

Social categories in the public sector:
When they are applied and
how to reduce discriminatory use

Thorbjørn Sejr Guul

Social categories in the public sector:
When they are applied and
how to reduce discriminatory use

PhD Dissertation

Politica

© Forlaget Politica and the author 2018

ISBN: 978-87-7335-224-3

Cover: Svend Siune

Print: Fællestrykkeriet, Aarhus University

Layout: Annette Bruun Andersen

Submitted December 22, 2017

The public defense takes place March 23, 2018

Published March 2018

Forlaget Politica

c/o Department of Political Science

Aarhus BSS, Aarhus University

Bartholins Allé 7

DK-8000 Aarhus C

Denmark

Table of Contents

Acknowledgment	7
Chapter 1: Introduction	9
Chapter 2: Theoretical argument	15
What do we know about how to reduce discriminatory use of social categories?.....	15
Why social distance matters for discriminatory use of social categories	17
How cognitive resources affects discriminatory use of social categories	19
Limited attention as explanation for discriminatory use of social categories	20
Chapter 3: Methodology	23
Design and outcomes.....	23
Case selection.....	27
Data, sample selection, and recruitment.....	27
Chapter 4: Results	29
When are social categories used discriminatorily?	29
How can discriminatory use of social categories be reduced?	32
Chapter 5: Conclusion	35
Chapter 6: Discussion.....	37
Limitations.....	37
Political and practical implications	39
Theoretical implications and future research	41
Literature	45
Summary	53
Dansk resumé	55

Acknowledgment

Writing a PhD is a huge challenge, and I could not possibly have finished it without the help from numerous people. I first wish to thank Simon Calmar Andersen and Michael Rosholm for outstanding supervision during the past three years. Michael, particularly for your expertise in labor economics. Simon, particularly for your expertise in school research and for introducing me to the world of academia within public administration. This dissertation would not have been possible without your help. Thank you for all the critical feedback that has improved the quality of this dissertation tremendously. I also owe my co-authors a great deal: Maria Humlum, Simon Calmar Andersen, Anders Villadsen, and Jesper Wulff. Thank you for making our joint work possible. I feel very privileged to have had the opportunity to work with such skilled researchers and have learned a great deal from the process. The collaboration has both increased the quality of the research tremendously and made the process much more enjoyable – thanks for that.

I also feel privileged to have had the opportunity to be part of the Department of Political Science at Aarhus University – a department characterized by both a friendly environment and high academic standards. Thanks particularly to the members of the public administration section for great companionship during conferences and long section meetings and for invaluable and highly constructive feedback on my project and papers. Also thanks to the PhD group for both enriching academic and social activities during the years of my PhD. Writing a PhD could have been a very lonely process, however, thanks to the excellent environment at the Department it has been nothing like that. My office mates Mathias and Mads deserve special mention. Mathias particularly for an excessive amount of coffee runs during the first semester of my PhD; Mads particularly for sharing all the ups and downs in life related to the process of writing a dissertation for the remaining years of my PhD. Thank you both for many great academic as well as non-academic discussions. Special thanks also go to Jakob for great data collection collaboration and for both discussions closely linked to the work of my PhD and matters pretty removed. During my PhD, I have enjoyed the additional privilege of being part of Tryg-Fonden's Centre for Child Research, which has both enabled discussions with and generated high quality feedback from people from different academic disciplines and many great non-academic discussions as well. Thanks particularly to the junior research group in both regards. Last but by no means least, thanks to Malene Poulsen and Ruth Ramm for helping out with receipts and

bookkeeping, to Annette Andersen, Anja Dalsgaard, Lone Winther and Catherine McDonald for excellent language editing and to Helle Bundgaard for excellent help in relation to my teaching duties.

Various people outside Aarhus University also need mentioning. Particularly all the school principals, teachers, students, mentors and mentees. Without your contributions to my research, I would not have been able to write this dissertation. Also thanks to Gøsta Esping-Andersen and the demosoc research group at Pompeu Fabra University for your extremely welcoming attitude and for giving me the opportunity for a very rewarding research stay. Moreover, thanks for making my first paper presentation a great experience with helpful feedback.

Finally, but most importantly this PhD would not have been possible without the support from my family. Especially, I want to thank my parents for nurturing my curiosity about the world from a very stage early stage and my wife, Louise, for unconditional support through the at times rather frustrating process of writing a PhD. But most importantly for making my life with our little family outside academia the greatest privilege of all.

Chapter 1: Introduction

One of the virtues of modern bureaucracy, according to Weber, is that the bureaucracy functions in a spirit of formalistic impersonality and by *sine ira et studio* – without hatred and passion: “*Everyone is subject to formal equality of treatment; that is, everyone in the same empirical situation.*” (Weber 1947, 340). However, studies of modern bureaucracy show that case managers sanction ethnic minority welfare clients harder than ethnic majority welfare clients for the same deviant behavior (Schram et al. 2009). Similarly, teachers consider children’s problems as more worrisome if they belong to another social class than themselves (Harrits and Møller 2014), and ethnic minority job applicants with the same qualifications as majority applicants receive fewer callbacks for public job vacancies (Villadsen and Wulff 2017). In fact a substantial number of studies find that social categories – broadly defined as categories of people which acquire their meaning by contrast with other categories (Hogg 2001, 56) – such as gender, ethnicity and social class, affects the treatment citizens receive in the public sector. Research suggests that such social categories tend to make individuals exaggerate perceived similarities between the in-group to which they belong and accentuate perceived (stereotypical) differences to out-group members (Shelton, Richeson, and Dovidio 2013). In some instances, differential treatment might be desirable such as when a teacher makes an extra effort to help a student that is falling behind in class. However, as the examples above show, differential treatment based on social categories can turn into outright discrimination. In this dissertation, I define discriminatory use of social categories as cases in which we cannot attribute differential treatment to objective differences in needs or qualifications – to use Weber’s words, people in the same empirical situation. For example, when an ethnic minority applicant with the same skills as a majority applicant does not receive a callback for a job interview, while the majority applicant does.

Aside from the immediate detrimental effects of differential treatment for the individual, this is problematic for several other reasons. As mentioned, the promise of modern bureaucracy is “*to eradicate prejudicial behavior through universalistic treatment*” (Lipsky [1980] 2010). In addition, misperceptions of citizens’ needs and abilities might lead to inefficient service delivery or affect the citizens directly by creating different expectations and thereby, self-fulfilling prophecies (Steele and Aronson 1995; Rosenthal and Jacobson 1968; Jussim and Harber 2005). Finally, (negative) experiences in direct encounters

with government institutions can have severe consequences for political efficacy and participation (Soss 1999; Schneider and Ingram 1993; Hjortskov, Andersen, and Jakobsen Forthcoming).

Considering insights from social psychology it is not particularly surprising that social categories affect public employees' behavior. As summarized by Lau and Redlawsk, humans are *"limited information processors" (...)* who have become quite adept at applying a variety of information "shortcuts" to make reasonable decisions with minimal cognitive effort in all aspects of their lives." (2001, 952). Considering that resources in the public sector are scarce because no price mechanism limits the demand for public service (Lipsky [1980] 2010), we should expect to find that public employees, like other human beings, sometimes use social categories as such shortcuts when they interact with citizens. Particularly since studies of implicit discrimination find that discrimination might be completely unintentional and outside the discriminator's awareness (Bertrand, Chugh, and Mullainathan 2005). While this suggests that such behavior is partly the result of an automatic process, studies in social psychology find that it is individually controllable and responsive to social structures (Fiske 1998). Still, few studies examine how to reduce differential treatment by politically controllable means (Paluck and Green 2009; Hardin and Banaji 2013; Bertrand and Duflo 2016). *This dissertation examines when social categories are applied and most importantly, how to reduce their discriminatory use in the public sector.*

To answer the question this summary proposes a theoretical model that can explain when discriminatory use of social categories occurs and how to reduce such use. To build this model, I combine the findings from this dissertation's papers which draw on public administration theories about representative bureaucracy (Keiser et al. 2002; Meier and Nigro 1976), street level bureaucrat coping (Lipsky [1980] 2010; Tummers et al. 2015), and performance information (Rockoff et al. 2012; Moynihan and Landuyt 2009) with related insights from social psychology. We cannot both build and test a theoretical model based on the same data and my model obviously does not account for all possible sources and explanations for when discrimination occurs and how to reduce such use. Still, the model provides a first attempt to make a public administration model based on insights from the research in this dissertation as well as related research areas that is capable of guiding the development of interventions that hopefully will reduce discriminatory use of social categories.

More specifically, the dissertation argues that the salience of social distance affects discriminatory use of social categories. That is, the more accentuated differences between social categories is the more discriminatory use of these categories. Lack of cognitive resources amplifies such use meaning when

public employees are under cognitive pressure or when they are cognitively depleted, and I argue that one particular reason for this is that it limits attention to the influence of social categories in stereotypical and consequently discriminatory ways. Correspondingly, I argue that reducing social distance, releasing cognitive resources or focusing attention on potentially discriminatory use of social categories might reduce such use. The dissertation does not directly examine the validity of these three psychological elements, however, from this model and in combination with the public administration theories mentioned above, the model has a number of implications concerning discriminatory use of social categories in the public sector and how to reduce such use. Specifically, the dissertation suggests that reducing social distance by *making the bureaucracy more representative* of the citizens they serve, releasing cognitive resources by *reducing workload* or focusing attention on discriminatory use by *providing systematic information* that increase awareness of discriminatory influence of social categories can reduce discriminatory use. In addition to this summary the dissertation consists of the following papers:

Table 1. Overview of the articles in the dissertation

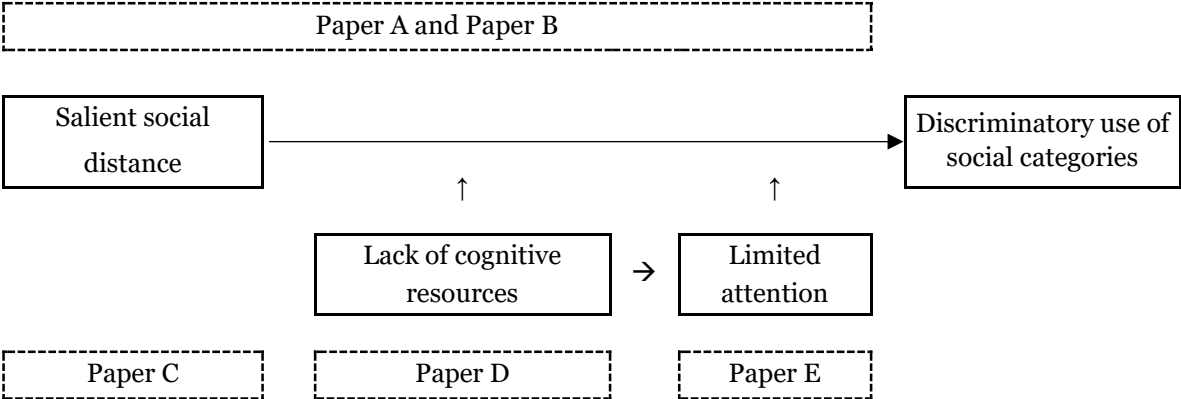
Articles	Short title ^{a)}
(Any) barriers for representative bureaucracy? How ethnicity affects public managers' evaluation of applicants. Under review.	Barriers
Does good performance reduce bad behavior? Antecedents of ethnic discrimination in public organizations. (Co-authored with Anders R. Villadsen and Jesper Wulff) Working paper.	Performance
The individual-level effect of gender match in representative bureaucracy. Accepted for publication in <i>Public Administration Review</i> .	Gender match
Minority discrimination at the front line. Combined survey- and field experimental evidence. (Co-authored with Simon Calmar Andersen). Invited for revise and resubmit in <i>Journal of Public Administration Research and Theory</i> .	Discrimination
Reducing the achievement gap between students of high and low socioeconomic status. Evidence from a field experiment (Co-authored with Simon Calmar Andersen and Maria Humlum). Working paper.	Misperceptions

a. The short titles are used throughout the rest of the summary

Figure 1 summarizes the overall argument of the dissertation and places the papers in relation to this argument. While it is difficult to test implications of all elements of the model in one paper, I argue that the evidence across papers supports the proposed model. Paper A and B examine the overall argument of the dissertation and particularly focus on answering the question of when public employees use social categories in discriminatory ways. Paper C, D, and E

each have a more specialized focus on the three basic elements in the argument. In addition, they not only provide evidence on when discriminatory use occurs, but also provide evidence concerning how to reduce discriminatory use. The dissertation combines quasi-, field- and survey-experimental evidence in support of the proclaimed model. These results speak to a literature about public sector discrimination and have important implications for how to combat this kind of injustice. Within the specific papers I explain how individually, they also contribute to the literature on representative bureaucracy, street level bureaucrat coping, and performance information respectively, on which the overall argument draws.

Figure 1. Illustration of the papers relationship to the overall argument of the dissertation.



While I present the argument in general terms, the empirics of the dissertation has a specific scope. Obviously any focus of scope also implies limitations. The dissertation focuses on the social categories of gender, ethnicity and social class, and discriminatory use by public managers and street level bureaucrats in the educational sector. Street level bureaucrats are defined as public employees that interact directly with citizens and have substantial discretion in how they execute their work (Lipsky [1980] 2010, 3). This focus provides several advantages. First, the educational setting, in particular primary and lower secondary public schools, represents one of the largest public-sector employers and accounts for significant levels of public spending across countries (Meier, Wrinkle, and Polinard 1999). Second, public managers and street level bureaucrats evaluate citizens in situations in which they are able to apply social categories. Public managers in particular, when they choose which applicants they call for an interview and subsequently who they decide to hire, and street level bureaucrats particularly when they interact and make decision in relation to the citizens they serve. Third, I focus on the social categories of gender, ethnicity, and social class because existing research repeatedly finds discriminatory use of these social categories (e.g. Rangvid 2015; Harrits and

Møller 2014; Keiser et al. 2002; Villadsen and Wulff 2017). I would argue that determining when discriminatory use occurs and how to reduce such use is most relevant in categories where we actually know discriminatory use often occurs. I will discuss the strength and limitations of this focus in more detail at the end of the dissertation.

Finally, discriminatory use contrasts with purposeful use. I completely acknowledge that social categories – like other categories – can be used in purposeful ways and enable fast, frugal and sometimes even better decisions without gathering and processing huge amounts of information (Gigerenzer and Gaissmaier 2011). This might for instance reduce complexity in public organizations and increase efficiency. However, people tend to accentuate perceived (stereotypical) differences to out-group members. Using social categories as informational shortcuts might therefore not lead to valid conclusions about an individual's needs and qualifications. In addition, even when there is a valid relationship between a social category and needs or qualifications, it might still not be justified to rely on this relationship. For instance, it might be true that a student's gender contains relevant information about the student's abilities if boys on average are slightly worse at reading than girls (Beuchert and Nandrup 2014). However, if a teacher's perception of a student's abilities is based on this when objective information is available, this still has detrimental effects from the perspective of the individual; the individual receives differential treatment as a result of their ethnicity, gender, or social class.

In the next chapter I review existing studies on why social categories are used in discriminatory ways. From this literature, I infer the overall argument of the dissertation concerning when we can expect discriminatory use and how to reduce such use. In chapter 3 I explain the design of my papers before summarizing the evidence in favor of my argument in chapter 4. I conclude on the findings in chapter 5 and discuss the results in chapter 6.

Chapter 2: Theoretical argument

What do we know about how to reduce discriminatory use of social categories?

This chapter accounts for existing explanations concerning why social categorizations are used in discriminatory ways and what we know about how to reduce such use. The chapter draws on existing literature on discrimination, categorization, and stereotyping. Concerning the first, a substantial literature shows that in particular, people belonging to racial or ethnic minority groups experience discrimination. For instance a study in the U.S. shows that job seekers with a black sounding name receive fewer callbacks for job interviews than white job seekers with the same skills (Bertrand and Mullainathan 2004). Similarly job seekers with ethnic minority names receive fewer callbacks in a Scandinavian context (Midtbøen 2016; Carlsson and Rooth 2007). In a similar vein, on average Americans reward white welfare applicants more for being hard working and punish them less for being lazy than their black counterpart (DeSante 2013). Also white politicians are more likely to respond to potential voters with a white sounding name than to a black sounding name (Butler and Broockman 2011). Allport – the father of the modern understanding of prejudice – argued that categorizing is unavoidable: “(...) *the human mind must think with the aid of categories.*” ([1954] 1979, 20). With that in mind it is not surprising that a number of studies show that public employees also apply social categories in a discriminatory way. Case workers punish welfare recipients with a black sounding name harder than their white counterparts for the same misconduct (Schram et al. 2009), voting officials respond more often to questions related to voter registration from citizens with white sounding names compared to a black sounding names (White, Nathan, and Faller 2015), and the same is true for Latinos compared to whites in relation to help in a housing program (Einstein and Glick 2017). However, as Lipsky argues in his seminal work *Street level Bureaucracy*: “*It would be as much of a mistake to infer that ethnic or racial appeals always prevail in affecting discretionary judgements as that they never prevail.*” ([1980] 2010, 108). Therefore, to understand when discrimination based on social categories such as ethnicity occurs, we need to understand why discrimination occurs in general.

Traditionally the literature on discrimination has suggested three different but not necessarily mutually exclusive explanations for discriminatory behavior. The first perspective suggests that discrimination is taste-based (Becker 1957). This means that discrimination occurs simply because people with one set of social characteristics dislike people belonging to other social groups. The second perspective argues that discrimination happens because belonging to a specific social group is associated with specific (negative) attributes due to a statistically greater or smaller prevalence of these attributes in the social group in question compared to other social groups (Phelps 1972). The third perspective suggests that discrimination occurs because of implicit prejudice based on social stereotypes (Hardin and Banaji 2013). This dissertation primarily builds on the latter understanding of discrimination as something unconscious and outside the discriminator's awareness (Bertrand, Chugh, and Mullainathan 2005). As mentioned this does not mean that the other explanations are redundant. However, this perspective has decades of considerable support in the literature on stereotypes in social psychology (Devine 1989; Dovidio 1986; Fiske 1998) but also support in studies of case workers (Schram et al. 2009), employers (Rooth 2010) and regarding other administrative processes such as the writing of majority opinions in state supreme courts (Christensen, Szmer, and Stritch 2012). This suggests that the last perspective might prove most fruitful for understanding why public employees use social categories in discriminatory ways.

While discrimination prevails in numerous different settings as the reviewed studies above indicate, how to reduce discriminatory use by politically controllable means remains an important question with few well-known answers. Several reviews across disciplines support this claim. For instance, a review of psychological studies of prejudice concludes: "*Notwithstanding the enormous literature on prejudice, psychologists are a long way from demonstrating the most effective ways to reduce prejudice. Due to weaknesses in the internal and external validity of existing research, the literature does not reveal whether, when, and why interventions reduce prejudice in the world.*" (Paluck and Green 2009, 360). Similarly, a recent review of economic studies concerning field experiments on discrimination concludes: "*While field experiments in the last decade have been instrumental in documenting the prevalence of discrimination, field experiments in the future decade should aim to play as large of a role in isolating effective methods to combat it.*" (Bertrand and Duflo 2016, 85). They further argue that the dearth of studies on how to combat discrimination is surprising, since a rich theoretical and lab-based psychological literature provides ample suggestions concerning how this could be done. However, findings from the laboratory should be applied cau-

tiously to the real world (Spencer, Charbonneau, and Glaser 2016). Interventions designed and tested in the laboratory rarely resemble real world situations or correspond to factors which can be modified by politically controllable means. Recent work suggests that the most fruitful approach might be to preclude discriminative behavior rather than directly reducing implicit biases of which people are not even aware (ibid.). This dissertation therefore takes the psychological literature on social categories and stereotypes as an outset and combines this with insight from public administration studies on representative bureaucracy, street level bureaucrat coping, and performance information in order to answer the question of when discriminatory use occurs and how discriminatory use of social categories can be reduced.

Why social distance matters for discriminatory use of social categories

The following section explains why social distance might matter for discriminatory use of social categories. As mentioned, social categories can broadly be defined as categories of people which acquire their meaning by contrast with other categories (Hogg 2001, 56). A key finding concerning social categories in the literature on social psychology is that they accentuate perceived (stereotypical) differences between social categories and exaggerate perceived similarities within groups (Hogg 2001, 59). This process also applies to the basic distinction between the self (in-group) and other groups (out-groups) (Shelton, Richeson, and Dovidio 2013). Thus, when an in-group/out-group dimension becomes salient, public employees might tend to apply stereotypes to a greater degree and ultimately use social categories in more discriminatory ways.

A number of sociological studies find that social distance between public employees and the citizens they serve directly affects their interaction with these citizens. For instance Harrits and Møller find a tendency for middle-class street-level bureaucrats to evaluate otherwise similar children to have more worrisome problems when they have a lower class (or even higher class) background (2014). Similarly Ridgeway (2014) argues that street level bureaucrats “(...) create a context in which the implicit interactional rules are better understood and more familiar to middle-class petitioners than to working class ones”. Quantitative studies also show that teachers tend to assess students from different social categories differently. Rangvid shows, for instance, that male students, students with low educated parents, and ethnic minority students receive lower teacher grades compared to girls, students with high educated parents, and ethnic majority students even though they receive the

same grades at their final exams (2015). This work also shows how social categories might be particularly relevant in relation to public employees due to structural conditions. Teachers, for example, usually have a middle-class background, whereas they tend to serve a quite diverse population of students.

Representative bureaucracy theory relies on the same underlying assumption of the relevance of social distance between the bureaucracy and the citizens they serve. The theory proposes that the public administration should represent the population they serve on demographic characteristics (Mosher 1968). At the core of the theory is the distinction between passive and active representation. The bureaucracy passively represents the population when it resembles the population's diversity in terms of demographic characteristics. The bureaucracy actively represents the population when bureaucrats use their discretion to eliminate discrimination and pursue the interests of the group of citizens they represent (*ibid.*). The theory presents several arguments for this. First, the literature argues that bureaucrats will treat citizens belonging to their own social category better than bureaucrats belonging to different categories – either because other bureaucrats discriminate against outgroup citizens and bureaucrats belonging to the same group do not, or because bureaucrats belonging to a specific group favor in-group citizens (Lim 2006). Second, bureaucrats might better understand the needs of citizens with a similar background to their own (*ibid.*). Third, social diversity might affect other bureaucrats' views of citizens belonging to other social groups (Pitts 2005). Finally, the presence of bureaucrats belonging to certain social categories might both increase citizens' perceived legitimacy of the service (Riccucci, Ryzin, and Lavena 2014) and their willingness to contribute to its delivery (Riccucci, Van Ryzin, and Li 2016). The underlying argument appears strongly in line with the argument derived from the social psychology literature: Salient social distance might result in differential treatment of citizens.

While a substantial literature supports the overall relationship between passive representation and improved outcomes for represented citizens, most studies in this literature analyze organizational level effects of passive representation on outcomes of the represented citizens, which makes the underlying reason for these results open to interpretation (Andrews, Ashworth, and Meier 2014, 19; Atkins, Fertig, and Wilkins 2014, 504; Bradbury and Kellough 2011, 164) and thereby also the question of whether reducing social distance really reduces discriminatory use of social categories in the public sector. In addition, most studies examine the role of representation in settings where bureaucrats and citizens seldom interact one-on-one – particularly in schools (e.g. Meier, Wrinkle, and Polinard 1999; Atkins, Fertig, and Wilkins 2014; Atkins and Wilkins 2013). By contrast, many street-level bureaucrats such as

case workers and health workers do deliver their service on a one-on-one basis. While previous studies support that passive representation affects bureaucrat values (Ricucci and Meyers 2004) and citizens' perceptions of interaction quality (Gade and Wilkins 2013) in such settings, few studies examine how passive representation or in other words, the social distance itself affects bureaucrat behavior and ultimately citizen outcomes.

In sum, salient social distance between public employees and citizens might affect discriminatory use of social categories. A number of sociological studies support the influence of such distance on public employees' attitudes and behavior towards citizens belonging to other social categories. Building on the same underlying assumption, the theory of representative bureaucracy suggests that decreasing the social distance between the bureaucracy and the citizens they serve will decrease discriminatory use of social categories. However, even when social distance is salient, studies in social psychology find that stereotypical reactions are individually controllable and responsive to social structures (Fiske 1998) This suggests that other factors might reduce or amplify discriminatory use of social categories.

How cognitive resources affects discriminatory use of social categories

This section explains why we should consider cognitive resource as a factor that might reduce or amplify discriminatory use of social categories. This argument is based on social psychological literature on stereotypes and public administration literature concerning how street level bureaucrats cope with high workload. An important finding in the literature on stereotypes is that people can avoid the influence of stereotypes when they are appropriately motivated and have sufficient cognitive resources (Fiske 1998). Social psychological lab experiments show that individuals activate stereotypes less often under cognitive load (when they are conducting a demanding task), but when they do they more often apply these stereotypes to the subject in question (Gilbert and Hixon 1991). In a similar manner, studies of ego-depletion show that individuals also apply stereotypes more often after a cognitively depleting task (Govorun and Payne 2006). A number of correlational studies from real world situations also support this pattern. For instance, a sample of police recruits were more likely to apply racial stereotypes after reporting a night of poor sleep (Ma et al. 2013). Similarly, a study of physicians found that they applied more implicit stereotypes after a dangerously crowded or extremely busy work shift than before (Johnson et al. 2016).

Applying this knowledge in more general terms to the public sector raises cause for concern. Without a price mechanism, street level bureaucrats often

face a demand that exceeds available resources (Lipsky [1980] 2010) and quick decisions are therefore often needed. As a consequence, street-level bureaucrats may tend to respond to citizens from different social categories based on the use of stereotypical categorizations because they can work as heuristics that gives fast but inaccurate information about a citizen compared to assessing the same citizen individually. A recent review of the literature on street level bureaucrat coping mechanisms supports that street level bureaucrats use different strategies to make ends meet and reduce their workload (Tummers et al. 2015). This might explain why non-whites are sanctioned more than white welfare recipients within the same district (Keiser, Mueser, and Choi 2004) and prison staff disproportionately punish black males more than white males (Olson 2016). However, as the lab experiments above indicate, changes in workload might also affect the extent to which such differential treatment occurs. Maynard-Moody and Musheno argue that stereotypical categorizations by street level bureaucrats tend to be reinforced when there is a need to make quick judgements (2003). This is contrasted by the detailed account they produce for clients when they are given more time (ibid.). In a similar vein Dias and Maynard-Moody argue that street level bureaucrats want to make individual assessments that demand details, emotional interaction with clients and that discovers the individual client's areas of interest (2007).

A few existing studies support the notion that the level of discrimination varies with organizational circumstances. For instance, studies show that differences in administrative processes affect gender and racial biases (Wenger and Wilkins 2009; Christensen, Szmer, and Stritch 2012). Soss, Fording, and Schram similarly find in a sample of case workers in Florida, that recent trends of New Public Management reforms have increased discriminatory behavior (2011). While the studies mentioned above plausibly argue for the relevance of workload for discriminatory behavior, none of them test whether reducing workload indeed reduces discriminatory use of social categories in the public sector.

Limited attention as explanation for discriminatory use of social categories

So far, I have argued that salient social distance affects use of social categories, and that people will rely more on such categories when they lack cognitive resources. But why is that? Building on the understanding of discrimination as something unintentional and outside of the discriminator's awareness, one explanation relates to limited attention to how social categories influence public

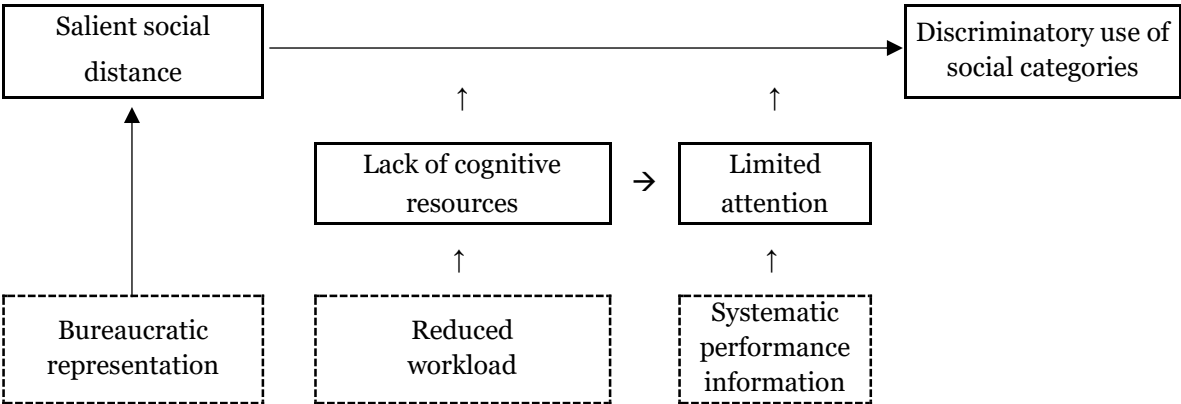
employees' perceptions of citizens. Studies on processing of performance information have examined for decades how limited attention shapes decision making. For instance, Simon argued in his seminal work *Administrative Behavior* that it requires conscious attention to prevent what he labeled habitual behavior, which means automatic and unconscious response to stimulus (Simon [1947] 1997, 100). Also, more recent work argues that limited attention might explain why apparently important aspects of a production function are not considered (Hanna, Mullainathan, and Schwartzstein 2014). The educational production function might prove particularly complex (Figlio 1999). In addition, limited attention might have particular importance in public organizations in general because they are multipurpose organizations and consensus concerning the relative importance of goals is often lower than in the private sector, where most stakeholders agree that financial success is essential (Boyne, Andrews, and Walker 2006, 15). The risk of overlooking or paying insufficient attention to some factors – including some students in the classroom – may therefore be particularly relevant in this setting.

The literature on performance information in public administration argues how systematic performance information can shape and inform decisions. This question has received considerable attention in private sector studies for at least half a decade (Cyert and March 1963). While studies of the use of systematic performance information coupled with financial incentives to improve performance on the measured dimensions finds strategic and perverse responses to such systems (Jacob and Levitt 2003; Figlio and Loeb 2011), scholars have suggested that performance information can also be used for learning purposes (Moynihan 2005; Bryson, Berry, and Kaifeng Yang 2010). Moynihan and Landuyt argue for instance, that implementing systematic information systems and giving the employee on the front line flexibility to do what they find best, increases potential for organizational learning (2009). Similarly a number of studies show how school principals shift priorities (Nielsen 2014) and update their perceptions of their employees (Rockoff et al. 2012) as a response to new inputs. Other existing studies show that performance information affects levels of innovation (Nicholson-Crotty, Nicholson-Crotty, and Fernandez 2017; Salge 2011). However, these studies do not examine whether performance information affects discriminatory behavior such as callback discrimination, nor do they examine how systematic performance information provided to employees on the front line might change their attention to ways in which their perceptions of citizens based on social categories could be inaccurate.

In sum, I argue that the salience of social distance affects discriminatory use of social categories; lack of cognitive resources amplify such use because

it limits attention to discriminatory use. Correspondingly, I argue that by reducing social distance by bureaucratic representation, releasing cognitive pressure by reducing workload or changing attention by providing systematic performance information, we can accomplish less discriminatory use of social categories in the public sector. Figure 2 summarizes the argument (similar to figure 1) and the proposed interventions for reducing discriminatory use.

Figure 2. Illustration of the overall argument of the dissertation and the elements that might reduce discriminatory use of social categories.



Chapter 3: Methodology

Design and outcomes

Providing credible answers to the question of when public employees use social categories in a discriminatory manner and how to reduce such use is not a simple task. First, simply establishing whether differential treatment occurs requires a valid measure of preferences and behavior. Social desirability complicates this since public employees might prove reluctant to recognizing that differential treatment occurs since “(...) *discrimination is outlawed and egalitarian norms are widely endorsed*” (Schram et al. 2009, 401). In addition even if they are willing to admit to differential treatment, they might not even be aware of unintended discrimination (Bertrand, Chugh, and Mullainathan 2005).

Second, public employees and citizens often (self-) select into specific organizations and positions. Students (or their parents) with specific characteristics self-select into specific school districts and principals assign teachers with specific characteristics into classrooms with specific challenges (Dieterle et al. 2015). Similarly, we can expect principals to assign more demanding tasks that imply higher workloads to more capable teachers, and teachers more aware of specific socioeconomic groups of students might systematically differ from their colleagues. Identifying exogenous sources of variation in social distance, cognitive resources and attention is therefore a necessary condition for establishing the relevance of these factors.

To address the challenges, this dissertation combines quasi-, field- and survey-experimental designs to both uncover when differential treatment – even though it is unintentional – occurs, and to support that the hypothesized factors actually reduce discriminatory use of social categories. More specifically, I use survey and field experimental approaches to measure discriminatory use without social desirability bias. In *Barriers*, I conduct a survey experiment with public school managers to uncover whether their assessment of a potential teacher at their school is dependent on the ethnicity of the applicant and of the social distance between the teacher and the students. To elaborate, I embedded a survey experiment in a school manager survey. I presented all school managers with a description of a potential teacher for their school. I randomized the minority status of the teacher and the students that the teacher was supposed to teach. As an outcome, I asked the school managers to

assess the potential teacher on several different dimensions related to the applicant's abilities, social fit, and overall fit within the managers' school.

In *Performance*, we build on a previous field experiment where applications with minority and non-minority names are sent to real job vacancies (see Villadsen and Wulff 2017 for a more detailed account). We use the callback rate for job vacancies at public schools from this previous study as an outcome measure to estimate the level of differential treatment based on ethnicity in the callback situation. Conducting field experiments is costly both for researchers and society, so using the same field experiments for different purposes is a way of generating more knowledge without additional cost. This approach is widespread and has resulted in significant new knowledge (e.g. Dee 2004; Chetty et al. 2011 use of Tennessee's Student Teacher Achievement Ratio (STAR) project. See Schanzenbach 2006 for a review of studies based on STAR). We combine this with administrative school data to examine whether school performance correlates with level of discrimination.

In *Gender match*, I examine whether matching bureaucrat and citizen by gender as a case of social distance affects the outcome of a mentor program. The ultimate goal of the program is to get the mentees enrolled in an educational program. I also examine the effects on the mentors' time use on different mentees, and mentees' reported effort towards the final goal of becoming enrolled in an educational program. In this study, I rely on an as good as random matching of mentors with mentees (job seekers) in an unemployment program. In addition to bolstering the results, I capitalize on the fact that the mentors in the program have several mentees and use within-mentor fixed effects. While this combination takes many potential alternative explanations for the results into account (particular mentor quality), selection is still possible. However, I show that the matched and unmatched jobseekers resemble each other on important observable characteristics such as unemployment and educational enrollment history, which supports that we can in fact consider the matching as good as random.

In *Discrimination*, we conduct three survey experiments and a combined survey- and field experiment with teachers to examine how workload affects their willingness to include a student with a minority background in their classroom. As primary outcome, we use the teachers' reported willingness to accommodate a potentially problematic student in their classroom (or at the school in general in one variation). We argue in line with *Barriers*, that the survey experimental approach reduces possible social desirability bias and thus provides a better measure of possible discriminatory attitudes. A possible limitation to the design is that we do not measure how the potentially discriminatory attitudes affect behavior. I will return to this limitation in the final discussion of the summary.

Finally, in *Misperceptions* we conduct a field experiment where we randomly assign teachers and their classrooms to an intervention (the intervention is a follow up study on a previous field experiment reported in Andersen, Humlum, and Nandrup 2016). While the intervention consists of several elements of particular interest, the intervention provides the teachers with monthly information on all their students' abilities in order to examine whether increased attention to the needs of the students might change perceptions and ultimately return to schooling for students belonging to different social classes. We operationalize social class in terms of socioeconomic status based on parents' education level. While income inequality is low in Denmark, educational inequality is still substantial and thus we argue, most relevant to study (Landersø and Heckman 2016). We use student reading test scores as the ultimate outcome of the treatment but also include teacher perceptions of student abilities. Table 2 shows an overview of the papers' research design.

Table 2. Overview of the papers' research design

Short Title	Social category	Setting	Design	Data/sample	Outcome
Barriers	Ethnicity	Public Schools	Survey experiment	A sample of Danish school principals, survey data	School managers' assessment of potential teacher
Performance	Ethnicity	Public Schools	Field experiment and cross-section	Administrative data, job vacancies/callback data	Callback for job interview
Gender match	Gender	Welfare service	Quasi-experiment/fixed effects estimation	Administrative data, bureaucrat survey, client survey	Educational enrollment, bureaucrat time use and citizen effort
Discrimination	Ethnicity	Public Schools	Survey- and field experiment	3 samples of teachers, survey data	Teachers' willingness to include a student in their classroom
Misperceptions	Social class/ socio-economic status	Public Schools	Field experiment	Administrative data on students, (teacher) survey data	Student reading test scores and teachers' misperception of student abilities

Case selection

I apply these designs in the educational sector and particular in a primary and lower secondary school setting. This setting has several advantages. Public schools represent one of the largest public employers and accounts for significant levels of public spending across countries. Correspondingly, a huge proportion of the population enrolls in public schools (80 percent of all children in Denmark, but also in the U.S.). Early encounters with discrimination may have detrimental long-term effects which could be particularly valuable to bring to light. In addition, educational success might have important consequences for life chances. Further, relatively large ethnic, gender and social class diversity exists within the school setting. Finally, schools are comparable organizations, with outcomes such as test results that are clearly measureable.

As an exception *Gender Match* builds on data from a mentor program. However, the program focused on young unemployed jobseekers without vocational qualifications and the goal of success was to get the jobseekers enrolled in an educational program. For this reason, several of the same arguments apply to this setting – moving young jobseekers in the direction of being able to support themselves in the long term can affect life chances substantially. In addition, this setting provides several other advantages. In particular, it allows us to assess the impact of social (gender) distance in a credible way and allows a test of one of the central elements of the dissertation’s argument in a different setting from the rest of the papers. Again, as mentioned in the introduction, any focus of scope implies limitations which I will discuss further in the final chapter of the summary.

Data, sample selection, and recruitment

I used several different strategies to select and recruit the participants in the dissertation’s papers. In *Barriers*, I simply invited all Danish public school managers to participate in the survey in which the survey experiment was embedded. While the experimental setup strengthened the internal validity of the study, the response rate was approximately 17 % which might cause for concern in relation to the external validity. The school managers who decided to participate might differ in systematic ways from those that did not, however, the participants did not differ from the population of Danish public schools on a number of school characteristics. In *Performance*, the job vacancies were focused in two regions of Denmark – Central Jutland and the Capital region – to ensure that the applications were not rejected because the applicant lived too far away from the employer. All job vacancies at public schools between

February and July 2015 in these regions were initially part of the sample. However, in the analysis, performance data is only available for schools with grade 9 students. While which school had job vacancies that semester may not be truly random and this might threaten the external validity, we show that the schools in our sample resemble the performance of other schools in these regions quite well. The data for *Gender match* was collected as part of the evaluation of a large scale field experiment. The participating job seekers were randomly assigned to either a mentor or the regular unemployment program in the job centers to which the young job seekers belonged. In the analysis, I focus on the job seekers assigned to a mentor in eight job centers. These job centers were widely spread across Denmark. In *Discrimination* we use three different samples of teachers in the four experiments. We recruited these teachers from Danish public schools across the country. The school principal or the administrative head of the school service made the signup decision. In the first two samples, they signed up to participate in two large scale field experiments not related to our study. We embedded the three survey experiments in the pre-surveys for these larger field experiments. In the third sample, we also recruited the schools to participate in a large scale experiment and embedded a survey experiment in a post treatment teacher survey. Across the three samples all Danish public schools were initially invited to participate. However, in the last two samples it was a requirement that the schools had a substantial number of bilingual students at the school since one purpose of the field experiments was to improve learning outcomes for this selected group. Finally, we recruited the participants in *Misperceptions* as part of the same field experiment as in the last experiment in *Discrimination*, only in this paper we focused on the participating students and on a limited part of the larger study. The results follow in the next chapter.

Chapter 4: Results

When are social categories used discriminatorily?

This section summarizes the findings across the papers in the dissertation regarding discriminatory use of social categories. In *Performance*, using a field experiment, we show that when applying for teacher job vacancies in Danish primary and secondary schools, minority applicants receive substantially fewer callbacks than majority applicants. On average, an applicant with a Danish sounding name is approximately 60 % more likely to receive a callback for a job interview than an applicant with an ethnic minority sounding name (Vil-ladsen and Wulff 2017). This shows that the approach reveals discrimination in this setting. In relation to the overall argument of the dissertation, management staff in Danish public schools are primarily of Danish descent, thus the salience of social distance in terms of ethnicity is plausibly at play. In addition, while we do not have the data to test the dissertation's argument directly, we theoretically argue for a relationship between organizational performance and both lack of cognitive resources and limited attention to discrimination. More specifically, we argue that low performing schools discriminate substantially more than high performing schools because low performing organizations might encounter higher demands for change by both policy makers and citizens to improve organizational processes. This puts them under more pressure which reduces the available cognitive resources, therefore amplifying discriminatory use of social categories. We also argue that performance might correlate with attention levels: performance below expected levels diverts attention in the organization towards well-known methods; as opposed to when an organization is doing well and is more open to new inputs. Though we completely acknowledge that performance is not a random organizational characteristic and that we therefore cannot give the relationship between performance and discrimination a causal interpretation, our findings supports this argument; we find that organizations with poor performance, measured as school grade point average at the students' final exam, discriminate more (a Danish applicant is almost 80 % more likely to receive a callback than a minority applicant), whereas the discrimination among top performing schools is indistinguishable from no discrimination.

In *Barriers*, I also examine how ethnic categories affect school managers' assessment of applicants. I argue that discrimination caused by distaste or statistical discrimination should also prove manifest in the survey setup. On the contrary, if implicit discrimination influences the school managers, it might not turn up in this setup since the decision-making context differs from the real world callback decision context. In line with previous research, I argue that going through a stack of applications (Bertrand, Chugh, and Mullainathan 2005) might differ substantially from assessing one applicant in a survey experimental setup. Again, though I do not directly test this, I argue that the cognitive resources might be less depleted in the latter case. I also argue that the social distance between the teacher and the students they are supposed to teach might affect the school managers' assessment of an applicant. I find that in the survey experiment the school managers on average do not assess minority applicants less favorably than majority applicants regardless of the social distance to the students. Thereby the findings suggest that under the right circumstances, school managers do not on average discriminate in their assessments. However, on a side note, as an exemption, the least experienced managers do react to the cues. Discussion of the relevance of this follows in the final section of the dissertation.

Gender match provides evidence regarding the role of social distance. In this paper, I examine the effect of on gender a quasi-random matching of mentors and young job seekers without a vocational education. As a direct measure of how the mentors treat jobseekers of their own gender compared to the opposite gender, the mentors in this program registered how much time they spend on each job seeker. Since the mentors had no formal possibilities of granting special treatment to job seekers, I argue that the prioritization of more time on specific job seekers captures the bureaucrat's potential differential treatment well. The mentors were required to stay in touch both by email and phone and keep regular (weekly) meetings with the jobseekers. If the social distance argument is correct, the mentors should spend more time on job seekers of their own gender and in particular in settings where the gender category is most salient. I argue that gender should be most salient and affect the interaction most in the personal meetings. This is exactly what I find in the paper. While the findings do not clarify whether these effects are caused by more neutral and less discriminatory treatment by same gender mentors or by favoritism of in-group mentees, the results suggest that the social distance in the studied context lead to differential treatment.

Discrimination adds additional evidence concerning the role of workload. In a survey experiment, we ask a sample of teachers whether they think it would be wise to accommodate a problematic student in their classroom and randomly assign teachers to ethnic minority or majority named students. We

find that the teachers are less willing to accommodate the student in their classroom if the student has a minority sounding name. However, in another survey experiment with the same sample of teachers, we ask the teachers about accommodating more problematic students at the school in general and again randomize whether the example of a problematic student has a majority or a minority sounding name. We find that the teachers do not react to the minority cue when the question is about the school in general. This suggests a relationship between the workload implications for the teacher and whether they are susceptible to the ethnic status of the student.

Misperceptions provides another piece of evidence regarding the role of attention. In the study, we asked a sample of 4th and 5th grade teachers to rank the reading skills of all their students. We made a similar ranking based on standardized reading tests taken by the students in advance of the teachers' ranking. Importantly, the teachers had access to this data before we asked them to rank their students. We find that the teachers ranking correlates more with the standardized test ranking for high-SES students compared to low-SES students. We measure socioeconomic status as whether at least one of the parents of the student has a college degree (vocational or academic). This suggests that the teachers are less aware of low-SES students' abilities because their perceptions are less accurate. The consequence is a reduced ability to correctly tailor the instruction to the low-SES students' needs. As a side note, the most experienced teachers misperceive the students' skills to a lesser extent. Again, discussion of the relevance of this follows in the final section of the dissertation.

In sum, the papers show that the social categories gender, ethnicity and socioeconomic status affect how citizens (students and job seekers) are treated by public employees (mentors, teachers and school managers). *Gender match* shows how the social distance between mentors and jobseekers influences the degree of interaction. *Discrimination, Performance and Barriers* show that the use of social categories varies with contextual and individual factors linked to lack of cognitive resources and limited attention. And finally, *Misperceptions* shows that inattention to social categories not only biases perceptions of citizens – it might also affect the accuracy of these perceptions. As stated in the introduction the dissertation's papers test implications of the theoretical model rather than the individual elements directly. While the model might simplify the real-world mechanisms, I would argue that the model's relevance ultimately depends on the model's ability to explain by which means discriminatory use of social categories can be reduced. The next section summarizes the support for this across papers.

How can discriminatory use of social categories be reduced?

In this section I present the evidence across papers concerning how discriminatory use of social categories can be reduced. As mentioned above, in *Gender match* I show that the mentors in a program for young job seekers with no vocational education spend more time on job seekers of their own gender. On the one hand, this demonstrates a differential treatment by the mentors that, one could argue, should be eliminated. On the other hand however, in the paper I further examine the effects of gender match between mentor and job seeker on the job seekers' effort to get enrolled in an educational program and ultimately their probability of getting enrolled in such a program within the first 6 months of their initial meeting with a mentor. The gender matched job seekers are 5.6 %-points more likely to get enrolled. At first glance, this seems like a rather small effect, however, when we take into account the enrollment rate among the non-matched jobseekers, it corresponds to more than a 50 % increase in enrollment rate. The results also show that the job seekers assigned to a mentor of the same gender indeed display a greater effort. In addition, the difference in time use appears to have an insignificant effect on the probability of enrollment whereas the citizens' effort seems to account for some of the positive effect. This suggests that when the social distance is lower (when a citizen interacts with a mentor of their own gender) it improves the outcome for the individual citizen. Building on existing literature on representative bureaucracy, I argue that better understanding and/or communication between same gender mentors and job seekers might account for this finding. This implies that lowering the social distance between public employees and citizens by making the public employees more representative of the general population or more deliberately pairing public employees with specific characteristics together with citizens with the same characteristics might indeed reduce discriminatory behavior and improve citizen outcomes. Alternatively, the results suggest that training public employees to improve their ability to understand and communicate with citizens with different characteristics, thereby reducing the social distance, could also be effective.

In *Discrimination*, we argue that since the use of social categories in the survey experiments described above suggests that workload implications affect discriminatory use of the category ethnicity, reducing employee workload might reduce discriminatory use. In support of this argument we conduct a third survey experiment that shows that the teachers are more willing to include a problematic minority student in their classroom if additional resources follow with the student. More significantly and in line with the overall argu-

ment of the dissertation, we argue that reducing workload and thereby increasing the individual teacher's available cognitive resources can reduce discriminatory attitudes towards ethnic minority students. We capitalize on a large-scale field experiment that randomly assigns teachers to an effective workload reduction. We operationalize a reduced workload for teachers by assigning more preparation time and more time with students, but with the same objectives and curriculum of the class. We show that the teachers assigned to reduced workload do not react to a minority cue concerning introducing a problematic student into their classroom after the intervention period, while a bias against the minority student persists in the control group. Since this final survey experiment is conducted after the intervention period and the reduced workload period is terminated, we interpret this result as caused by the reduced cognitive stress that follows from the reduction in workload.

In *Misperceptions*, we argue that the misperceptions of students with low socioeconomic background mentioned above might partially explain why these students systematically lag behind the other students in terms of reading abilities. To support this, we designed an intervention with the purpose of improving the teachers' ability to assess their students' needs. We created a simple reading test that the teachers should conduct once a month for four months. We coupled this with the implementation of a language comprehension course and additional instruction time. We show that the intervention reduced the teachers' misperceptions of low-SES students. We also show that the teachers least used to using tests in their instruction have the highest reduction of such misperceptions. The intervention also ultimately increases the reading abilities of low-SES students measured with standardized computer-based tests (Nationale tests)¹ and we show that this effect is primarily driven by the students that were previously misperceived by their teacher. The treatment improves the reading skills of low-SES students by 0.22 standard deviation. Initially the difference between low and high-SES students was 0.58 standard deviation. The improvement is thereby equal to 38 % percent of this achievement gap. These results suggest that providing systematic performance information to public employees that increases attention to students' needs and abilities can reduce this type of discriminatory use of social categories.

¹ See (Beuchert and Nandrup 2014) for a detailed account of the Danish National Test. Though these tests are low-stakes tests for the students, they account for approximately half of the variation in the students' final exam scores in the corresponding subject.

In sum, while the elements in the theoretical model are also not directly tested in relation to these results, the findings support the overall argument of the dissertation. Reducing social distance by matching public employees with citizens from the same social category, releasing cognitive resources by reducing workload, and focusing attention on social categories by providing systematic performance information reduces discriminatory use of social categories. In the following sections I conclude and discuss the findings.

Chapter 5: Conclusion

This dissertation aimed to answer the question of when social categories are applied and how to reduce their discriminatory use in the public sector: an important question if we want to combat discrimination by politically controllable means. We can derive at least two further arguments for the detrimental effects of such discrimination based on the dissertation's argumentation and findings. First, taking the findings from the representative bureaucracy theory into account, public sector recruitment discrimination might amplify possible discrimination against under-represented groups of citizens. Second, limited attention to the influence of social categories in the school setting might attribute to the reproduction of achievement differences between social groups.

The dissertation's overall answer to this question is derived from a theoretical model based on the findings across the dissertation's papers and related social psychological literature on social categories and stereotypes. The model suggests that the salience of social distance affects discriminatory use of social categories. Lack of cognitive resources amplifies use because it limits attention to the potential for social categories to be used in stereotypical and ultimately discriminatory ways. Correspondingly, reducing social distance, releasing cognitive resources or focusing attention on potential discriminatory use of social categories might reduce such use. More specifically, the dissertation suggests that building on public administration theories of representative bureaucracy, street level bureaucrat coping and performance information, that reducing social distance, making the bureaucracy more representative of the citizens they serve, releasing cognitive resources by reducing workload or focusing attention on discriminatory use by providing systematic information that increases awareness of discriminatory influence of social categories can reduce discriminatory use.

The results show that public employees (school managers) react to the ethnicity of a potential employee (teacher) to a greater extent when organizational performance is low. This is likely because cognitive resources are more limited due to increased pressure from politicians and citizens to improve performance. Additionally, attention to social categories are more limited because of a focus on organizational survival and more well-known approaches. Similarly, the results show that social categories influence public employees (school managers) more in a real-world application process in which they have to look through a stack of applications and cognitive resources might thereby be more depleted than in a survey experimental context where they only have

to evaluate one applicant. The results also show that gender match as an operationalization of social distance has an impact on mentors' time use on specific mentees, that workload implications affect teachers' willingness to include a minority student in their classroom, and that teachers' have systematically less accurate perceptions of low-SES students' reading abilities than of high-SES students.

Also in relation to the reduction of discriminatory use, the results show that reducing social distance by matching mentors and job seekers on gender effectively improves citizen outcomes, reducing workload by giving teachers more time to cover the same curriculum reduces differences in willingness to include minority and non-minority students in a class, and providing teachers with monthly test information on their students' skills and progression reduces differential perceptions of high- and low-SES students and ultimately the return to schooling for low-SES students. In the following section I discuss the limitations as well as the practical and political implications of the dissertation. Finally, I discuss the theoretical implications and questions for future research.

Chapter 6: Discussion

Limitations

While the results support the implications of the general model in the dissertation, a number of limitations need discussion. Also, as mentioned in the introduction, it is difficult to test the implications of all elements in the model in one paper, however I argue that the papers supplement each other and that the evidence across papers points in the same direction. Similarly, I argue that some shortcomings of the individual papers are addressed in other papers in the dissertation.

Particularly in the relation to the generalizability of the results, limitations need discussion. First, the main type of service considered, public schooling, might differ from other types of service in important ways. Teachers deliver a large portion of their service as classroom based instruction and thereby as one deliverer to many receivers of public service. This might make the use of social categories particularly relevant since it is costly to gather detailed information about each individual student. This might prove different from the work of for instance, case or health workers. However, *Gender match* shows that social categories also matter in a service delivered on a one-on-one basis.

Second, the decision-making context in which the papers identify differential treatment also exemplify situations to a great extent in which the public employees only have brief interaction or limited time to evaluate the citizen in question. For example, a teacher might react differently towards a minority student she sees every day in class compared to the scenarios tested in *Discrimination* where the case of a possible new student is subject to discussion. Similarly, inviting an applicant to an interview only based on application and CV might lead to different results than a study of public managers' treatment of employees within an organization. In fact, several studies argue that public employees such as street level bureaucrats' use of categorization is strongly tied to situations with a requirement for quick actions (Maynard-Moody and Musheno 2003; Dias and Maynard-Moody 2007). However, the *Misperception* paper shows that social categories also influence how well-informed teachers are about their own students with whom they interact on a daily basis. Similarly, *Gender match* shows the relevance of social categories in a service based on long-term interaction. Also, many public employees only interact briefly with the citizens they serve. Police officers, for instance, when they

decide to stop and search a vehicle, and similarly many health workers have only very brief interactions with the same patient.

Finally, all studies were conducted in a Danish context. Denmark might look like a rather unique setting in many dimensions. Particularly, it is a small and homogeneous country with a welfare state based on principles of universalism (Esping-Andersen 1990). However, the provision of public schooling is widespread across different countries (e.g. the U.S.). This strengthens the relevance of the results to other countries. Also, similar name cue effects as those identified in several of the dissertation's papers have been found in numerous U.S. studies (e.g. Bertrand and Mullainathan 2004; Schram et al. 2009; Butler and Broockman 2011; DeSante 2013; White, Nathan, and Faller 2015; Einstein and Glick 2017) as well as Scandinavian studies (e.g. Midtbøen 2016; Carlsson and Rooth 2007) which also speaks to a more universal relevance of this type of discriminatory use of social categories.

Another set of limitations relates to the design of the papers. In *Performance*, while we argue that performance relates to discrimination for reasons linked to cognitive resources and attention, we do not directly test this. Further, organizational performance is not plausibly a random characteristic of an organization. Thereby we cannot assert that it is performance as such that affects levels of discrimination. We handle this by taking alternative explanations for the relationship into our theoretical argumentation. In addition, we use a value-added measure of performance to examine how much of the relationship relating to absolute performance might also reflect other organizational characteristics such as the quality of input factors, compared to aspects more related to management quality. Related to this, in *Barriers* I do not directly test that the reason for the missing effect of ethnicity in the survey experiment relates to cognitive resources, nevertheless, I would argue that *Discrimination* shows the relevance of this in a more convincing way. A limitation in *Discrimination* on the other hand, is that all outcomes are survey based, thus we do not know whether the attitudes transform into behavior. *Misperception* however, shows that social categories also affect real world perceptions of students. Here a shortcoming is that we cannot experimentally separate the effects of the tests in our treatment from other elements in the treatment. Nevertheless, that the teachers least familiar with using data in their instruction update their perceptions of the students the most, and that the students previously misperceived by their teacher experience the greatest increase in reading scores supports the notion that the effects can be attributed to a substantial degree to the information part of the treatment. Finally, in *Gender match*, while the combination of as good as random matching with fixed effects takes many potential alternative explanations for the results into

account, selection is still possible. However, I show that the matched and unmatched jobseekers resemble each other on important observable characteristics such as unemployment and educational enrollment history, which supports that we indeed can consider the matching as good as random.

Political and practical implications

The findings have several implications for contemporary political debates. First, in a Danish context, teachers, researchers and politicians criticize the increasing number of tests implemented in Denmark. They particularly criticize that testing creates a feeling of failure among poor performing students. However, the results from *Misperceptions* show that testing also improves both teachers' perceptions of the students' needs and ultimately the return to schooling for students with low socioeconomic background. In addition, we actually show that the treatment does not increase the students' behavioral problems. On the contrary, we show that low-SES students in the treatment group report fewer behavioral problems than similar students in the control group. More focus on this learning aspect of testing could inform the ongoing debate. While an upper limit to how much testing helps might exist, within the sample of teachers that participated in the experiment two thirds used academic tests 0-4 times a year. The intervention involved monthly tests and thereby increased the use of tests substantially for a majority of teachers. In an international comparison, the use of tests in the Danish school system is rather low. The 2016 PISA report shows that only 33 % of Danish school managers report that their students are tested once a month or more frequently (OECD 2016, 134). By comparison, the OECD average is 65 % while the number in the U.S. is 90 % (ibid.). We might therefore be far from the upper limit in the Danish school system. These results also have implications regarding the use of performance information to improve performance. As mentioned, studies of the use of systematic performance information coupled with financial incentives show strategic and perverse responses to such systems (Jacob and Levitt 2003; Figlio and Loeb 2011). The results in *Misperceptions* suggest that low stakes tests in a system without financial incentives also might improve learning and ultimately performance.

Second, public employee workload is on the increase in many countries – particularly due to budgetary cuts following the financial crisis (Vaughan-Whitehead 2013). The findings suggest that this might lead to more unequal service delivery. However, since perceptions of workload and stress are subjective, future research must determine the long-term effects of this trend. In addition, I recognize that resources are notoriously scarce in the public sector. Suggesting a general decrease of the workload of all public employees is an

expensive way to decrease differential treatment. Whether it is worth the cost is in the end a political question and other methods of improving equality in the interaction between citizens and public employees might prove more cost-efficient in the long run.

Third, *Gender Match* suggests that it might be feasible to deliberately seek to match street level bureaucrats with citizens belonging to the same social categories – at least on gender. However, in practice this might prove difficult for several reasons. First, in many settings, such as the school setting, street level bureaucrats interact with several citizens at once. To match citizens and bureaucrats in such a setting demands that you make classrooms only for boys or girls. Further matching on other characteristics such as ethnicity would result in an extremely segregated school system. In such settings, it might make more sense to learn how to train bureaucrats to represent and communicate better with all citizens. Second, quite substantial segregation in terms of ethnicity and gender particularly in the job market makes it difficult in practice to make the bureaucracy more representative. The ratio between male and female nurses, for instance, is highly skewed and the same is true for the ratio between ethnic minority and majority teachers (Gjerberg and Kjølrsrød 2001; Villadsen and Wulff 2017). Third, and relatedly, increased emphasis on differences between social categories might reproduce or even amplify their importance.

Fourth, the *Barriers* paper shows that school managers are perfectly capable of making neutral judgements of teachers with different ethnic minority status regardless of the social distance to the target group. First, this implies that the findings from representative bureaucracy theory concerning the benefits of having a bureaucracy that mirrors the target group are not necessarily apparent for practitioners. Second, this implies that the usual recruitment process might contribute to producing differential treatment of different ethnic groups. While it might be costly, a review process that results in employers more deliberately considering each applicant might reduce the callback difference identified in *Performance*. Interestingly in *Performance*, we also suggest that good performance reduces discriminatory behavior. While we cannot determine with certainty that performance reduces discrimination as such, the results suggest that bad performance goes hand-in-hand with substantial discrimination.

Finally, across the papers the findings indicate that there might be an important trade-off between different goals for the delivery of public services. First, use of social categories as cues that enable fast and frugal responses to often complex real world circumstances might be a quite effective strategy for achieving the goal of getting the most out of scarce public resources. Particu-

larly concerning the discrimination found in *Misperceptions*, where the teachers do not in fact assess low-SES students systematically higher or lower than their actual level of abilities, but systematically less accurately. Second, this might conflict with the goal of equal treatment of all citizens belonging to different social categories. Finally, this might also conflict with a third goal related to equality in the outcomes of public service delivery. There might obviously be a trade-off between the three, and how to weight these goals is ultimately a political question. However, pursuing the individual goals might also have different long- and short-term effects. For instance, investments in more equal treatment of children or in more equal outcomes might be expensive in the short-term, but might also be an effective strategy in the long-term if such improved human capital investments translate into economic growth and increased tax revenues.

Theoretical implications and future research

This dissertation leaves at least four large and a number of smaller (but not less important) unanswered questions for future research. First, as I touch upon in the previous section, making the bureaucracy more representative or deliberately matching bureaucrats and citizens based on social categories might not always be feasible or may produce other undesirable side effects. While I argue that the results concerning social distance call for training of all bureaucrats to better serve all citizens, how this can be done is a difficult and unresolved question.

Another unresolved question relates to the social psychological literature. The literature suggests that individuals can abstain from applying stereotypes if they are appropriately motivated and have sufficient cognitive resources (Fiske 1998). While the cognitive resources receive considerable attention in this dissertation, the appropriate motivation needs more attention in relation to public employees' discriminatory use of social categories. This approach might link well to the existing public administration literature about public employees' agency roles (e.g. Wenger and Wilkins 2009; Maynard-Moody and Musheno 2000). Wenger and Wilkins (2009) describe three different perspectives on public employees' (specifically street level bureaucrats') motivation; 1) motivated to meet the clients' needs (citizen agents), 2) motivated to comply with administrative rules and hierarchy (state agents), and 3) motivated to punish unworthy clients (rogue agents). Determining which street level bureaucrats follow each of these motives, when they follow these motives and how to change these motives might be another way of reducing discriminatory use of social categories. Similarly, some of the existing findings con-

cerning discrimination do not support that social distance affects public employees' discriminatory use of social categories. For instance minority case workers discriminate just as much against minority citizens as majority case workers (Schram et al. 2009, 415). Similarly, housing officials do not seem more responsive to citizens of their own race (Einstein and Glick 2017, 109). Differences in motives might also explain these findings. The question might also be linked to research concerning public employees more general motivation; intrinsic, extrinsic and/or pro-social.

Third, the results concerning attention suggest that differential treatment can work in subtle ways. The most obvious examples relate to differential treatment of citizens that are equal in all other aspects other than belonging to a specific social category. However, the results in *Misperceptions* suggest a more subtle type of discrimination related to public employees' different levels of awareness of citizens from different social categories. While the consequence of differential awareness might not affect whether public employees on average treat citizens belonging to specific social categories in accordance with their qualifications and needs, for the individual, imprecise perceptions of these qualifications or needs compared to individuals belonging to other social categories leads to treatment that matches these qualifications or needs to a lesser extent. E.g. when teachers are less aware of the abilities of low-SES students compared to high-SES students, this might result in less well-tailored instruction to the first group of students. Why the differences in misperceptions are there in the first place is an unresolved question. One hypothesis is that it is because the teachers build their assessment of the student on other cues that correlate poorly with academic skills for low-SES students such the students' behavior in relation to homework, absenteeism, and/or classroom activity – whether this hypothesis or alternative ones provide the answer must be addressed in future research. Furthermore, how the goal of improving the teachers' perceptions of the students' skills can most effectively be achieved is unclear. Would it for instance be just as effective to simply tell the teachers for which students they misperceived the skills? Also, an interesting result outside the main argument of the dissertation is that the less experienced school managers and teachers respectively in *Barriers* and in *Misperceptions* respond to a greater extent to social categories or misperceive the students' skills. This suggests that experience might be another factor that reduces discriminatory use of social categories. These findings are in line with a simple learning model based on past experience. However, the test results in *Misperceptions* were equally available for all teachers. Future research must uncover when this type of learning reduces discriminatory use of social categories and when we need to consider other needs.

Fourth, based on the discussion of the generalization of the results also leaves room for specific hypotheses for exploration in future research, particularly concerning other implications of limited attention. The type of interaction (one-on-one vs. many-to-one) and the length of the interaction (short-term vs. long-term) might influence how much attention public employees can allocate to the individual citizen and thereby the use of social categories in discriminatory ways. Based on the discussion in the former section, we could expect that instances of one public employee related to many citizens for a short time span will be most influenced by social categories. The callback situation resembles this scenario quite well. The influence might prove smaller, but still large when the interaction is one-on-one but only for a short period of time. This is how many health workers and police officers interact with citizens, for instance. The effects might be smaller when the public employee interacts with the same group of citizens for a longer time period, in the way that teachers do, and even smaller for public employees that interact one-on-one with the same citizen for long time periods, such as in the mentor program described above. These hypotheses obviously only make sense when everything else is kept constant. The elements identified as crucial to the level of discriminatory use of social categories in this dissertation might change the expected effects considerably. Again, whether these hypotheses hold remains an empirical question for future research. Table 3 summarizes the hypotheses for future research:

Table 3. Hypotheses regarding the type of service and discriminatory use of social categories

		Type of interaction	
		One-to-one	One-to-many
Length of interaction	Short	Medium high	High
	Long	Low	Medium low

Along the same line of reasoning, these factors might also influence how we choose to organize the delivery of public services within the same type of service. Thus, all else being equal, this might imply that increased class size, lower ratio between daycare personnel and children, and higher caseloads for case workers might all lead to more discriminatory use of social categories. Again, whether these factors indeed affect discriminatory use of social categories and

if so, how large these effects are is ultimately an empirical question for future research.

Finally, the dissertation leaves a number of smaller but still highly important questions unanswered. First, the relevance of social categories other than gender, ethnicity and social class needs more research. For instance, age in relation to hiring and firing within public organizations. Second, while I argue that the setting of my studies has advantages concerning generalizability, how well the dissertation's argument travel across sectors and countries is ultimately an empirical question for future research. Finally, possible intersectionality between the different social categories needs more attention in future research: how combinations of social categories affects their use for instance.

In sum this dissertation takes some of the first steps towards a public administration theory of discriminatory use of social categories in the public sector and how to reduce such use by politically controllable means. The quest for future research is to find more and sustainable political means to reduce discriminatory use of social categories.

Literature

- Allport, Gordon W. [1954] 1979. *The Nature of Prejudice: 25th Anniversary Edition*. Unabridged edition. Reading, Mass: Basic Books.
- Andersen, Simon Calmar, Maria Knoth Humlum, and Anne Brink Nandrup. 2016. "Increasing Instruction Time in School Does Increase Learning." *Proceedings of the National Academy of Sciences* 113 (27):7481–84. <https://doi.org/10.1073/pnas.1516686113>.
- Andrews, Rhys, Rachel Ashworth, and Kenneth J. Meier. 2014. "Representative Bureaucracy and Fire Service Performance." *International Public Management Journal* 17 (1):1–24.
- Atkins, Danielle N., Angela R. Fertig, and Vicky M. Wilkins. 2014. "Connectedness and Expectations: How Minority Teachers Can Improve Educational Outcomes for Minority Students." *Public Management Review* 16 (4):503–26. <https://doi.org/10.1080/14719037.2013.841981>.
- Atkins, Danielle N., and Vicky M. Wilkins. 2013. "Going Beyond Reading, Writing, and Arithmetic: The Effects of Teacher Representation on Teen Pregnancy Rates." *Journal of Public Administration Research and Theory* 23 (4):771–90. <https://doi.org/10.1093/jopart/muto01>.
- Becker, Gary Stanley. 1957. *The Economics of Discrimination*. Chicago, Illinois: University of Chicago Press.
- Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit Discrimination." *The American Economic Review* 95 (2):94–98.
- Bertrand, Marianne, and Esther Duflo. 2016. "Field Experiments on Discrimination." Working Paper 22014. National Bureau of Economic Research. <https://doi.org/10.3386/w22014>.
- 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. <https://doi.org/10.1257/0002828042002561>.
- Beuchert, Louise Voldby, and Anne Brink Nandrup. 2014. "The Danish National Tests – A Practical Guide." 2014–25. Economics Working Papers. Department of Economics and Business Economics, Aarhus University. <https://ideas.repec.org/p/aah/aarhec/2014-25.html>.
- Boyne, George A., Rhys Andrews, and Richard Walker. 2006. "Subjective and Objective Measures of Organizational Performance: An Empirical Exploration." In *Public Service Performance: Perspectives on Measurement and Management*, edited by George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole Jr, and Richard Walker. New York: Cambridge University Press.
- Bradbury, Mark, and J. Edward Kellough. 2011. "Representative Bureaucracy: Assessing the Evidence on Active Representation." *The American Review of Public Administration* 41 (2):157–67. <https://doi.org/10.1177/0275074010367823>.
- Bryson, John M., Frances S. Berry, and Kaifeng Yang. 2010. "The State of Public

- Strategic Management Research: A Selective Literature Review and Set of Future Directions.” *The American Review of Public Administration* 40 (5):495–521. <https://doi.org/10.1177/0275074010370361>.
- 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. <https://doi.org/10.1111/j.1540-5907.2011.00515.x>.
- Carlsson, Magnus, and Dan-Olof Rooth. 2007. “Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data.” *Labour Economics*, European Association of Labour Economists 18th annual conference CERGE-EI, Prague, Czech Republic 21-23 September 2006, 14 (4):716–29. <https://doi.org/10.1016/j.labeco.2007.05.001>.
- Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star.” *The Quarterly Journal of Economics* 126 (4):1593–1660. <https://doi.org/10.1093/qje/qjr041>.
- 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. <https://doi.org/10.1093/jopart/mus020>.
- Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood, New Jersey: Prentice-Hall, Inc.
- Dee, Thomas S. 2004. “Teachers, Race, and Student Achievement in a Randomized Experiment.” *The Review of Economics and Statistics* 86 (1):195–210. <https://doi.org/10.1162/003465304323023750>.
- 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. <https://doi.org/10.1111/ajps.12006>.
- Devine, Patricia G. 1989. “Stereotypes and Prejudice: Their Automatic and Controlled Components.” *Journal of Personality and Social Psychology* 56 (1):5–18.
- Dias, Janice Johnson, and Steven Maynard-Moody. 2007. “For-Profit Welfare: Contracts, Conflicts, and the Performance Paradox.” *Journal of Public Administration Research and Theory* 17 (2):189–211. <https://doi.org/10.1093/jopart/mul002>.
- Dieterle, Steven, Cassandra M. Guarino, Mark D. Reckase, and Jeffrey M. Wooldridge. 2015. “How Do Principals Assign Students to Teachers? Finding Evidence in Administrative Data and the Implications for Value Added.” *Journal of Policy Analysis and Management* 34 (1):32–58. <https://doi.org/10.1002/pam.21781>.
- Dovidio, John F. 1986. *Prejudice, Discrimination, and Racism*. Edited by Samuel L. Gaertner. Orlando: Academic Pr.
- Einstein, Katherine Levine, 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):100–116. <https://doi.org/10.1111/ajps.12252>.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, N.J: Princeton University Press.
- Figlio, David, and Susanna Loeb. 2011. “School Accountability.” In *Handbooks in Economics*, edited by Eric A. Hanushek, Stephen J. Machin, and Ludger Woessmann. The Netherlands: North-Holland: Elsevier.
- Figlio, David N. 1999. “Functional Form and the Estimated Effects of School Resources.” *Economics of Education Review* 18 (2):241–52. [https://doi.org/10.1016/S0272-7757\(98\)00047-8](https://doi.org/10.1016/S0272-7757(98)00047-8).
- Fiske, Susan T. 1998. “Stereotyping, Prejudice, and Discrimination.” In *The Handbook of Social Psychology*, edited by Susan T Fiske, Daniel T. Gilbert, and Lindzey Gardner, 4th edition, 357–411. New York, NY: McGraw-Hill.
- Gade, Daniel M., and Vicky M. Wilkins. 2013. “Where Did You Serve? Veteran Identity, Representative Bureaucracy, and Vocational Rehabilitation.” *Journal of Public Administration Research and Theory* 23 (2):267–88. <https://doi.org/10.1093/jopart/mus030>.
- Gigerenzer, Gerd, and Wolfgang Gaissmaier. 2011. “Heuristic Decision Making.” *Annual Review of Psychology* 62 (1):451–82. <https://doi.org/10.1146/annurev-psych-120709-145346>.
- Gilbert, Daniel, and J. Gregory Hixon. 1991. “The Trouble of Thinking: Activation and Application of Stereotypic Beliefs.” *Journal of Personality and Social Psychology* 60 (April):509–17. <https://doi.org/10.1037/0022-3514.60.4.509>.
- Gjerberg, Elisabeth, and Lise Kjølørød. 2001. “The Doctor–nurse Relationship: How Easy Is It to Be a Female Doctor Co-Operating with a Female Nurse?” *Social Science & Medicine, The Physicians Role in Transition*, 52 (2):189–202. [https://doi.org/10.1016/S0277-9536\(00\)00219-7](https://doi.org/10.1016/S0277-9536(00)00219-7).
- Govorun, Olesya, and B. Keith Payne. 2006. “Ego—Depletion and Prejudice: Separating Automatic and Controlled Components.” *Social Cognition* 24 (2):111–36. <https://doi.org/10.1521/soco.2006.24.2.111>.
- Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. “Learning Through Noticing: Theory and Evidence from a Field Experiment.” *The Quarterly Journal of Economics* 129 (3):1311–53. <https://doi.org/10.1093/qje/qju015>.
- Hardin, C., and M. R. Banaji. 2013. “The Nature of Implicit Prejudice.” In *The Behavioral Foundations of Public Policy*, edited by Eldar Shafir. Princeton University Press Princeton, CA. <https://www.google.com/books?hl=da&lr=&id=tyWqtKjyx3cC&oi=fnd&pg=PA13&dq=hardin+and+banaji+the+nature+of+implicit+prejudice&ots=VQM8R4fB67&sig=X92GynJbICssul51mbq2oYU-uVg>.
- Harrits, Gitte Sommer, and Marie Østergaard Møller. 2014. “Prevention at the Front Line: How Home Nurses, Pedagogues, and Teachers Transform Public Worry into Decisions on Special Efforts.” *Public Management Review* 16

- (4):447–80. <https://doi.org/10.1080/14719037.2013.841980>.
- Hjortskov, Morten, Simon Calmar Andersen, and Morten Jakobsen. Forthcoming. “Encouraging Political Voices of Underrepresented Citizens through Coproduction. Evidence from a Randomized Field Trial.” *American Journal of Political Science*.
- Hogg, Michael A. 2001. “Social Categorization, Depersonalization, and Group Behavior.” In *Blackwell Handbook of Social Psychology: Group Processes*, edited by Michael A. Hogg and R. Scott Tindale, 56–85. Blackwell Publishers Ltd. <https://doi.org/10.1002/9780470998458.ch3>.
- Jacob, Brian A., and Steven D. Levitt. 2003. “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating.” *The Quarterly Journal of Economics* 118 (3):843–77. <https://doi.org/10.1162/00335530360698441>.
- 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. <https://doi.org/10.1111/acem.12901>.
- Jussim, Lee, and Kent D. Harber. 2005. “Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies.” *Personality and Social Psychology Review* 9 (2):131–55. https://doi.org/10.1207/s15327957pspr0902_3.
- 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. <https://doi.org/10.1111/j.0092-5853.2004.00072.x>.
- Keiser, Lael R., Vicky M. Wilkins, Kenneth J. Meier, and Catherine A. Holland. 2002. “Lipstick and Logarithms: Gender, Institutional Context, and Representative Bureaucracy.” *American Political Science Review* 96 (3):553–64. <https://doi.org/10.1017/S0003055402000321>.
- Landersø, Rasmus, and James J. Heckman. 2016. “The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the U.S.” Working Paper 22465. National Bureau of Economic Research. <https://doi.org/10.3386/w22465>.
- Lau, Richard R., and David P. Redlawsk. 2001. “Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making.” *American Journal of Political Science* 45 (4):951–71. <https://doi.org/10.2307/2669334>.
- Lim, Hong-Hai. 2006. “Representative Bureaucracy: Rethinking Substantive Effects and Active Representation.” *Public Administration Review* 66 (2):193–204. <https://doi.org/10.1111/j.1540-6210.2006.00572.x>.
- Lipsky, Michael. [1980] 2010. *Street-Level Bureaucracy, 30th Ann. Ed.: Dilemmas of the Individual in Public Service*. 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. <https://doi.org/10.1080/01973533.2013.840630>.
- Maynard-Moody, Stephen Williams, and Michael Craig Musheno. 2003. *Cops, Teachers, Counselors: Stories from the Front Lines of Public Service*. University of Michigan Press.
- Maynard-Moody, Steven, and Michael Musheno. 2000. "State Agent or Citizen Agent: Two Narratives of Discretion." *Journal of Public Administration Research and Theory* 10 (2):329–58. <https://doi.org/10.1093/oxfordjournals.jpart.a024272>.
- Meier, Kenneth J., Robert D. Wrinkle, and J. L. Polinard. 1999. "Representative Bureaucracy and Distributional Equity: Addressing the Hard Question." *The Journal of Politics* 61 (4):1025–39. <https://doi.org/10.2307/2647552>.
- Meier, Kenneth John, and Lloyd G. Nigro. 1976. "Representative Bureaucracy and Policy Preferences: A Study in the Attitudes of Federal Executives." *Public Administration Review* 36 (4):458–69. <https://doi.org/10.2307/974854>.
- Midtbøen, Arnfinn H. 2016. "Discrimination of the Second Generation: Evidence from a Field Experiment in Norway." *Journal of International Migration and Integration* 17 (1):253–72. <https://doi.org/10.1007/s12134-014-0406-9>.
- Mosher, Frederick C. 1968. *Democracy and the Public Service*. New York: OUP.
- Moynihan, Donald P. 2005. "Goal-Based Learning and the Future of Performance Management." *Public Administration Review* 65 (2):203–16. <https://doi.org/10.1111/j.1540-6210.2005.00445.x>.
- Moynihan, Donald P., and Noel Landuyt. 2009. "How Do Public Organizations Learn? Bridging Cultural and Structural Perspectives." *Public Administration Review* 69 (6):1097–1105. <https://doi.org/10.1111/j.1540-6210.2009.02067.x>.
- Nicholson-Crotty, Sean, Jill Nicholson-Crotty, and Sergio Fernandez. 2017. "Performance and Management in the Public Sector: Testing a Model of Relative Risk Aversion." *Public Administration Review* 77 (4):603–14. <https://doi.org/10.1111/puar.12619>.
- Nielsen, Poul Aaes. 2014. "Learning from Performance Feedback: Performance Information, Aspiration Levels, and Managerial Priorities." *Public Administration* 92 (1):142–160.
- OECD. 2016. *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*. Paris: OECD Publishing.
- Olson, Jeremiah C. 2016. "Race and Punishment in American Prisons." *Journal of Public Administration Research and Theory* 26 (4):758–68. <https://doi.org/10.1093/jopart/muw026>.
- 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 (1):339–67. <https://doi.org/10.1146/annurev.psych.60.110707.163607>.
- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The*

- American Economic Review* 62 (4):659–61.
- Pitts, David W. 2005. “Diversity, Representation, and Performance: Evidence about Race and Ethnicity in Public Organizations.” *Journal of Public Administration Research and Theory* 15 (4):615–31. <https://doi.org/10.1093/jopart/mui033>.
- Rangvid, Beatrice Schindler. 2015. “Systematic Differences across Evaluation Schemes and Educational Choice.” *Economics of Education Review* 48 (Supplement C):41–55. <https://doi.org/10.1016/j.econedurev.2015.05.003>.
- Riccucci, Norma M., and Marcia K. Meyers. 2004. “Linking Passive and Active Representation: The Case of Frontline Workers in Welfare Agencies.” *Journal of Public Administration Research and Theory* 14 (4):585–97. <https://doi.org/10.1093/jopart/muh038>.
- Riccucci, Norma M., Van Ryzin, and Cecilia F. Lavena. 2014. “Representative Bureaucracy in Policing: Does It Increase Perceived Legitimacy?” *Journal of Public Administration Research and Theory* 24 (3):537–51. <https://doi.org/10.1093/jopart/muu006>.
- Riccucci, Norma M., Gregg G. Van Ryzin, and Huafang Li. 2016. “Representative Bureaucracy and the Willingness to Coproduce: An Experimental Study.” *Public Administration Review* 76 (1):121–30. <https://doi.org/10.1111/puar.12401>.
- Ridgeway, Cecilia L. 2014. “Why Status Matters for Inequality.” *American Sociological Review* 79 (1):1–16. <https://doi.org/10.1177/0003122413515997>.
- Rockoff, Jonah E., Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. 2012. “Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools.” *The American Economic Review*, 10–1257.
- Rooth, Dan-olof. 2010. “Automatic Associations and Discrimination in Hiring: Real World Evidence.” *Labour Economics*, 523–534.
- Rosenthal, Robert, and Lenore Jacobson. 1968. “Pygmalion in the Classroom.” *The Urban Review* 3 (1):16–20. <https://doi.org/10.1007/BF02322211>.
- Salge, Torsten O. 2011. “A Behavioral Model of Innovative Search: Evidence from Public Hospital Services.” *Journal of Public Administration Research and Theory* 21 (1):181–210. <https://doi.org/10.1093/jopart/muq017>.
- Schanzenbach, Diane Whitmore. 2006. “What Have Researchers Learned from Project STAR?” *Brookings Papers on Education Policy*, no. 9:205–28.
- Schneider, Anne, and Helen Ingram. 1993. “Social Construction of Target Populations: Implications for Politics and Policy.” *American Political Science Review* 87 (2):334–47. <https://doi.org/10.2307/2939044>.
- 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. <https://doi.org/10.1177/000312240907400304>.
- Shelton, J. Nicole, Jennifer A. Richeson, and John F. Dovidio. 2013. “Biases in Interracial Interactions.” In *The Behavioral Foundations of Public Policy*, edited by Eldar Shafir. Princeton University Press Princeton, CA. <https://www.google.com/books?hl=da&lr=&id=tyWqtkjyx3cC&oi=fnd&pg=PA>

- 13&dq=hardin+and+banaji+the+nature+of+implicit+preju-
dice&ots=VQM8R4fB67&sig=X92GynJbICssul51mbq20YU-uVg.
- Simon, Herbert A. [1947] 1997. *Administrative Behavior, 4th Edition: A Study of Decision-Making Processes in Administrative Organisations*. 4 Rev Ed edition. New York: Free Press.
- Soss, Joe. 1999. "Lessons of Welfare: Policy Design, Political Learning, and Political Action." *American Political Science Review* 93 (2):363–80. <https://doi.org/10.2307/2585401>.
- 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. <https://doi.org/10.1093/jopart/muq095>.
- Spencer, Katherine B., Amanda K. Charbonneau, and Jack Glaser. 2016. "Implicit Bias and Policing." *Social and Personality Psychology Compass* 10 (1):50–63. <https://doi.org/10.1111/spc3.12210>.
- Steele, C. M., and J. Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African Americans." *Journal of Personality and Social Psychology* 69 (5):797–811.
- 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):1099–1126. <https://doi.org/10.1093/jopart/muu056>.
- Vaughan-Whitehead, Daniel, ed. 2013. *Public Sector Shock: The Impact of Policy Retrenchment in Europe*. Cheltenham, UK ; Northampton, MA, USA : Geneva, Switzerland: International Labor Office with Edward Elgar.
- Villadsen, Anders, and Jesper Wulff. 2017. "Is the Public Sector a Fairer Employer? Ethnic Employment Discrimination in the Public and Private Sectors." *Academy of Management Discoveries*, September, amd.2016.0029. <https://doi.org/10.5465/amd.2016.0029>.
- Weber, Max. 1947. *The Theory Of Social And Economic Organization*. New York: Oxford University Press.
- 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. <https://doi.org/10.1093/jopart/mum044>.
- 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. <https://doi.org/10.1017/S0003055414000562>.

Summary

A substantial number of studies find that social categories, broadly defined as categories of people which acquire their meaning by contrast with other categories, such as gender, ethnicity and social class, affect the treatment citizens receive in the public sector. This has detrimental effects for the individual but also conflicts with the promise universalistic treatment embedded in modern bureaucracy. In addition, misperceptions of citizens' needs and abilities might lead to inefficient service delivery and it might affect citizens directly through creating different expectations and thereby self-fulfilling prophecies. Finally, (negative) experiences in direct encounters with government institutions can have severe consequences for political efficacy and participation. Considering insights from social psychology, it is not particularly surprising that public employees – like other human beings – sometimes use social categories as shortcuts when evaluating citizens. This is particularly the case because public employees often face demand that exceeds the available resources, and because studies of implicit discrimination find that discrimination might be completely unintentional and outside the discriminator's awareness. While this suggests that such behavior is partly the result of an automatic process, studies in social psychology find that people can avoid the influence of stereotypes when they are appropriately motivated and have sufficient cognitive resources. However, few studies examine how to reduce discriminatory treatment by politically controllable means. This is surprising since a rich theoretical and lab-based psychological literature provides ample suggestions concerning how this could be done. This dissertation examines when social categories are applied and importantly, how their discriminatory use in the public sector can be reduced.

To answer this question, the dissertation proposes a theoretical model that is able to explain when discriminatory use of social categories occurs and how to reduce such use. The models build on a combination of the findings from the dissertation's five separate papers that draw on public administration theories about representative bureaucracy, street level bureaucrat coping and performance information, and on insights from social psychology concerning social categories and stereotypes. More specifically, the dissertation argues that salient social distance affects discriminatory use of social categories. Lack of cognitive resources amplify such use because it limits attention to the influence of social categories in stereotypical and consequently discriminatory ways. Correspondingly, the dissertation suggests that reducing social distance, releasing cognitive resources or focusing attention on potentially discriminatory use of social categories might reduce such use. The dissertation does not

directly test the relevance of these three psychological elements, however, the model has a number of implications concerning discriminatory use of social categories in the public sector and how this can be reduced. Specifically, the dissertation suggests that reducing social distance by making the bureaucracy more representative of the citizens they serve, releasing cognitive resources by reducing workload or focusing attention on discriminatory use by providing systematic information that increases awareness of discriminatory influence of social categories can reduce discriminatory use.

The dissertation combines quasi-, field- and survey-experimental evidence and finds support for the proclaimed model. The results have direct implications in relation to reducing discriminatory use of social categories by the proposed interventions as well as several additionally important implications. First, the results suggest that not only might it be possible to reduce discriminatory use of social categories, but also ultimately, to improve citizen outcomes by deliberately matching street level bureaucrats with citizens belonging to the same social categories and consequently reducing the social distance. Second, public employee workload is on the increase in many countries – particularly due to budgetary cuts following the financial crisis. A side effect of this might be more discriminatory use of social categories. Third, the results suggest that systematic performance information provided in the format of low stakes tests in a system without financial incentives might improve learning but also ultimately reduce the gap between low and high socioeconomic status students.

In sum, this dissertation takes some of the first steps towards a comprehensive theory of discriminatory use of social categories in the public sector and how such use might be reduced by politically controllable means. The quest for future research is to find more and sustainable political means to reduce discriminatory use of social categories.

Dansk resumé

En lang række studier viser, at sociale kategorier bredt defineret som kategorier af mennesker, der opnår mening gennem deres modsætning til andre kategorier, såsom køn, etnicitet og social klasse påvirker den behandling borgere får i den offentlige sektor. Dette kan have direkte negative konsekvenser for det enkelte individ, men det strider også med normen om lige behandling af borgerne i et moderne bureaukrati. Derudover kan fejlopfattelser af borgeres behov og evner føre til en inefficent levering af offentlig service, og det kan direkte påvirke borgerne gennem forskelle i forventninger og dermed følgende selvopfyldende profetier. Endelig kan (negative) oplevelser i direkte kontakt med offentlige institutioner have alvorlige konsekvenser for borgeres politiske effektivitetsfølelse og politiske deltagelse. Konsulterer man den socialpsykologiske litteratur er det ikke overraskende at offentligt ansatte ligesom andre mennesker nogle gange bruger sociale kategorier som informationsmæssige genveje når de skal danne sig et indtryk af borgere. Især fordi offentligt ansatte ofte står over for en efterspørgsel, der er større end de tilgængelige ressourcer, og fordi studier af implicit diskrimination viser, at diskrimination kan være komplet uintenderet, og at den der diskriminerer ofte er helt ubevidst om dette. Mens dette tyder på, at en sådan adfærd delvist bygger på en automatisk reaktion viser studier i socialpsykologi, at mennesker kan undgå at blive påvirket af stereotyper, når de er tilpas motiverede for det og har tilstrækkelige kognitive ressourcer. Få studier viser imidlertid, hvordan man kan reducere diskriminerende forskelsbehandling med politisk kontrollerbare redskaber. Dette er overraskende siden en omfattende teoretisk og laboratoriebaseret psykologisk litteratur giver mange bud på, hvordan dette kan gøres. Denne afhandling undersøger, hvornår sociale kategorier bliver anvendt, og hvordan man kan reducere diskriminerende brug af dem i den offentlige sektor.

For at besvare dette spørgsmål opstiller afhandlingen en teoretisk model til at forklare, hvornår diskriminerende brug forekommer, og hvordan man kan reducere sådan brug. Modellen bygger på en kombination af resultaterne fra afhandlingens fem papirer, der trækker på offentlig forvaltningsteorier om repræsentativt bureaukrati, markarbejderes coping-adfærd og performance information og på indsigter hentet i den socialpsykologiske litteratur om sociale kategorier og stereotypopfattelser. Mere specifikt argumenteres der i afhandlingen for, at fremtrædende social distance føre til mere diskriminerende brug af sociale kategorier. Manglende kognitive ressourcer forstærker sådan brug fordi det begrænser opmærksomheden på den indflydelse sociale kategorier kan have gennem stereotype opfattelser og ultimativt gennem diskriminerende forskelsbehandling. Tilsvarende argumenteres der i afhandlingen for,

at man ved at mindske den sociale distance, ved at frigøre kognitive ressourcer og ved at sætte mere fokus på potentielt diskriminerende brug af sociale kategorier kan reducere denne brug. Betydningen af disse tre faktorer testes ikke direkte i afhandlingen, men ud fra modellen kan en række implikationer udledes angående diskriminerende brug af sociale kategorier, og hvordan dette kan mindskes. Specifikt foreslås det i afhandlingen, at reduktion af den sociale distance ved at gøre forvaltningen mere repræsentativ i forhold til de borgere, som forvaltningen betjener, at frigørelse af kognitive ressourcer ved at begrænse arbejdspresset og øget fokus på diskriminerende brug af sociale kategorier ved hjælp af systematisk performance information, der øger opmærksomheden på diskriminatorisk brug af sociale kategorier, kan reducere diskriminerende brug.

I afhandlingen kombineres kvasie- felt- og survey-eksperimentelle metoder, og der findes støtte til den opstillede model. Resultaterne har direkte implikationer i forhold til, hvordan man kan reducere diskriminerende brug af sociale kategorier ved hjælp af de foreslåede interventioner, men har også en række andre vigtige implikationer. For det første viser resultaterne, at det ikke kun reducerer diskriminerende brug af social kategorier, men også ultimativt forbedre borgeres resultater aktivt at matche dem med en markarbejder, der tilhører samme sociale kategori og som en konsekvens mindske den sociale distance. For det andet er arbejdspresset på offentligt ansatte stigende i mange lande – især på grund af offentlige besparelser i kølvandet på finanskrisen. Resultaterne peger på, at dette kan føre til en mere ulige levering af offentlig service. For det tredje viser resultaterne, at systematisk performance information i form af testresultater, der ingen betydning har for hverken elevernes fremtidsmuligheder eller skolernes finansiering, kan forbedre læring og ultimativt mindske forskellen mellem elever med høj og lav socioøkonomisk baggrund.

Opsummerende tager denne afhandling nogle af de første skridt imod en samlet teori om diskriminerende brug af sociale kategorier i den offentlige sektor, og om hvordan man kan reducere diskriminerende brug med politisk kontrollerbare redskaber. Målet for fremtid forskning er at finde flere og langtidsholdbare politiske redskaber, der kan bruges til at reducere diskriminerende brug af sociale kategorier.