OXFORD

Article

Reducing Minority Discrimination at the Front Line—Combined Survey and Field Experimental Evidence

Simon Calmar Andersen, Thorbjørn Sejr Guul

Aarhus University

Address correspondence to the author at sca@ps.au.dk.

Abstract

Despite laws of universalistic treatment, bureaucrats have been shown to discriminate against minorities. A crucial question for public administration is how bureaucracies can be organized in ways that minimize illegitimate discrimination. Especially, since theories suggest that prejudices happen unintentionally and particularly under high workload, bureaucrats' working conditions may be important. Four randomized experiments support the notion that bureaucrats discriminate as a way of coping with high workload. Most notably, a field experiment randomly assigned teachers to reduced workloads by giving them resources to have more time with the same group of students. In a subsequent survey experiment—using a fictitious future scenario unrelated to the resources provided in the field experiment—discrimination was minimized in the field treatment group, but persisted in the control group. The results thereby support the notion that even though discrimination among bureaucrats does not (only) occur in a reflective manner it can be reduced by altering the way bureaucrats' work is organized.

Introduction

The promise of modern bureaucracy is "to eradicate prejudicial behavior through universalistic treatment" (Lipsky 1980, 109; see also Weber 1922). However, a number of studies show that bureaucrats tend to use coping strategies that bend or break the rules in different ways and discriminate against groups of clients (e.g., Baviskar and Winter 2017; Keiser, Mueser,

and Choi 2004; Olson 2016; see Tummers et al. 2015 for a recent review; Tummers 2016). A large body of research in psychology on stereotyping demonstrates different general psychological mechanisms that may lead to discrimination (e.g., Fiske 1998; Hardin and Banaji 2013). However, the task for public administration as a design science (cf. Simon 1969) is to sort out how to organize bureaucracies in ways that minimize illegitimate discrimination—even if it is based on irrational or unconscious psychological processes in the individual.

In his seminal work on street-level bureaucrats, Lipsky (1980) theorized that bureaucrats typically do not discriminate because they are biased or racist, but because their work—that involves heavy workloads due to ever-high demands for their services—creates conditions for de facto discrimination: "Client differentiation may take place because, confronted with heavy work loads and apparently impossible tasks,

street-level bureaucrats seek ways to maximize personal or agency resources, or they attempt to succeed with some clients when they cannot succeed with all." (1980, 107). As recent studies note, the workload of public sector employees are increasing in many countries, particularly due to cutbacks after the financial crisis (Vaughan-Whitehead 2013). The possible relationship between workload and discrimination therefore becomes all the more important. Unfortunately, little if any research examines how interventions in street-level bureaucrats' working conditions affect discrimination (Spencer, Charbonneau, and Glaser 2016). Indeed, field experimental evidence of prejudice reduction in general is very limited (Paluck and Green 2009).

The primary aim of the research studies presented here is to examine how workload relates to discrimination. We use theories from psychology and economics to substantiate the hypothesis that workload affects discrimination. Besides the possibility that discrimination is a consequence of sheer taste-based preferences (Becker 1957)—a view that does not find many supporters in the field—discrimination is often argued to build on either statistical (correct or incorrect) beliefs about traits correlated with clients' group membership (Phelps 1972) or on implicit biases based on unconscious and automatic psychological processes that escape people's intentions (e.g., Devine 1989) and depend on contextual factors (Christensen, Szmer, and Stritch 2012; Wenger and Wilkins 2009). In the theory section, we argue that both the two latter views would support the notion that discrimination is related to bureaucrats' workload.

However, workload—defined as the relationship between tasks and resources—can be operationalized and changed in many different ways.[1] We use four randomized survey and field experiments on school teachers (one of Lipsky's (1980) recurrent examples of street-level bureaucrats) to test the relationship between workload and discrimination in different ways. In the first experiment, we show that discrimination based on putatively ethnic minority name cues (which has often been documented in the United States) is also present in Denmark. Second, we compare the first experiment to a situation in which the teacher's decision to include a new student does not directly affect the teachers' own tasks. The third experiment shows that when more resources follow the potentially demanding students, teachers are more willing to include such students in their classroom. Finally, and most notably, in a combined field and survey experiment, we show

that after a period of reduced workload (in the field-treatment group) teachers do not discriminate against minorities in a hypothetical future scenario, whereas this discrimination is present in the field-control group, whose workload was not reduced.

This study advances existing knowledge in several important ways. First and foremost, we show how the organization of street-level bureaucrats' work affects their discrimination against minorities. Second, even if it is not the main purpose of this study, we also believe the field experiments give new evidence to the research on the psychological mechanisms underlying discrimination. In the discussion section following the presentation of the empirical studies, we argue that the results support both the statistical and implicit bias theories (but not the racism theory).

Any instances of unintentional discrimination only emphasize the need for studying effective interventions to reduce such biases. To quote Lipsky again: "If public officials were simply biased or racist, and if their prejudices were regularly manifested in behavior, the problem of bias in bureaucracy would be more pernicious but easier to root out" (Lipsky 1980, 109). In the concluding section, we discuss the viability of the workload-reducing intervention studied here, but first, we present the theoretical and empirical background of the study. Then, we present the general design of the four experiments, which is followed by a sequential presentation of each of them.

Theory and Existing Evidence

Prior Studies of Bureaucrats' Discrimination

The principle of modern constitutional states should guard against differential treatment, but a recent review shows that a substantial literature predominantly based on qualitative studies supports that bureaucrats use coping strategies that bend or break the rules on how to treat clients (Tummers et al. 2015). Strategies for coping range from bending rules to meeting the needs of the clients, to rationing services to limit clients access, or to referring some clients to other organizations as a way of reducing the street-level bureaucrats' own workload (Lipsky 1980; Tummers et al. 2015). By way of example, it has been shown that nonwhites are sanctioned more than white welfare recipients within the same district (Keiser, Mueser, and Choi 2004). Similarly, prison staff disproportionately punish black males more than white males (Olson 2016). Besides its immediate consequences, such behavior is problematic, as studies of political learning have shown that negative and stigmatizing experiences in direct encounters with government institutions can have severe consequences for political efficacy and participation (Schneider and Ingram 1993; Soss 1999). Further studies have shown

---

1 Workload defined as the relationship between tasks and resources are sometimes referred to as "caseload." In order not to refer to students as "cases," we use the general term "workload" similarly to, for example, Tummers et al. (2015). We note that we do not measure the perceived or cognitive workload of the teachers.

that recent trends of New Public Management reforms have increased discriminative behavior (Soss, Fording, and Schram 2011). Likewise, the administrative process has been found to affect the level of gender and racial bias (Christensen, Szmer, and Stritch 2012; Wenger and Wilkins 2009). This indicates that bureaucratic discrimination of minority citizens is contingent on organizational factors.

A literature more specifically concerned with teachers' beliefs about and treatment of black students suggests the existence of a white teacher bias against black students (Downey and Pribish 2004; Ferguson 2003; Yarnell and Bohrnstedt 2017). In support of this view, a study concerning teachers' assignment of students to gifted service shows that black students are more likely to be assigned to gifted service if they have a black teacher compared with a white (Nicholson-Crotty et al. 2016). As an exception to this general finding, a recent conjoint experiment shows that teachers are more willing to help an African American or Latino student compared with a Caucasian student. This is true for both white and African American teachers (Jilke and Tummers 2018). One explanation for the opposite result of the latter study may be that teachers were asked explicitly to compare students of different ethnic backgrounds. While this method (conjoint experiment) is suitable to test effects of socially undesirable attitudes, it may be less suited to test for implicit, unintentional biases.

Recently, a number of experimental studies in the United States have shown that street-level bureaucrats discriminate against minorities simply based on name cues (Einstein and Glick 2017; Schram et al. 2009; White, Nathan, and Faller 2015).[2] Studies of discrimination by street-level bureaucrats based on name cues outside the United States show more mixed results. For instance, Jilke, Van Dooren, and Rys (2018) find no discrimination by public officials in elder care centers in Flanders. Similarly, Grohe, Adam, and Knill (2016) find no systematic discrimination based on ethnicity in German local government. Contrary to this, Hemker and Rink (2017) find differences in the quality of answers that German majority and minority citizens receive from German welfare offices (but no discrimination in response rates), and in a survey vignette experiment, Pedersen, Stritch, and Thuesen (2018) find that employment agency caseworkers in Denmark are

more likely to recommend sanctions for clients with putatively ethnic minority names than Danish names.

In sum, the vast majority of more general studies on the use of coping strategies are conducted in the United States (43 %) or in the United Kingdom (24 %) (Tummers et al. 2015), the literature on teacher beliefs also primarily rely on US studies, and the name cue studies conducted outside the United States find mixed results. It is therefore an open question whether same kind of discrimination based on name cues will be found outside the United States.

More importantly to public administration as a design science, however, is how to organize bureaucracies in ways that reduce such tendencies toward discrimination. In the words of Simon (1947, 240–1):

> The need for an administrative theory resides in the fact that there are practical limits to human rationality, and that these limits are not static, but depend upon the organizational environment in which the individuals' decisions take place. The task of administration is so to design this environment that the individual will approach as close as practicable to rationality (judged in terms of the organization's goals) in his decisions.

Building on this perspective, we will employ existing research on human psychology to substantiate our hypothesis about one way that organizations may reduce discrimination. We do not aim to test the underlying theories but argue that what we see as the two dominating theories of discrimination both support the same organizational solution: reducing workloads.

Theoretical Relationship between Workload and Discrimination

At least three explanations for discriminating behavior can be distinguished. First, some older studies suggested that discrimination against ethnic minorities is caused by sheer racism (Becker 1957). While this is a possibility, we do not find much empirical evidence for this explanation. In what has been called "modern racism," whites' political opinions about racial issues has been explained by racial animus (Butler and Broockman 2011; DeSante 2013). However, this is not necessarily caused by conscious, intentional racism. For instance, DeSante (2013) uses a survey experiment to manipulate the names of applicants for state assistance. He finds that a representative sample of US adults perceive black applicants of state assistance as less deserving than otherwise similar white applicants. When the survey participants are informed that the applicants are either hard working or lazy, white applicants gain more from being hard working and are punished less for being lazy. Furthermore, this differential treatment of blacks and whites is more pronounced for citizens

---

2 In a recent meta-analysis of primarily elected, but also nonelected public officials, Costa (2017) confirms that citizens with racial/ethnic minority names are significantly less likely to receive a response from the public officials. The same approach has been used to show that minorities are discriminated by employers responding to job applications (Bertrand and Mullainathan 2004), politicians responding to potential voters (Butler and Broockman 2011), and citizens evaluating the deservingness of welfare applicants (DeSante 2013).

with higher levels of racial resentment. Yet, this could be the result of implicit biases rather than sheer racism, and DeSante (2013) himself labels this as implicit racism.

Two alternative explanations seem to have more empirical support, and both, we will argue, support the notion that individual bureaucrats' workload is related to discrimination. One explanation suggests that bureaucrats react based on conscious beliefs about the relative workload implied by different groups of clients. These beliefs may be more or less statistically correct. This explanation sees discrimination as a conscious and rational cost-benefit analysis based on relatively accurate beliefs and on the fact that the cost of obtaining more information does not outweigh the benefit of a more precise evaluation of a client or a customer (Phelps 1972). A German study of ethnic callback discrimination finds a substantial difference in callback rates, but importantly, this difference disappears when they consider a subsample of applications including a letter of recommendation. They interpret this as evidence of statistical discrimination since the difference disappears when more specific information (other than the applicant's ethnicity) is available (Kaas and Manger 2011). As noted by Jilke, Van Dooren, and Rys (2018), statistical discrimination in public service is a form of cream skimming or cropping—that is, prioritizing the citizens who are easy to serve and avoiding the citizens least likely to succeed. From this perspective, we deduce that bureaucrats will evaluate new clients in terms of how it affects their own workload. Bureaucrats would be less willing to serve clients who, statistically, can be expected to require more work. However, in situations where serving the clients does not affect the individual bureaucrat but other bureaucrats or other organizations, they should be indifferent to different groups of clients.

Another type of explanation suggests that discrimination works on a subtler level where people unwittingly and unintentionally exert discrimination against minority groups (Hardin and Banaji 2013). Such unconscious discrimination can be a result of either heuristics or aversive racism. This may happen because of a basic human need to make categorizations to perceive and act promptly to phenomena in the world (Gigerenzer and Gaissmaier 2011). These categorizations build on prevailing cultural stereotypes in society that may be more or less statistically correct (depending on the environment) and the stereotypes are internalized by bureaucrats and thereby in a nonconscious way influence the decisions they make when they implement policies (Schram et al. 2009; Soss, Fording, and Schram 2008). Studies of labeling suggest that when citizens are categorized in accordance with such stereotypes, little further discretion is exercised (Gilboy

1991). Several studies in the United States have indeed shown the existence of such attitudes. For instance, Dovidio, Evans, and Tyler (1986) show in a psychological experiment that students connect stereotypically black adjectives (such as musical, sensitive, lazy, and imitative) with black people faster (about 0.1 s) than with white people and vice versa, which indicates a subconscious web of cognitive associations. In another psychological experiment, Devine (1989) shows that after priming students with words stereotypically associated with black people (such as nigger, poor, afro, jazz, slavery, blues, rhythm, Africa), students presented with a vignette evaluated a race-unspecified person as more hostile than students primed with words stereotypically associated with white people. Though the heuristic explanation and aversive racism explanation differ in their argumentation, they both speak of subtler implicit biases.

We argue that also theories of implicit bias would suggest that a higher workload is related to more discrimination—but for other reasons than the statistical discrimination perspective. The implicit use of cues and stereotypes seems to depend on the psychological state of the individual (e.g., Bodenhausen 1990; Fazio 1990). The individual's motivation and ability to process information in a specific situation, for instance, affects how much a person rely on stereotypical judgment (Bodenhausen 1990). A main finding is that when appropriately motivated and with sufficient cognitive resources, people can avoid the influence of stereotypes in their conscious evaluations of others (Fiske 1998). For instance, experimental work has shown that subjects under cognitive business are less likely to activate stereotypes but more likely to apply such stereotypes if they indeed are activated (Gilbert and Gregory 1991). Studies of ego-depletion (Muraven and Baumeister 2000) have found related effects. Subjects engaged in a cognitive depleting task were just as likely to activate stereotypes as subjects in the control group but were more likely to apply these stereotypes afterwards (see also Richeson and Shelton 2003; Govorun and Payne 2006). Despite recent failure to replicate the ego-depletion effect on a computer task (Hagger et al. 2016), these findings suggest that people apply stereotypes to a larger extent when they have to make decisions under cognitive load but also in situations with low cognitive load if they have previously conducted a depleting task.

Some correlational evidence supports the notion that public employees apply stereotypes when under higher workload or more stressed. For instance, an observational study shows that police recruits apply stereotypes more when they had slept less the night before (Ma et al. 2013). Similarly, a study in emergency departments measured the level of implicit racial biases

among the physicians before and after ending their shift. The physicians who reported that they had experienced an extremely busy or dangerously overcrowded shift reported higher levels of implicit racial biases after the shift than before. Similarly, even stronger differences were found for physicians who reported that they had cared for particularly many patients (>10) (Johnson et al. 2016). Relating these findings to the issue of street-level bureaucrats' discrimination would suggest that unintentional prejudices or heuristics have greater influence on bureaucrats' attitudes and decisions when they feel pressured or stressed by high levels of workload.

Lipsky (1980) argues that in the absence of price mechanisms, bureaucracies typically face a demand that exceeds available resources. Bureaucrats will therefore sometimes need to limit access to their services, and they often have to do this based on a limited set of information. Ethnic minorities that may be subject to racial animus may also statistically represent a larger workload for the bureaucrats. However, the same discriminating behavior may have several causes. As mentioned, the present studies are not designed to test these underlying mechanisms, but after presenting the results, we discuss how the experiments may nevertheless provide some insights on the underlying mechanisms.

In sum, evidence suggests that minorities are being discriminated in many different settings in the United States and in some settings outside the United States as well. There are reasons to believe that workload will affect discrimination—either because of explicit cost-benefit analyses based on statistical discrimination or because a higher "objective" workload will affect the cognitive load of bureaucrats, which again will make unintentional, implicit use of stereotypes more likely. Yet, while most research has focused on the psychological mechanisms, reviews of this literature show that there is very little rigorous interventional research on how to reduce discrimination outside of the laboratory (Spencer, Charbonneau, and Glaser 2016) and apparently no research on how organizing bureaucrats' workloads affect their levels of discrimination (for reviews, see Hardin and Banaji 2013; Paluck and Green 2009).

Design

To elicit bureaucrats' attitudes and decisions we use survey experiments that are less biased by social desirability than standard survey formats. We build on previous studies randomly presenting cases to respondents with minority- and majority-sounding names (Bertrand and Mullainathan 2004; Butler and Broockman 2011; DeSante 2013; Einstein and Glick 2017; Schram et al.

2009; White, Nathan, and Faller 2015). We present four studies all using school teachers in Denmark. Teachers are one of Lipsky's (1980, 3) recurrent examples of street-level bureaucrats defined as public service workers that interact directly with citizens and have substantial discretion in the execution of their work. In addition, teachers' possible discriminating behavior is important to study for several reasons. First, public schooling is one of the most prevalent types of public service delivery. This is true across different welfare regimes. Focusing on schools thus increases the generalizability. Second, schools employ one of the largest groups of public employees and consume a huge part of public spending (Meier, Wrinkle, and Polinard 1999). Finally, following the findings from political learning (Schneider and Ingram 1993; Soss 1999), differential treatment of minorities at this stage can have important negative consequences in the long run.

Our general approach is to present teachers with hypothetical cases of students with special needs or discrediting information and ask the teachers to evaluate these cases. The studies are conducted in a Scandinavian context where the racial categories black, white, and Latino do not apply adequately to societal stereotypes and the actual population at large. Instead, immigrants or descendants of immigrants are the minority group of interest. We have used the names "Ahmed" and "Yousef" that are widespread among immigrants from the Middle East to indicate immigrant students, and "Anders" and "Mathias" that are two of the most common Danish names to indicate ethnic, Danish students. Certain names relate very clearly to immigrants from specific regions. We expect this association to affect the street-level bureaucrats' judgments of students. We do not claim that differential treatment of "Ahmed" and "Anders" necessarily reflects racism. However, we do claim that Ahmed is being discriminated if he is treated worse than Anders only because of his name. Some names might also be associated with high or low socioeconomic status, so we cannot conclude that discrimination against ethnic minority names only implies discrimination based on ethnicity. However, to reduce the influence of socioeconomic profiles associated with the names we use, we explicitly describe the socioeconomic status of the student in all hypothetical cases. In study 1, we also carry out a separate test of whether socioeconomic status itself affects discrimination.

We present four studies using three different study samples. Table 1 gives an overview of the four studies and three samples. Studies 1, 2, and 3 use survey experiments and study 4 combines survey and field experimental evidence. We recruited teachers from Danish public schools across the country, and, in general, the teachers had primarily an ethnic, Danish, or

a western background. Across all four studies, only 15 teachers with an immigrant background participated and no more than eight people within a single study. In all four studies, the school principals or the administrative head of school service signed up the teachers. In studies 1, 2, and 3, they did that to participate in other nonrelated field experiments related to improving learning outcomes for the students. The survey experiments were embedded in the presurveys for these other studies and are therefore expected to be unaffected by the unrelated field experiments. Study 4 was a field experiment providing teachers with more resources to teach the students. As part of the experiment, we collected survey data among the teachers after the field intervention period. A survey experiment was embedded in this survey.[3]

Table 1 shows that teachers from a wide range of schools (102–230) and municipalities (18–41 out of 98) were recruited. The sample schools were situated in cities as well as areas that are more rural. About 24%–37% of the schools were situated in the four largest city municipalities (>120,000 inhabitants). The teachers all taught classes from the fourth to the sixth grade level. The average share of students with an immigrant (or immigrant descendant) background at the school level was between 11% and 23% across the three samples, but with roughly the same amount of variation with a standard deviation close to 0.2. In all three samples, the response rates were high (teachers were asked by their principals to participate because the surveys were used for larger studies). Attrition will be described in more detail in the presentation of each study below. Finally, it should be mentioned that some teachers were sampled several times. The share of teachers from sample 1 that are also included in sample 2 or 3 is very low (<3%). Excluding these teachers from the analysis does not change the results. Studies 2 and 3 have more overlap. However, including the treatment indicators from study 3 in study 4 does not change the results (supplementary appendix table A9). Each of the four studies is presented below.

---

3 We carefully considered the ethics in this procedure. School principals have the authority to decide that teachers should take part in projects that attempt to improve school performance. Teachers were informed that the surveys were part of the research that should assess the effect of these projects. School principals therefore asked their teachers to respond to the surveys, which probably contributed to the high response rates. However, if teachers did not want to respond to specific questions, neither school principals nor anyone else outside the research team were informed about which teachers that did not respond. In order not to bias the results, teachers were not informed about the specific purpose of each question in the survey. Their responses were held confidential so no principal or anyone else outside the research team would be informed about the response of the individual teachers. Furthermore, by design the survey experiment cannot tell the researchers anything about the individual teacher's discriminatory attitudes.

Studies 1 and 2—Replication of Minority Bias and the Relevance of Workload

The first study tests whether discrimination similar to that identified in previous US studies (e.g., Einstein and Glick 2017; Schram et al. 2009; White, Nathan, and Faller 2015) replicates in the Danish context. This test presents the teachers with the case of a problematic student and randomly assign immigrant- and nonimmigrant-sounding names to the student. In this experiment, we ask the teacher to what extent she believes it is wise to include the student in her own class—which would affect her workload. Even if teachers do not formally have the authority to allocate students between classrooms, their opinion is likely to affect the school principal's decision. A comparative study of school principals in Texas and Denmark show that teachers have more influence in Denmark (Meier et al. 2015).

We attached a discrediting marker to the student in the vignette (that the student had socio-emotional difficulties) to create a realistic scenario where teachers would be asked about their willingness to include the student in the classroom and to reduce the association between the name and any socioeconomic characteristics. The teachers were asked to evaluate the statement on a Likert scale from 1 to 5 (mean = 3.23: SD = 1.07). Table 2 presents the exact wording of the case.

Table 3 shows that the experimental groups are balanced in the outset, which indicates that the randomization was successful. The response rate on the survey item used in study 1 was 94.9% (teachers were asked by their principals to participate because the survey was used as premeasure for a larger study). Attrition was not significantly different between treatment and control condition.

Study 2 tests whether teachers also discriminate against students with minority names if they are asked about their more general attitude toward moving students with special needs into ordinary education—a situation that would not affect their own workload directly. The second study is embedded in the same survey as study 1, so that we are able to directly compare teachers' responses in the two situations (with workload implications in study 1 and without workload implications in study 2). Just as in study 1, the teachers are presented with the case of a problematic student and discrediting markers are attached. To separate this from study 1, we created two different, but still similar cases. The student in study 2 has the diagnosis ADHD instead of socioemotional difficulties. ADHD students may imply higher workload than students with socioemotional difficulties as the ADHD diagnosis is usually associated with externalizing behavior. However, this should only make the test more conservative as the least demanding marker is attached to the student that affects the teacher's workload directly.

Table 1. Overview of Sampled Teachers

Study characteristics	Sample 1 (Study 1/Study 2)	Sample 2 (Study 3)	Sample 3 (Study 4)
Number of teachers	890	258	196
Number of schools	230	123	102
Number of municipalities	18	41	37
Share of schools in four largest city municipalities	0.37	0.24	0.25
Teaching grade	6th grade	4th grade	4th/5th grade
Conducted	Autumn 2012	Autumn 2013	Spring 2014
Share of immigrant students	Mean = 0.11 Std. = 0.15	Mean = 0.23 Std. = 0.18	Mean = 0.22 Std. = 0.17
Response rate	94.9 % (study 1) 96.9 % (study 2)	76.0 %	82.1 %
Share from sample 1		0.02	0.03
Share from sample 2			0.36

Table 2. Study 1—Experimental Treatments

Imagine a different situation. Your principal tells you that a new student, [**Mathias/Yousef**], is going to start in one of the 6th grade classes at the school. [**Mathias's/Yousef's**] dad is an engineer and his mum is a nurse. [**Mathias/Yousef**] has socioemotional difficulties, and hence learning disabilities.
Your principal asks if [**Mathias/Yousef**] could be accommodated in your class. To what extent do you agree or disagree that this is wise?

Note: Bold in square brackets: Randomized cue.

Table 3. Study 1—Balance across Experimental Conditions

Variables	Control	Treatment
Female	0.65	0.65
Age	47.04	47.78
College degree	0.90	0.88
Log earnings[a]	12.30	12.27
No valid information	0.04	0.05

Note: No statistical significant differences at 5% level. *N* = 890.
[a]We decided to use log earning throughout the article to reduce potential problems with extreme observations. However, using the raw earnings do not change the results.

Study 2 was placed before study 1 in the survey so that any consistency bias (that responses reflect attempts to respond consistently rather than honestly to any specific question) would work against finding a difference between willingness to include students in the school and in the classroom. The exact wording of the case is presented in table 4.[4] The teachers were asked to evaluate the statement on a five-point Likert scale (mean = 2.30; SD = 1.11).

4 In a 2 × 2 design, half of the teachers were assigned to another unrelated cue. Since it is not relevant and does not interact with the minority cue, it is not presented here.

Table 4. Study 2—Experimental Treatments

[**Anders/Ahmed**] is in 6th grade. He lives with his dad and mom and elder sister. Both parents are working, the family is well-functioning, and the parents are extending their support to [**Anders/Ahmed**] in his school work. [**Anders/Ahmed**] is assigned to special needs support besides the ordinary education. [**Anders/Ahmed**] is diagnosed with ADHD, and his behavior is dominated by attention problems and hyperactivity. His diagnosis reflects that he easily loses concentration and often interrupts the teaching. Apart from that, he is a boy with a normal intelligence who does well in the social relationships with his schoolmates.
Imagine that your principal in general wants to move more students such as [**Anders/Ahmed**] from special needs education into ordinary education. To what extent do you agree or disagree that this is wise?

Note: Bold in square brackets: Randomized cue.

Table 5 shows that the experimental groups are well-balanced in the outset, which indicates that the randomization was successful. The response rate on the survey item was 96.9%, and attrition was not significantly different between treatment and control condition.

The results from studies 1 and 2 are presented in table 6. Model 1 in the table shows that the minority name cue affects teachers' willingness to include the student in their own classroom negatively and significantly ($p < .002$). Including covariates does not change the results, and using ordered logistic regression produces the same results (supplementary appendix table A1). This result confirms previous studies from the United States showing that minorities tend to be treated worse. Therefore, students with immigrant names are discriminated against in similar ways as blacks or Latinos. The effect size is roughly equal to a quarter of a scale point on the Likert scale and also equal to a 0.22 standard deviation increase.

Table 5. Study 2—Balance across Experimental Conditions

Variables	Control	Treatment
Female	0.65	0.66
Age	47.37	47.44
College degree	0.88	0.90
Log earnings	12.38	12.19
No valid information	0.04	0.05

Note: No statistical significant differences at 5% level. *N* = 890.

To examine whether discrimination relates specifically to immigrant names, study 1 was conducted as a 2 × 2 design. Half of the teachers were randomly assigned to a case where the parents were described as having low socioeconomic status. The socioeconomic status marker itself had no significant effect, and we found no interaction effect between immigrant name and socioeconomic status marker (supplementary appendix table A10). This supports the notion that teachers react to the ethnicity implied by the names rather than any socioeconomic status associated with the names.

Model 3 in the table shows that the effect of the minority cue in study 2 concerning including the student in the school rather than the classroom is positive, very close to zero, about one-tenth of the effect size in study 1, and insignificant. Including covariates does not change the coefficients substantially. Using ordered logistic regression produces the same result (supplementary appendix table A2).[5] This suggests that when including students does not have direct implications for teachers' own workload, they do not react to the name cues.

To estimate whether the effects in study 1 (include student in own classroom) and study 2 (include in school) are significantly different, we combined the two survey experiments in one model. This implies that each respondent appears twice in the dataset. We therefore cluster the standard errors at the teacher level. Table 7 below presents the results from this regression.

---

5 We also examined whether any treatment effect depends on the share of minority students at the school. If teachers were characterized by racism but did not react to it in study 2 simply because the probability of having an immigrant student in their own classroom (or even at their school) was much lower in the general policy scenario in study 2, we might expect that teachers at schools with high shares of immigrants reacted more negatively to the immigrant name cue. However, as shown in supplementary appendix table A3, the interaction term between share of immigrants at the school and the name cue is positive and insignificant suggesting that teachers at schools with many immigrants are more positive. Due to the selection of teachers into schools with different shares of immigrants, we do not give any causal interpretation of the interaction term. We just note that it does not support the notion that teachers react to this reflective form of racism.

The interaction term in model 1 (table 7) confirms that the teachers responded significantly more negatively to the minority cue when they were asked to include the student in their own classroom (*p* < .014). Including covariates does not change the results (model 2). Neither does restricting the sample to teachers included in both studies (models 3 and 4). The interaction effect is also significant using ordered logistic regression (supplementary appendix table A4). Figure 1 graphically presents the results from studies 1 and 2.

Study 3—Organizational Response to Additional Workload

Study 3 investigates whether the decision to include a boy with a minority name is contingent on the resources provided to the bureaucrats and thereby the workload implications. Since the sample in study 3 is relatively small and the power only sufficient to test a few conditions, the study does not test whether the effect of resources is different for minority and majority students. The study tests the effect of two organizational initiatives compared to a control group. The control group is presented with a problematic immigrant student situation similar to the situation from study 1. The first treatment informs the teacher that the principal will monitor the teacher's test results closely and provide feedback on changes in performance. The second treatment informs the teacher that additional resources consisting of a co-teacher will accompany the inclusion of the student in her class. If the resources are important, we would expect the willingness to include the student to be higher in the resource treatment group.

The exact wording is presented in table 8. Respondents were asked to evaluate the statement on a Likert scale from 1 to 5 (mean = 3.44: SD = 1.11). Table 9 confirms that the randomization created three fairly balanced experimental groups. The response rate among the teachers assigned to treatment was 76.0%. Therefore, the balance in the sample of analysis differs, but the experimental groups do not differ significantly at the 5% level.

Figure 2 shows the results of study 3 regarding the role of organizational initiatives (based on supplementary appendix table A5). The monitor cue has a negative but insignificant effect on the willingness to include the student with a minority name in the classroom. The resource cue has a positive effect as expected (*p* < .051). When covariates are included in the model it is significant (*p* < .043). Using ordered logistic regression produces the same result (supplementary appendix table A5, models 3–4). Consequently, the willingness to include a student with an immigrant name appears to

Table 6. Studies 1 and 2—Effect of Minority Cue and In-Class or In-School Cue Variant on Willingness to Include Student (1–5)

Variables	In Class		In School	
	(1)	(2)	(3)	(4)
Minority student	−0.238*	−0.233*	0.0153	0.0188
	(0.0743)	(0.0742)	(0.0725)	(0.0729)
Female		0.0104		0.0752
		(0.0802)		(0.0763)
Age		−0.00478		−0.0119*
		(0.00377)		(0.00367)
College degree		−0.0993		−0.100
		(0.142)		(0.151)
Log earnings		−0.00259		−0.0131
		(0.0256)		(0.0328)
No valid information		−0.591		−0.994*
		(0.325)		(0.427)
Constant	3.345*	3.693*	2.292*	3.064*
	(0.0530)	(0.285)	(0.0536)	(0.399)
Observations	845	845	862	862
Adjusted R^2	0.011	0.012	−0.001	0.014
Joint F-test	10.24*	2.554*	0.0445	2.499*

*Note:* Standard errors clustered at the school level in parentheses. OLS = ordinary least squares regression.
*$p < .05$, two-sided test (coefficients).

Table 7. Studies 1 and 2 Combined—Effect of Minority Cue and In-Class or In-School Cue on Willingness to Include Student in School (1–5)

Variables	All Teachers		Only Teachers in Both Studies	
	(1)	(2)	(3)	(4)
Minority	0.0153	0.0203	0.0104	0.0158
	(0.0755)	(0.0750)	(0.0771)	(0.0766)
In class	1.053*	1.052*	1.048*	1.047*
	(0.0671)	(0.0677)	(0.0681)	(0.0687)
Minority × in class	−0.253*	−0.251*	−0.241*	−0.241*
	(0.102)	(0.103)	(0.104)	(0.104)
Female		0.0437		0.0473
		(0.0644)		(0.0653)
Age		−0.00839*		−0.00844*
		(0.00291)		(0.00295)
College degree		−0.1000		−0.0934
		(0.120)		(0.124)
Log earnings		−0.00757		−0.00911
		(0.0259)		(0.0275)
No valid information		−0.792*		−0.792*
		(0.320)		(0.334)
Constant	2.292*	2.851*	2.297*	2.868*
	(0.0529)	(0.301)	(0.0541)	(0.315)
Observations	1707	1707	1670	1670
Adjusted R^2	0.157	0.165	0.156	0.164
Joint F-test	145.9*	58.04*	142.4*	56.59*

Note: Standard errors clustered at the teacher level in parentheses. OLS = ordinary least squares regression.
*$p < .05$, two-sided test (coefficients).
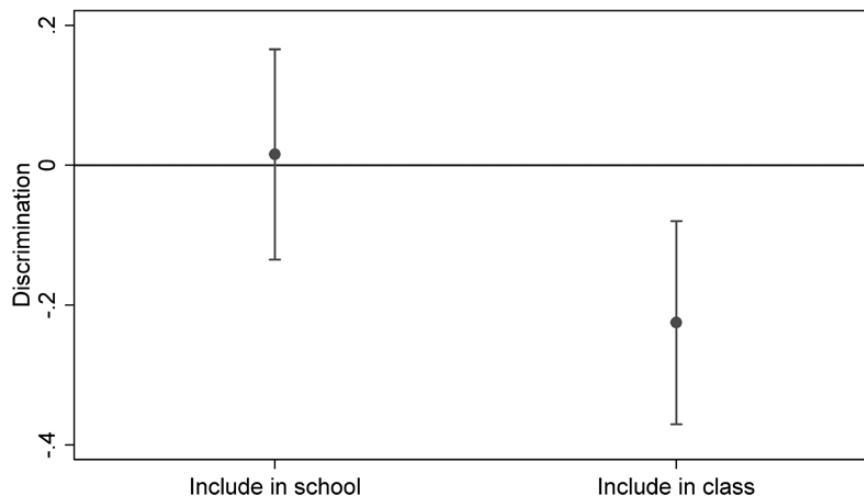
Figure 1. Comparing the Effects of Including a Minority Student in Class and in School (1–5). *Note:* Marginal effects of minority cue from combined model 95% confidence intervals. The two treatment effects (include in school and include in class) are also significantly different from each other (see interaction term in table 7).

be larger if the inclusion is accompanied with additional resources. The positive effect is estimated to be a 0.36 standard deviation increase in the willingness to include the immigrant student. Since we did not vary the ethnicity of the student in this experiment, we cannot conclude that more resources will reduce discriminatory behavior toward minority students. However, the experiment suggests that the evaluation is sensitive to resource concerns that affect the workload. The effect of workload on discriminative behavior will be tested in the fourth and final combined field and survey experimental study.

Study 4—Reducing Workload and Bias

Finally, in study 4, we test whether actually providing teachers with reduced workload reduces minority bias. We operationalize a reduced workload for teachers by assigning more preparation time and more time with the students but with the same objectives and curriculum of the class. This is similar to Tummers and colleagues' description of high workload as when "Frontline workers try to deliver the same standard of services to a great amount of people in a short timeframe" (Tummers et al. 2015, 1110). We utilize a field experiment that provided teachers with resources to organize two additional lessons in reading and writing.[6] This intervention reduced workload by giving teachers more time to work with the same number of people and the same standard of services.

To measure discrimination, we combine this field experiment with a survey experiment similar to study 1. The exact wording of the survey experiment is presented in table 10. The teachers were asked to evaluate the statement on a Likert scale from 1 to 5 (mean = 3.13: SD = 1.05).

The field experiment randomly assigned additional resources to classes and teachers or to a control group. This resource allocation came in three different versions. The three versions were randomly allocated. The first was implemented in the fourth grade and consisted only of resources needed to implement additional instruction time. The second resource allocation was also implemented in the fourth grade and consisted of resources to be used on additional instruction time and of four monthly reading tests aimed at low performing students. Finally, the third version of the resource treatment was implemented in the fifth grade and, on top of the second package, it included a language comprehension course. Regardless of the version of resource allocation, all the teachers experienced a substantial increase in instruction time while the overall objectives of their work remained constant. Because the instruction time increased while the objectives remained the same, the workload decreased. Teachers had more time to cover the same curriculum. Teachers were fully compensated for the additional instruction time including time to prepare the additional lessons. In relation to

---

6 The field trial was run as a follow-up to a trial the previous year including the same schools. When signing up for the trial in the beginning, schools were promised to have some kind of treatment in 1 of the 2 years. This agreement imposed some restrictions on the randomization in the follow-up trial the second year. The design is described in the supplementary appendix. Important to note is

that there was no self-selection into any of the treatment groups. Supplementary appendix table A7 presents different robustness checks. Between studies 1 and 4 a national reform affecting all schools increased teachers' workload and was followed up by a lock out. The level of discrimination between studies 1 and 4 may therefore not be directly comparable, due to changes in workload.

Table 8. Study 3—Experimental Treatments

Imagine your principal tells you that a new student, Ahmed, is going to start in one of the 4th grade classes at the school. Ahmed's mum and dad are both unemployed. In addition, Ahmed experiences socioemotional difficulties and hence learning disabilities.

Your principal asks you to consider making room for Ahmed in your class.

[Control group]	[Monitor treatment] To make sure this addition does not affect the rest of the students, your principal will monitor your test results closely and provide feedback on changes in your performance	[Resource treatment] To make sure this addition does not affect the rest of the students, your principal will use municipality resources on a co-teacher 8 hours per week

To what extent do you agree or disagree that this is wise?

Table 9. Study 3—Balance across Experimental Conditions

Variables	Control	Monitor	Resource
Female	0.66	0.74	0.78
Age	47.77	46.78	46.17
College degree	0.92	0.93	0.93
Log earnings	12.18	12.36	12.51
No valid information	*n < 5*	*n < 5*	*n < 5*

Note: No statistical significant differences at 5% level. $N = 258$.
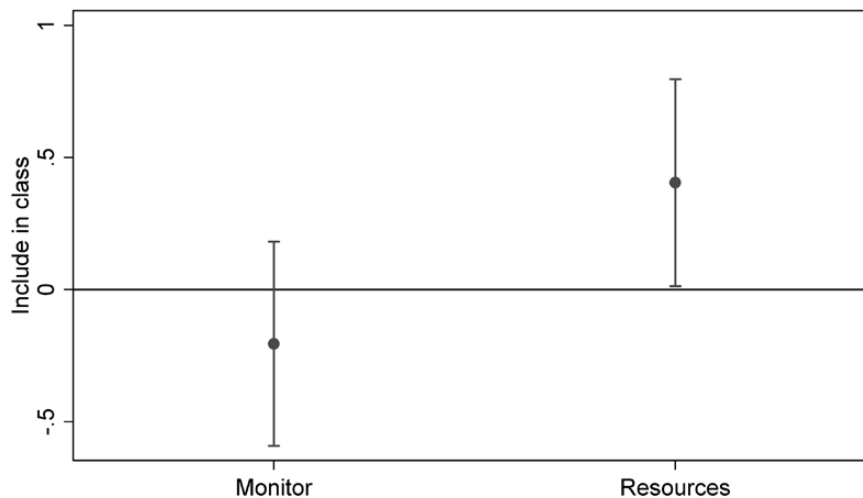
Figure 2. Effect of Monitor and Resource Cue on Willingness to Include in Class (1–5). *Note:* Marginal effects of monitor and resource cue with 95% confidence intervals. The two treatment effects (monitor and resources) are also significantly different from each other (test not shown).

our research question about the effect of workload, we do not expect any differences between the three versions of the treatment, so we pool them together in the analyses to increase the statistical power.

Attrition in relation to the survey experiment is unevenly distributed across cues and resource allocation (table 11, fourth row). A particular concern relates to whether the field-experimental treatment caused different attrition rates. However, when we compare the overall attrition rate in the field-experimental

treatment and control group, we find no significant difference (table 11, second row). In supplementary appendix table A11, we show the *p* values for all pairwise comparisons of groups which makes it clear that it is the response rate for the group of teachers who received the minority cue in the control group that is significantly different from the response rate in the other groups.

Attrition within the survey experimental groups could bias the results. However, looking at observed

covariates, there are no significant differences between the groups (table 12; the *p* values for all pairwise comparisons are shown in supplementary appendix table A12). The groups are also well-balanced in the outset (supplementary appendix table A6). As a robustness analysis, we also test that including controls for covariates does not affect the results (table 13, model 3). Another concern is that not all teachers assigned to the field treatment group implemented the additional instruction time as intended. However, we estimate the intention to treat effect, which may be a conservative estimate involving fewer assumptions than, for instance, estimating the treatment effect on the treated.

Table 13 summarizes the findings from the combined field and survey experiment in study 4. Model 1 confirms the finding from study 1 (and previous research) that the minority cue has an overall negative effect on the willingness to include a student with special needs in the classroom ($p < .004$). The inclusion of the interaction term between assignment to additional resources and the minority cue shows, however, that the effect of the minority cue is dependent on the field experimental treatment. The negative effect of the minority cue in model 2 estimates the effect of the minority cue for the teachers that were not assigned to additional resources. It should be noted that the effect of the minority cue is substantially larger when we only consider the teachers of classes who are not assigned to additional resources (compared with the average effect for both field experimental conditions in model 1). The interaction term in model 2 shows that the effect is statistically smaller for the teachers assigned to additional resources ($p < .020$). Model 3 confirms that including covariates does not change the results and using ordered logistic regression produces the same result (supplementary appendix table A7). As mentioned, some teachers from study 3 also participated in study 4, but controlling for study 3 treatment does not change the results (supplementary appendix table A9).

Figure 3 presents the marginal effects graphically. The figure shows that the effect of the minority cue is virtually eliminated when teachers have been provided with resources for more lessons before the survey experiment. The effect is maintained in the control group of the field experiment that continued as usual in terms of resources. This confirms that bureaucrats' discrimination is affected by their workload.

It is important to note that the survey experiment was a hypothetical example relating to the future.

Table 10. Study 4—Experimental Treatments

Imagine that your principal tells you about a new student [**Mathias/Yousef**] who is about to start in a 4th grade class at your school. He lives with his parents and older sister. Both parents are on welfare benefits and have been so for the last couple of years. In addition, there have been incidents of alcohol-related violence in the family. You are told that [**Mathias/Yousef**] was lagging behind the other students in his previous class and that he has concentration problems. Your principal asks you whether [**Mathias/Yousef**] could be accepted in your class. To what extent do you agree or disagree that it would be wise to accept him in your class?

Note: Bold in square brackets: Randomized cue.

Table 11. Study 4—Response Rates

Field Experiment	Control Group		Resource Treatment	
Response rate	0.79		0.86	
Survey experiment	Mathias	Yousef	Mathias	Yousef
Response rate	0.90	0.67	0.84	0.86

Note: N = 196.

Table 12. Study 4—Balance Table

Variables	Control Group		Resource Treatment	
	Mathias	Yousef	Mathias	Yousef
Female	0.89	0.90	0.86	0.88
Age	45.11	45.23	45.86	43.26
College degree	0.87	0.97	0.89	0.86
Log earnings	12.09	12.51	12.35	11.76
No valid information	*n < 5*	*n < 5*	*n < 5*	*n < 5*

Note: No statistical significant differences at 5% level. N = 161.

Therefore, teachers who were provided with resources were not less biased toward minority names because they had more resources to cope with that student. In the next section, we discuss the interpretation of the results.

Table 13. Effect of Minority Cue and Additional Resources on Willingness to Include Student in Class (1–5)

Variables	(1)	(2)	(3)
Minority	–0.483*	–0.853*	–0.819*
	(0.161)	(0.215)	(0.217)
Resource	0.0680	–0.268	–0.251
	(0.160)	(0.215)	(0.223)
Minority × resource		0.709*	0.675*
		(0.299)	(0.311)
Female			0.255
			(0.237)
Age			–0.00325
			(0.00825)
College degree			–0.474
			(0.350)
Log earnings			0.0391
			(0.0374)
No valid information			–0.121
			(0.760)
Constant	3.326*	3.478*	3.336*
	(0.144)	(0.158)	(0.432)
Observations	161	161	161
Adjusted R^2	0.040	0.063	0.061
Joint F-test	4.933*	5.775*	2.743*

*Note:* Standard errors clustered at the school level in parentheses. OLS.

*$p < 0.05$, two-sided test (coefficients).

Discussion

All four studies presented here in different ways support the notion that the teachers' workload affects the extent to which they discriminate against students with putatively minority names. As mentioned in the theory section, there are three approaches to why people discriminate against minorities: explicit racism, statistical discrimination, and implicit bias. Even though the studies presented here are not designed to test these theories—but rather to test how the organization of teachers' working conditions affect their discrimination—we believe the results shed some new light on each of them.

First, in study 2, where the inclusion of the students did not have direct implications for the individual teacher's own workload, we saw no signs of discrimination. We believe that this result speaks against the theory of explicit racism. If teachers had racial animus, they should also be opposed to the inclusion of ethnic minorities even if it did not have implications for their own workload. In study 1, where the inclusion of the student in the classroom directly affects the teacher's workload, we do see discrimination. The combined results of studies 1 and 2, therefore, provide more support for the two theoretical approaches predicting that discrimination is related to workload either because of statistical assessments or through the cognitive load on the bureaucrats.

Second, study 3 shows that teachers do react to reduced workloads by having more resources in the classroom, but it does not show clearly whether this is because of statistical assessments or more implicit associations. However, in both studies 1 and 3, effects on workload are hypothetical, which might suggest that teachers react based on some form of statistical
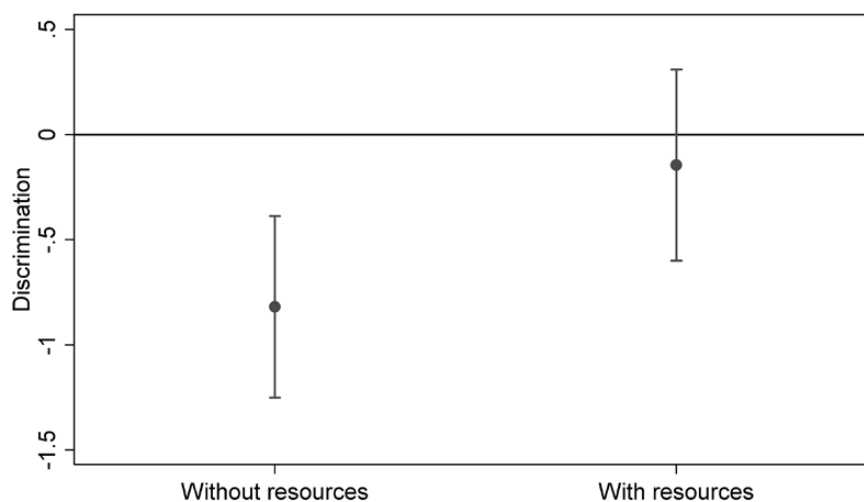
Figure 3. Effect of Minority Cue and Additional Resources on Willingness to Include Student in Class (1–5). *Note:* Marginal effects of minority cue in field experimental conditions with and wihtout extra resources 95% confidence intervals. The two treatment effects (without resources and with resources) are also significantly different from each other (see interaction term in table 13).

discrimination rather than because of changes in cognitive workload.

In study 4, teachers had reduced workloads during the period of the intervention. Their levels of discrimination were measured using a fictitious, future scenario unrelated to field experimental intervention that was about to end at the time when the teachers were surveyed. We do not have a measure of whether the objective reduction in workloads affects the subjective, cognitive load of the teachers, but the fact that the field treatment made them react differently to a future, fictitious scenario suggests that this discrimination was not based on statistical cost-benefit analyses of the implications for their own workload. It was rather due to a situation with less work pressure, making them less prone to implicit biases. In sum, the results provide some support to both the statistical discrimination and the implicit bias, but not so much the explicit racism approach.

Another theoretical question is to what extent teachers react to the ethnicity implied by the names and to what extent other factors such as socioeconomic status is affecting their reactions. In all vignettes, we also provided teachers with information about the social backgrounds of the students, which we believe should reduce how much teachers relate different socioeconomic status to the names. Furthermore, in a separate test, we do not find any indications that teachers react to information about the socioeconomic status or that this interacts with the names of the students. We therefore believe that the results indicate that teachers react to the ethnic minority implied by the names "Yousef" and "Ahmed."

To what extent the results would replicate and generalize to other countries and other occupations is—as always—difficult to know. Study 4 had uneven attrition between the experimental conditions, which is always a concern. However, the facts that attrition did not seem to be caused by exposure to the field experimental treatment, that the experimental groups balance on baseline characteristics, and that the results are robust to the inclusion of covariates make this less of a concern. The name-based discrimination that we find is an implicit replication of many studies from the United States, but whether the effects of workload on discrimination would also replicate back in the United States cannot be known without further empirical evidence.

Conclusion

The promise of modern bureaucracy is to eradicate discrimination by (among other things) the strict use of legal principles to regulate the access to public services and benefits (Weber 1922). However, when clients' demand for public services exceeds available resources, bureaucrats tend to use coping strategies to manage the demand, which may (re-), create the discrimination that bureaucracy was supposed to eradicate (Lipsky 1980). Further, if this type of discriminative behavior is susceptible to levels of workload recent trends of increased workload in the public sector might have consequences for discriminating behavior.

In the four studies reported here, we find that public school teachers' discrimination of students with putatively immigrant names is closely linked to the teachers' workload. They are more inclined to refer a child to other classrooms than their own (a typical coping strategy), if the child carries a putative immigrant, minority name than if the he has a typical native name. However, teachers do not discriminate against the children if the question is about including them at normal schools more generally, which does not directly affect their own workload. And they are more inclined to include a student with immigrant name if resources in the form of a co-teacher follow the child—or if they have just experienced a period with more resources to prepare and teach more lessons each week (without changes in the curriculum). Even though the studies presented here do not provide any definitive evidence we conclude on this basis that teachers in our study de facto discriminated against students with ethnic minority names (even if it may be completely unintended) but that this discrimination was reduced by the way their work was organized and, more specifically, by the way the workload was distributed. Even though we cannot know what it was about the names that teachers reacted to (additional tests suggest it is not their socioeconomic status), treating students differently only because of their names is a matter of discrimination.

Further research is needed to assess whether the reductions in discrimination we find here is really a result of reduced workload, and if so, whether that effect is driven by statistical discrimination or more implicit biases. Indeed, this research is only a first step in a more general research agenda asking how public administrations can be organized in ways that reduce illegitimate discrimination. One question for this research is the degree to which the results found here can be replicated in other settings. We find consistent results in four related, but different studies, but we do not know the degree to which the results generalize to other types of public organizations and services. However, many other street-level bureaucrats, for instance, social workers, police officers, and health care professionals, have to evaluate citizens with limited information and are coping with high workloads. On that basis, the type of discrimination identified here may apply to these other types of street-level

bureaucrats, and reducing workload might have similar effects (at least in the short run).

Even if future research confirms that workload has a direct effect on discrimination, we do not contend that the solution to minority discrimination would be to increase the available resources. Whether the resources needed to reduce or eliminate discrimination are worth their effect is ultimately a political question. We also note that the perceived workload (the cognitive workload) of bureaucrats may be relative, so any permanent or universal increase of resources may make the immediate effect disappear when bureaucrats mentally adapt to a new situation. As suggested by Lipsky (1980), when resources increase at one place, demand may follow the supply (in lack of a price mechanism in many public organizations). Whether this prediction is true is of course ultimately an empirical question. Consequently, the more general question for future public administration research is to find other modes of organization of the bureaucracy that may eliminate discriminating decision making—in a sustainable way. The name cue discrimination we find in these studies is also found consistently in American studies among employers, politicians, general citizens, and bureaucrats. This speaks to the universality of minority discrimination. The quest is to find equally universal means to reduce such discrimination.

Supplementary Material

Supplementary data are available at *Journal of Public Administration Research and Theory* online.

References

Baviskar, Siddhartha, and Søren C. Winter. 2017. Street-level bureaucrats as individual policymakers: The relationship between attitudes and coping behavior toward vulnerable children and youth. *International Public Management Journal* 20(2): 316–53.

Becker, Gary S. 1957. *The economics of discrimination*. Chicago: University of Chicago Press.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94(4): 991–1013.

Bodenhausen, Galen V. 1990. Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science* 1(5): 319–22.

Butler, Daniel M., and David E. Broockman. 2011. Do politicians racially discriminate against constituents? A field experiment on state legislators. *American Journal of Political Science* 55(3): 463–77.

Christensen, Robert K., John Szmer, and Justin M. Stritch. 2012. Race and gender bias in three administrative contexts: Impact on work assignments in state supreme courts. *Journal of Public Administration Research and Theory* 22(4): 625–48.

Costa, Mia. 2017. How responsive are political elites? A meta-analysis of experiments on public officials. *Journal of Experimental Political Science* 4(3): 241–254.

DeSante, Christopher D. 2013. Working twice as hard to get half as far: Race, work ethic, and America's deserving poor. *American Journal of Political Science* 57(2): 342–56.

Devine, Patricia G. 1989. Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56(1): 5–18.

Dovidio, John F., Nancy Evans, and Richard B. Tyler. 1986. Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology* 22(1): 22–37.

Downey, Douglas B., and Shana Pribesh. 2004. When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education* 77(4): 267–82.

Einstein, Katherine L., and David M. Glick. 2017. Does race affect access to government services? An experiment exploring street-level bureaucrats and access to public housing. *American Journal of Political Science* 61(1): 100–16.

Fazio, Russell H. 1990. Multiple processes by which attitudes guide behavior: The mode model as an integrative framework. *Advances in Experimental Social Psychology* 23:75–109.

Ferguson, Ronald F. 2003. Teachers' perceptions and expectations and the Black–White test score gap. *Urban Education* 38(4): 460–507.

Fiske, Susan T. 1998. Stereotyping, prejudice, and discrimination. In *The handbook of social psychology* ed. Daniel T. Gilbert, Susan T. Fiske and Gardner Lindzey, vols. 1–2, 4th ed., 357–411. New York: McGraw-Hill.

Gigerenzer, Gerd, and Wolfgang Gaissmaier. 2011. Heuristic decision making. *Annual Review of Psychology* 62(1): 451–82.

Gilbert, Daniel T., and J. Gregory. 1991. The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology* 60(4): 509–17.

Gilboy, Janet A. 1991. Deciding who gets in—Decision-making by immigration inspectors. *Law & Society Review* 25(3): 571–99.

Govorun, Olesya, and B. Keith Payne. 2006. Ego-depletion and prejudice: Separating automatic and controlled components. *Social Cognition* 24(2): 111–36.

Grohs, Stephen, Christian Adam, and Christoph Knill. 2016. Are some citizens more equal than others? Evidence from a field experiment. *Public Administration Review* 76(1): 155–64.

Hagger, Martin S., Nikos Chatzisarantis, H. Alberts, C. O. Anggono, C. Batailler, A. R. Birt, R. Brand, et al. 2016. A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science* 11(4): 546–73.

Hardin, Curtis D., and Mahzarin R. Banaji. 2013. The nature of implicit prejudice: Implications for personal and public policy. In *The behavioral foundations of public policy*, ed. Eldar Shafir, 13–31. Princeton: Princeton Univ. Press.

Hemker, Johannes, and Anselm Rink. 2017. Multiple dimensions of bureaucratic discrimination: Evidence from german welfare offices. *American Journal of Political Science* 61(4): 786–803.

Jilke, Sebastian, Wouter van Dooren, and Sabine Rys. 2018. Discrimination and administrative burden in public service markets: Does a public–private difference exist? *Journal of Public Administration Research and Theory* 28(3): 423–439.

Jilke, Sebastian, and Lars Tummers. 2018. Which clients are deserving of help? A theoretical model and experimental test. *Journal of Public Administration Research and Theory* 28(2): 226–238.

Johnson, Tiffani J., Robert W. Hickey, Galen E. Switzer, Elizabeth Miller, Daniel G. Winger, Margaret Nguyen, Richard A. Saladino, and Leslie R. M. Hausmann. 2016. The impact of cognitive stressors in the emergency department on physician implicit racial bias. *Academic Emergency Medicine* 23(3): 297–305.

Kaas, Leo, and Christian Manger. 2011. Ethnic discrimination in Germany's labour market: A field experiment. *German Economic Review* 13(1): 1–20.

Keiser, Lael R., Peter R. Mueser, and Seung-Whan Choi. 2004. Race, bureaucratic discretion, and the implementation of welfare reform. *American Journal of Political Science* 48(2): 314–27.

Lipsky, Michael. 1980. *Street-level bureaucracy: The dilemmas of the individual in public service*. New York: Russell Sage Foundation.

Ma, Debbie S., Joshua Correll, Bernd Wittenbrink, Yoav Bar-Anan, N. Sriram, and Brian A. Nosek. 2013. When fatigue turns deadly: The association

between fatigue and racial bias in the decision to shoot. *Basic and Applied Social Psychology* **35**(6): 515–24.

Meier, Kenneth, Simon Calmar Andersen, Laurence J. O'Toole Jr, Nathan Favero, and Søren C. Winter. 2015. Taking managerial context seriously: Public management and performance in US and Denmark schools. *International Public Management Journal* **18**(1): 130–50.

Meier, Kenneth J., Robert D. Wrinkle, and Jerry L. Polinard. 1999. Representative bureaucracy and distributional equity: Addressing the hard question. *The Journal of Politics* **61**(4): 1025–39.

Muraven, Mark, and Roy F. Baumeister. 2000. Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin* **126**(2): 247–59.

Nicholson-Crotty, Sean, Jason A. Grissom, Jill Nicholson-Crotty, and Christopher Redding. 2016. Disentangling the causal mechanisms of representative bureaucracy: Evidence from assignment of students to gifted programs. *Journal of Public Administration Research and Theory* **26**(4): 745–57.

Olson, Jeramiah C. 2016. Race and punishment in American prisons. *Journal of Public Administration Research and Theory* **26**(4): 758–68.

Paluck, Elizabeth Levy, and Donald P. Green. 2009. Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology* **60**:339–67.

Pedersen, Mogens Jin, Justin M. Stritch, and Frederik Thuesen. 2018. Punishment on the frontlines of public service delivery: Client ethnicity and caseworker sanctioning decisions in a Scandinavian welfare state. *Journal of Public Administration Research and Theory* **28**(3): 339–354.

Phelps, Edmund S. 1972. The statistical theory of racism and sexism. *The American Economic Review* **62**(4): 659–61.

Richeson, Jennifer A., and J. Nicole Shelton. 2003. When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science* **14**(3): 287–90.

Schneider, Anne, and Helen Ingram. 1993. Social construction of target populations: Implications for politics and policy. *American Political Science Review* **87**(2): 334–47.

Schram, Sanford F., Joe Soss, Richard C. Fording, and Linda Houser. 2009. Deciding to discipline: Race, choice, and punishment at the frontlines of welfare reform. *American Sociological Review* **74**(3): 398–422.

Simon, Herbert A. 1947. *Administrative behavior*. New York: Macmillan.

Simon, Herbert A. 1969. *The sciences of the artificial*. Cambridge: MIT Press.

Soss, Joe. 1999. Lessons of welfare: Policy design, political learning, and political action. *American Political Science Review* **93**(2): 363–80.

Soss, Joe, Richard C. Fording, and Sanford F. Schram. 2008. The color of devolution: Race, federalism, and the politics of social control. *American Journal of Political Science* **52**(3): 536–53.

Soss, Joe, Richard Fording, and Sanford F. Schram. 2011. The organization of discipline: From performance management to perversity and punishment. *Journal of Public Administration Research and Theory* **21**(suppl 2): i203–32.

Spencer, Katherine. B., Amanda K. Charbonneau, and Jack Glaser. 2016. Implicit bias and policing. *Social and Personality Psychology Compass* **10**:50–63.

Tummers, Lars. 2016. The relationship between coping and job performance. *Journal of Public Administration Research and Theory* **27**(1): 150–62.

Tummers, Lars L. G., Victor Bekkers, Evelien Vink, and Michael Musheno. 2015. Coping during public service delivery: A conceptualization and systematic review of the literature. *Journal of Public Administration Research and Theory* **25**(4): 1099–126.

Vaughan-Whitehead, Daniel. 2013. *Public sector shock: The impact of policy retrenchment in Europe*. Cheltenham, UK: Edward Elgar Publishing.

Weber, Max. 1922. *Wirtschaft Und Gesellschaft*. Tübingen: Mohr.

Wenger, Jeffrey B., and Vicky M. Wilkins. 2009. At the discretion of rogue agents: How automation improves women's outcomes in unemployment insurance. *Journal of Public Administration Research and Theory* **19**(2): 313–33.

White, Ariel R., Noah L. Nathan, and Julie K. Faller. 2015. What do I need to vote? Bureaucratic discretion and discrimination by local election officials. *American Political Science Review* **109**(1): 129–42.

Yarnell, Lisa M., and George W. Bohrnstedt. 2017. Student–teacher racial match and its association with black student achievement: An exploration using multilevel structural equation modeling. *American Educational Research Journal* **55**(2): 287–324.