: A Meta-Analysis of Experimental Research Findings." *Australian Journal of Education* 60(3):245–56. doi: [10.1177/0004944116664618](https://doi.org/10.1177/0004944116664618).

- Mattison, Erica, and Mark S. Aber. 2007. "Closing the Achievement Gap: The Association of Racial Climate with Achievement and Behavioral Outcomes." *American Journal of Community Psychology* 40(1):1–12. doi: [10.1007/s10464-007-9128-x](https://doi.org/10.1007/s10464-007-9128-x).
- Mickelson, Roslyn Arlin. 2003. "When Are Racial Disparities in Education the Result of Racial Discrimination? A Social Science Perspective." *Teachers College Record* 105(6):1052–86. doi: [10.1111/1467-9620.00277](https://doi.org/10.1111/1467-9620.00277).
- Mitani, Hajime. 2018. "Principals' Working Conditions, Job Stress, and Turnover Behaviors Under NCLB Accountability Pressure." *Educational Administration Quarterly* 54(5):822–62. doi: [10.1177/0013161X18785874](https://doi.org/10.1177/0013161X18785874).
- Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." *American Journal of Political Science* 62(3):760–75. doi: [10.1111/ajps.12357](https://doi.org/10.1111/ajps.12357).
- National Research Council of the National Academies. 2004. *Measuring Racial Discrimination*. National Academies Press.
- Oberfield, Zachary W., and Matthew B. Incantalupo. 2021. "Racial Discrimination and Street-Level Managers: Performance, Publicness, and Group Bias." *Public Administration Review* 81(6):1055–70. doi: [10.1111/puar.13376](https://doi.org/10.1111/puar.13376).
- O'Connor, Carla, Amanda Lewis, and Jennifer Mueller. 2007. "Researching 'Black' Educational Experiences and Outcomes: Theoretical and Methodological Considerations." *Educational Researcher* 36(9):541–52. doi: [10.3102/0013189X07312661](https://doi.org/10.3102/0013189X07312661).
- Okonofua, Jason A., and Jennifer L. Eberhardt. 2015. "Two Strikes: Race and the Disciplining of Young Students." *Psychological Science* 26(5):617–24. doi: [10.1177/0956797615570365](https://doi.org/10.1177/0956797615570365).
- Owens, Jayanti. 2022. "Double Jeopardy: Teacher Biases, Racialized Organizations, and the Production of Racial/Ethnic Disparities in School Discipline." *American Sociological Review* 87(6):1007–48. doi: [10.1177/00031224221135810](https://doi.org/10.1177/00031224221135810).
- Pager, Devah, and Hana Shepherd. 2008. "The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets." *Annual Review of Sociology* 34(1):181–209. doi: [10.1146/annurev.soc.33.040406.131740](https://doi.org/10.1146/annurev.soc.33.040406.131740).
- Pfaff, Steven, Charles Crabtree, Holger L. Kern, and John B. Holbein. 2021. "Do Street-Level Bureaucrats Discriminate Based on Religion? A Large-Scale Correspondence Experiment among American Public School Principals." *Public Administration Review* 81(2):244–59. doi: [10.1111/puar.13235](https://doi.org/10.1111/puar.13235).
- Priest, Naomi, Yin Paradies, Brigid Trenerry, Mandy Truong, Saffron Karlsen, and Yvonne Kelly. 2013. "A Systematic Review of Studies Examining the Relationship between Reported Racism and Health and Wellbeing for Children and Young People." *Social Science & Medicine* 95:115–27. doi: [10.1016/j.socscimed.2012.11.031](https://doi.org/10.1016/j.socscimed.2012.11.031).
- Quillian, Lincoln, Devah Pager, Ole Hexel, and Arnfinn H. Midtbøen. 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):10870–75. doi: [10.1073/pnas.1706255114](https://doi.org/10.1073/pnas.1706255114).
- Quinn, David M. 2020. "Experimental Evidence on Teachers' Racial Bias in Student Evaluation: The Role of Grading Scales." *Educational Evaluation and Policy Analysis* 42(3):375–92. doi: [10.3102/0162373720932188](https://doi.org/10.3102/0162373720932188).
- Rivera, Lauren A., and András Tilcsik. 2023. "Not in My Schoolyard: Disability Discrimination in Educational Access." *American Sociological Review* 88(2):284–321. doi: [10.1177/00031224221150433](https://doi.org/10.1177/00031224221150433).

- Smalls, Ciara, Rhonda White, Tabbye Chavous, and Robert Sellers. 2007. "Racial Ideological Beliefs and Racial Discrimination Experiences as Predictors of Academic Engagement Among African American Adolescents." *Journal of Black Psychology* 33(3):299–330. doi: [10.1177/0095798407302541](https://doi.org/10.1177/0095798407302541).
- Sue, Derald Wing, Christina M. Capodilupo, Gina C. Torino, Jennifer M. Bucceri, Aisha M. B. Holder, Kevin L. Nadal, and Marta Esquilin. 2007. "Racial Microaggressions in Everyday Life: Implications for Clinical Practice." *American Psychologist* 62(4):271–86. doi: [10.1037/0003-066X.62.4.271](https://doi.org/10.1037/0003-066X.62.4.271).
- Thapa, Amrit, Jonathan Cohen, Shawn Guffey, and Ann Higgins-D'Alessandro. 2013. "A Review of School Climate Research." *Review of Educational Research* 83(3):357–85. doi: [10.3102/0034654313483907](https://doi.org/10.3102/0034654313483907).
- Umaña-Taylor, Adriana J., and Nancy E. Hill. 2020. "Ethnic–Racial Socialization in the Family: A Decade's Advance on Precursors and Outcomes." *Journal of Marriage and Family* 82(1):244–71. doi: [10.1111/jomf.12622](https://doi.org/10.1111/jomf.12622).
- U.S. Census Bureau. 2012. "Genealogy Data: Frequently Occurring Surnames from Census 2000, Census Report Data File A: Top 1000 Names." Retrieved September 15, 2014 (https://www.census.gov/topics/population/genealogy/data/2000_surnames.html).
- Voight, Adam, Thomas Hanson, Meagan O'Malley, and Latifah Adekanye. 2015. "The Racial School Climate Gap: Within-School Disparities in Students' Experiences of Safety, Support, and Connectedness." *American Journal of Community Psychology* 56(3):252–67. doi: [10.1007/s10464-015-9751-x](https://doi.org/10.1007/s10464-015-9751-x).
- Vuolo, Mike, Christopher Uggen, and Sarah Lageson. 2018. "To Match or Not to Match? Statistical and Substantive Considerations in Audit Design and Analysis." Pp. 119–40 in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance, Methodos Series*, edited by S. M. Gaddis. Cham: Springer International Publishing.
- Wang, Fei, Cameron Hauseman, and Katina Pollock. 2022. "'I Am Here for the Students': Principals' Perception of Accountability amid Work Intensification." *Educational Assessment, Evaluation and Accountability* 34(1):33–56. doi: [10.1007/s11092-021-09368-6](https://doi.org/10.1007/s11092-021-09368-6).
- Wang, Jennifer, Janxin Leu, and Yuichi Shoda. 2011. "When the Seemingly Innocuous 'Stings': Racial Microaggressions and Their Emotional Consequences." *Personality and Social Psychology Bulletin* 37(12):1666–78. doi: [10.1177/0146167211416130](https://doi.org/10.1177/0146167211416130).
- Wang, Ming-Te, Daphne A. Henry, Leann V. Smith, James P. Huguley, and Jiesi Guo. 2020. "Parental Ethnic-Racial Socialization Practices and Children of Color's Psychosocial and Behavioral Adjustment: A Systematic Review and Meta-Analysis." *American Psychologist* 75(1):1–22. doi: [10.1037/amp0000464](https://doi.org/10.1037/amp0000464).
- West, Deborah L., Craig M. Peck, Ulrich C. Reitzug, and Elizabeth A. Crane. 2014. "Accountability, Autonomy and Stress: Principal Responses to Superintendent Change in a Large US Urban School District." *School Leadership & Management* 34(4):372–91. doi: [10.1080/13632434.2014.928683](https://doi.org/10.1080/13632434.2014.928683).
- West, Deborah L., Craig Peck, and Ulrich C. Reitzug. 2010. "Limited Control and Relentless Accountability: Examining Historical Changes in Urban School Principal Pressure." *Journal of School Leadership* 20(2):238–66. doi: [10.1177/105268461002000206](https://doi.org/10.1177/105268461002000206).
- Woo, Bongki, Dale Dagar Maglalang, Samuel Ko, Michael Park, Yoonsun Choi, and David T. Takeuchi. 2020. "Racial Discrimination, Ethnic-Racial Socialization, and Cultural Identities among Asian American Youths." *Cultural Diversity and Ethnic Minority Psychology* 26(4):447–59. doi: [10.1037/cdp0000327](https://doi.org/10.1037/cdp0000327).

Young, I. Phillip, and Julie A. Fox. 2002. "Asian, Hispanic, and Native American Job Candidates: Prescreened or Screened within the Selection Process." *Educational Administration Quarterly* 38(4):530–54. doi: [10.1177/001316102237671](https://doi.org/10.1177/001316102237671).