

Noortje de Boer  
Utrecht University

## Research Article

# How Do Citizens Assess Street-Level Bureaucrats' Warmth and Competence? A Typology and Test

Noortje de Boer is assistant professor in the School of Governance at Utrecht University. Her research addresses why street-level bureaucrats do what they do and how citizens see and react to street-level bureaucrats. She studies topics such as transparency, stereotypes, and frontline behaviors. Her research has been published in journals such as *Public Administration Review* and *Public Administration*.  
Email: n.c.deboer@uu.nl

**Abstract:** *Citizens encounter many street-level bureaucrats in their lifetime. How do they assess the traits of the bureaucrats they meet? Understanding citizens' assessments of bureaucrats is important, because citizens are not passive receivers of policies. This article explores citizens' classifications of street-level bureaucrats based on their core task. Using a factorial survey ( $n = 580$ ), three clusters of bureaucrats are identified: those who are regulation oriented, those who are service oriented, and those who are both regulation and service oriented. Then, the article tests how these three types of bureaucrats are assessed on warmth and competence and whether their gender matters. A between-subjects experiment ( $n = 1,602$ ) reveals that regulation-oriented bureaucrats are assessed as least competent and warm. Moreover, regardless of core task, female bureaucrats are assessed as warmer than males. Female and male bureaucrats are assessed as equally competent. This article shows that bureaucrats are stereotyped by citizens and discusses the implications for the public management literature.*

## Evidence for Practice

- While researchers and practitioners tend to focus on commonalities between street-level bureaucrats, focusing on differences—for instance, in terms of core task and gender—may be helpful to understand how citizens see and react to their encounters with bureaucrats.
- Much like street-level bureaucrats' stereotypical notions of citizens, citizens stereotype the bureaucrats they encounter. These stereotypes are based on cues related to the bureaucrats' core task and gender.
- Especially for regulation-oriented organizations and street-level bureaucrats, such as inspectorates, being aware of stereotypes may be useful when engaging with citizens because they are rated as the least competent and warm.

Citizens encounter many different kinds of street-level bureaucrats in their lifetime, such as nurses, customs officers, and youth councilors. How do citizens assess the bureaucrats they meet? Understanding how citizens assess street-level bureaucrats is important because citizens are not passive actors during service provision (de Boer 2020; Dubois 2010; Gofen 2015). Variations in assessments of bureaucrats by citizens may have implications for service provision generally and for the behavior of citizens during bureaucrat-citizen encounters specifically. To illustrate, matching individual-level characteristics, such as gender, can increase the effort of both the bureaucrat and the citizen (Guul 2018; Meier 2018; Riccucci et al. 2016). Citizens' judgments about whether bureaucrats have friendly (or not) *intentions* or the *ability* to do their job are important to understand, because variations in judgments may affect, for instance, how citizens behave toward bureaucrats or treat bureaucrats during public encounters (see Fiske, Cuddy, and Glick 2007; Pautz and

Wamsley 2012). For instance, citizens may film police misconduct, which obstructs enforcement activities (Potere 2012).

There is little scholarship on the citizen side of service provision. Scholars have focused primarily on how bureaucrats deal with, behave, and make decisions during citizen encounters (e.g., de Boer and Eshuis 2018; Dubois 2010; Jensen and Pedersen 2017; Lipsky 2010; Loyens and Maesschalck 2010; Maynard-Moody and Musheno 2000; Tummers et al. 2015). As a result, little is known about how citizens assess bureaucrats or the implications for citizens' assessments of the implementation of policies. This is surprising because, from the bureaucrats' perspective, there is evidence that how citizens are judged—for instance, in terms of trust, stereotypes, or deservingness—matters during service provision. How bureaucrats assess citizens plays a role, at least to some extent, in how bureaucrats make decisions about services (Harrits 2019; Jilke and Tummers 2018; Keiser 2010; Pedersen, Stritch, and Thuesen 2018;

This article argues that citizens, much like bureaucrats, may rely on mental shortcuts to evaluate the person they encounter (see Jilke and Tummers 2018; Raaphorst, Groeneveld, and Van de Walle 2018; Willis and Todorov 2006). Before the actual implications of citizens assessments of bureaucrats can be investigated, however, we need understand how citizens perceive different types of bureaucrats. Both social psychology and political science have long traditions of understanding how humans assess each other in terms of the traits of warmth and competence (e.g., Fiske, Cuddy, and Glick 2007; Laustsen and Bor 2017; Wojciszke, Bazinska, and Jaworski 1998). This article draws on this trait assessment literature and investigates how citizens assess street-level bureaucrats with different core tasks in terms of their warmth and competence. The purpose of this article is twofold. First, it explores how citizens classify different bureaucrats based on their core task. A typology is created using a factorial survey design ( $n = 580$ ). Second, using this typology, this article then tests how bureaucrats' warmth and competence are assessed using a within-subjects experiment ( $n = 1,602$ ).

Two theories with roots in psychology are tested, with each focusing on a different cue that is expected to influence citizens' stereotypes of street-level bureaucrats. First, Hayes's (2005, 2010) issue ownership theory is used to test the effect of a bureaucrat's core task. Issue ownership theory states that political parties "own" some social issues more than others, which, in turn, reflects how party members' traits are assessed. The Democratic Party, for instance, is associated with the issue of welfare because it often discusses vulnerable groups, and its members are, in turn, assessed as warmer than their Republican counterparts (Hayes 2005, 2010). It is theorized that, like political parties, public organizations "own" some issues more because of their core tasks, and this explains how bureaucrats' traits are assessed.

However, it is not just issue ownership that has an impact on stereotypical beliefs about bureaucrats by citizens. Therefore, Eagly and Kagan's (2002) role congruency theory is used to test the effect of a bureaucrats' gender. Role congruency theory states that there are gender stereotypes and these may align (i.e., congruence) or clash (i.e., incongruence) with someone's leadership role. It is theorized that some core tasks of bureaucrats are (in)congruent with gender stereotypes and that this, in turn, affects competence and warmth assessments by citizens.

The remainder of this article is structured as follows: First, the conceptual underpinnings and hypothesized relations are highlighted. Then, the methodological considerations are presented, followed by the results of two survey experiments. The article ends with a concluding section and discussion of the findings for understanding the citizen side of public service provision specifically and public management more generally.

## Conceptual Framework

Scholars seem to agree that warmth and competence are the two main dimensions along which humans assess one another's traits (e.g., Fiske, Cuddy, and Glick 2007; Laustsen and Bor 2017; Wojciszke, Bazinska, and Jaworski 1998). Fiske, Cuddy, and Glick (2007, 77) highlight that "the warmth dimension captures traits that are related

to perceived intent, including friendliness, helpfulness, sincerity, trustworthiness and morality, whereas the competence dimension reflects traits that are related to perceived ability, including intelligence, skill, creativity and efficacy." In other words, the warmth dimension encompasses how someone judges another person's *intentions*, and the competence dimension concerns how a person's *ability* is judged. However, in the trait literature, there is discussion of which trait is more dominant. Social psychologists argue—based on evolutionary pressures—that social perceptions are based, first, on whether someone has good intentions or not (i.e., the warmth dimension) and, second, on whether someone can behave according to intentions (i.e., the competence dimension) (Fiske, Cuddy, and Glick 2007). Political scientists argue—using the relevancy of tasks—for the exact same opposite dominance, since citizens evaluate candidates based on whether they are competent enough to fulfill their political tasks before evaluating their warmth (Funk 1999; Laustsen and Bor 2017).

## Issue Ownership

But *how* do citizens assess the warmth or competence of street-level bureaucrats? Different information cues have been related to variations in trait assessments, ranging from the party affiliation of political candidates to facial features (Hall et al. 2009; Olivola and Todorov 2010). Political scientists have found that trait assessments are rooted in the issues "owned" by the political party that an individual politician represents. An institution's "issue ownership" is defined as a "reputation for policy and program interest, produced by a history of attention, initiative, and innovation toward problems, which leads [citizens] to believe that one of the [institutions] is more sincere and committed to do something" (Petrocik 1996, 826). Issue ownership develops as a result of the reputation that political parties build regarding issues they *can* and *cannot* handle skillfully; this, in turn, is transferred to how politicians working for those parties are assessed regarding their traits (Goren 2007; Hayes 2005; Petrocik 1996). Essentially, this indicates that the social issues that political parties embody, such as health care or law enforcement, trickle down to how citizens judge the traits of politicians.

Hayes (2005, 2010) builds on the notion of issue ownership and has developed it regarding (1) how citizens evaluate the traits of political candidates and (2) how such trait evaluations impact vote choice. Hayes's (2005) issue ownership theory states that trait evaluations of presidential candidates (i.e., individuals) by the public are determined by the political parties (i.e., institutions) they represent. Trait ownership theory stresses that public perceptions of presidential candidates' traits relate to the issues their political parties campaign on and thus talk about frequently. Presidential candidates' traits are determined by issue ownership because the public associates specific character traits with candidates because they frequently publicly campaign on them. To put it differently, citizens judge the traits of presidential candidates predominantly based on the stereotypical notions they have of the political parties they represent and not on information cues related specifically to the candidate (Hayes 2005, 2010).

Public organizations do not campaign for votes, nor do they have candidates competing with one another. Nevertheless, the reputation of a (public) institution—a core component of Hayes's (2005) trait ownership theory—is widely acknowledged to be important for organizational behavior and public perceptions.

Organizational reputation influences bureaucratic behavior and institutional patterns such as the amount of autonomy and discretion (Busuioc and Lodge 2016; Carpenter and Krause 2012). Public organizations also manage and use their reputation to achieve more favorable public perceptions (Maor, Gilad, and Bloom 2013; Teodoro and An 2018; Wæraas 2017; Wæraas and Byrkjeflot 2012). Hence, in line with Hayes's (2005) theory of issue ownership, certain public organizations may be affected by the stereotypical thinking of citizens because of their reputation for handling societal issues, which, in turn, reflects how citizens assess the traits of the street-level bureaucrats they encounter (see Hvidman 2018).

Citizens encounter a range of street-level bureaucrats who are employed by public organizations with different core tasks. According to Jensen (2018), public organizations can be roughly classified as either regulation focused (e.g., police or inspectorates), service focused (e.g., schools or hospitals), or both (e.g., rehabilitation centers). More specifically, regulation-driven organizations "deliver obligations rather than services" (Sparrow 2000, 2), such as fines and sanctions. Service-oriented organizations offer services, such as care and unemployment benefits (Jensen 2018). Notably, public organizations are almost never perceived as fully regulation or service oriented. For example, a school inspector may fire a school director for not meeting requirements regarding budget plans but may also provide advice on how to improve student well-being. Regardless, public organizations are seen to "own" certain societal issues more than others (see Hayes 2005; Petrocik 1996) because there are differences in the precise job tasks that public organizations have and, in turn, *what* street-level bureaucrats deliver to citizens. In other words, it could be that the stereotypical notions depicted by Hayes (2005) are also present for public organizations when citizens assess street-level bureaucrats' traits because the nature of public organizations' core tasks give them ownership of some issues.

More specifically, Hayes (2005) found that, on the one hand, Democrats often campaign at locations involving vulnerable groups of citizens such as nursing homes, and therefore they "own" the issue of the welfare of the vulnerable. In turn, citizens perceive their presidential candidates to be warmer (labeled "empathetic" and "compassionate" by Hayes 2005) than their Republican counterparts. Republicans, on the other hand, focus on law and order more (labeled "leadership" by Hayes 2005) than Democrats. It is expected that this mechanism is also present for public organizations and their street-level bureaucrats. More specifically, service-oriented public organizations will be associated with vulnerable groups of citizens, such as the elderly, sick, and unemployed (i.e., the warmth dimension). Enforcement-oriented public organizations, on the other hand, will be associated with law and order because they punish wrongdoers and sanction accordingly (i.e., the competence dimension). In sum, based on issue ownership theory, it is expected that:

**Hypothesis 1a:** Bureaucrats employed by regulation-oriented organizations will be assessed as *more* competent than bureaucrats employed by service-oriented organizations.

**Hypothesis 1b:** Bureaucrats employed by regulation-oriented organizations will be assessed as *less* warm than bureaucrats employed by more service-providing organizations.

## Role Congruence

According to Hayes (2005, 2010), citizens draw on campaign information to assess whether political candidates behave in line with the traits "owned" by their political parties because they rarely meet the candidates in person. However, unlike political candidates, citizens *do* meet street-level bureaucrats face-to-face during service provision (Lipsky 2010; Maynard-Moody and Musheno 2000). These bureaucrats are an important source of episodic information that steers citizens' assessments (Olsen 2017b). To illustrate, if you call your local police station to report a theft, you may already have predispositions about the police as an institution, but you are also directly interacting with another human being, the police officer. Whether the police officer has a female and not a male voice and says she is a deputy and not a detective may provide you with episodic cues to draw on and, ultimately, assess the traits of the person you are dealing with.

Notably, because of technological advancements such as digitalization, there seems to be a shift toward more bureaucrat-citizen encounters that are, by design, low-information settings and therefore impersonal in nature. For instance, contact between bureaucrats and citizens is increasingly moderated by technology and face-to-face contact is being replaced by screen-to-screen contact, by telephone, email, or social media (Bovens and Zouridis 2002). In addition, for some citizens, especially nonvulnerable ones, the only times they encounter bureaucrats are in singular settings rather than repeated (i.e., plural) (see Black 2010; Boyne, Day, and Walker 2002). To illustrate, you may encounter a parking officer who is fining you for wrongfully parking your car. The next time that you get fined, it probably will not be the same parking officer. In these cases, citizens may rely only on name or voice when assessing the bureaucrat. In such low-information encounters, any identifiable information available to citizens to identify street-level bureaucrats matters because citizens' reliance on stereotypical thinking changes when more cues can be drawn on (Bauer 2013, 2015; Kunda and Spencer 2003).

Gender is a particularly powerful cue because it can be identified based on multiple sources of information, such as physical appearance, name, and voice. Moreover, gender cues are hard to hide. In this line of reasoning, Eagly and Karau's (2002) role congruence theory proposes that stereotypical characteristics associated with someone's gender role may be (in)congruent with other roles, such as leadership positions. Gender roles are "those shared expectations (about appropriate qualities and behavior) that apply to individuals on the basis of their socially defined gender" (Eagly 1997, 12) or "sets of norms that communicate what is generally appropriate for each sex" (Burn 1996, 3). Gender roles for females, for instance, include traits such as sensitive, warm, and compassionate (Bauer 2013, 2015). Following Eagly and Karau (2002), these gender roles and associated traits do not match the expected roles, for instance, of leaders (Eagly and Karau 2002; Konrad and Cannings 1997; Ritter and Yoder 2004). Moreover, women are generally characterized as warmer (see also Bauer 2015) and less competent than men (see Herrnson, Lay, and Stokes 2003; Huddy and Terkildsen 1993; Lawless 2004). Drawing on role congruence theory (Eagly and Karau 2002), it is expected that these stereotypes are also present for street-level bureaucrats. The second hypothesis, based on role congruence theory, is as follows:

**Hypothesis 2:** Female bureaucrats will be assessed as *more* warm and *less* competent than male bureaucrats.

Stereotypical notions of the public organizations and gender of bureaucrats can also interact. To illustrate, when you call the police station to report a break-in and get a female police officer on the phone, these two cues may be incongruent and, in turn, decrease the warmth assessment. However, when you call your local hospital with an emergency—a public organization “owning” the “warm” social issue of health care—and get a female nurse on the phone, these cues are congruent and, ultimately the cues may amplify the warmth assessment. In other words, the warmth of women may not correspond to the stereotypical notions associated with regulation-oriented organizations focused on law and order. Masculinity is associated with law and order-related topics such as economic and military issues (e.g., Huddy and Terkildsen 1993; Lawless 2004). Male bureaucrats may therefore be congruent with stereotypical associations of regulation-oriented organizations. The warmth of female bureaucrats, however, may be congruent with service-oriented organizations because they are associated with vulnerable groups but not with male bureaucrats. Hence, combining the theories on issue ownership and role congruence, it is expected that when the core task and gender are congruent (service oriented \* female and regulation oriented \* male), traits will be assessed higher. When core task and gender are *incongruent* (regulation oriented \* female and service oriented \* male), traits will be assessed lower.

**Hypothesis 3:** When issue ownership and gender are congruent, trait assessment effects will be larger than when issue ownership and gender are incongruent.

## Study 1

This study was used to explore how citizens classify different types of street-level bureaucrats based on their core task (i.e., regulation focused or service focused).

**Design.** A factorial survey was used, which is a method well suited to exploring the underlying principles of attitudes and perceptions. A factorial survey is a within-subjects experimental method in which respondents are exposed to X vignettes that vary across X dimensions based on their levels (Auspurg and Hinz 2014). One dimension with one level was varied, namely, the *type* of street-level bureaucrat. Thirty types of street-level bureaucrats were formulated based on six main public enforcement domains in the Netherlands: (1) public space; (2) environment, well-being, and infrastructure; (3) education; (4) public transport; (5) work, income, and health care; and (6) general investigation. To provide a balance between regulation- and service-oriented bureaucrats, half of the formulated street-level bureaucrats had official investigative authority (*buitengewoon opsporingsambtenaar*) and the other half did not (see appendix A).

**Procedure.** To prevent boredom, fatigue, or respondents identifying the manipulation, the vignette sample was divided into six decks of five vignettes (see appendix A). Respondents were randomly allocated to a deck. Within decks, the order of the vignettes was randomized. Since each of the 116 respondents rated five vignettes, the total analytical sample is 580 observations (Auspurg and Hinz 2014). After answering some demographic questions, respondents rated each vignette on a scale ranging from 1 to 10 for the

dependent variable of this study, the two core tasks. The item for the perceived regulation-oriented task was formulated as follows: “The core task of the [insert type of street-level bureaucrat] is enforcing laws and regulations.” The item for the perceived service-oriented task was formulated as “The core task of the [insert type of street-level bureaucrat] is providing services.”

**Respondents.** The factorial survey was distributed via Qualtrics panels in February 2019 and its software among a sample of the Dutch population (see table 1 for sample characteristics and the Supporting Information online for a detailed discussion of the panel and sample). Notably, citizens with higher education are slightly overrepresented, and those with no to little education are underrepresented. This should be taken into account when interpreting the findings.

## Results

First, for each of the 30 vignettes, the ratings of respondents on the two dependent variables (a regulation-focused core task and a service-focused core task) were used to calculate mean scores for each type of street-level bureaucrat (see appendix A). The structure of the data was explored using cluster analysis, which allows for the identification of optimal grouping of observations in clusters and dissimilarities across clusters. Hierarchical clustering was used since the optimal *n* of clusters is unknown (Rencher 2003). Interpreting the agglomeration statistics<sup>1</sup> of hierarchical cluster analysis, not two but three clusters were optimal<sup>2</sup> with a maximum silhouette width of 0.46. To identify which types of street-level bureaucrats belong to which cluster, *k*-means clustering was used<sup>3</sup> (Rencher 2003). Table 2 displays the mean scores of the three clusters on both core tasks. Sixty-five percent of the total variance is explained by the clustering.

The first cluster can be classified as primarily service oriented, which higher scores on the service-oriented (*M* = 4.95) than the regulation-oriented (*M* = 3.94) core task. Street-level bureaucrats

**Table 1** Sample Characteristics

		Study 1	Dutch Population <sup>a</sup>
Age (M)		39.3	40.1
Gender (%)	Female	25.9	49.4
	Male	74.1	50.6
Education level (%)	None	0	1.0 <sup>b</sup>
	Elementary	2.6	9.0
	High school	23.3	60.0 <sup>c</sup>
	MBO	24.1	
	Bachelor's	19.0	19.0
	Master's	29.3	11.0 <sup>d</sup>
Ethnic minority (%)	MBA	0.9	
	PhD	0.9	
	Yes	25.0	24.0
	No	75.0	76.0

<sup>a</sup>Based on 2019 data from the Dutch Central Bureau for Statistics (<http://www.cbs.nl>).

<sup>b</sup>Includes unknown.

<sup>c</sup>Includes both high school and MBO.

<sup>d</sup>Includes master's, MBA and PhD.

**Table 2** Cluster Means

Cluster	Mean Score Regulation Oriented	Mean Score Service Oriented
1	3.94	4.95
2	5.30	3.90
3	5.06	5.35



in this cluster include nurses, maternity assistants, and elementary school teachers. The second cluster is primarily regulation oriented, scoring high on the regulation-oriented ( $M = 5.30$ ) core task and low on the service-oriented ( $M = 3.90$ ) core task. Bureaucrats in this cluster include parking officers, customs officers, and health care inspectors. Finally, the third cluster scores high on both core tasks ( $M = 5.06$  on the regulation-oriented core task,  $M = 5.35$  on the service-oriented core task). This cluster includes youth care workers, debt councilors, and police officers. Figure 1 shows the composition of all three clusters; cluster 1 is indicated by dark gray, cluster 2 by medium gray, and cluster 3 by light gray.

### Study 2

The second study tested whether trait assessments of street-level bureaucrats are affected by the core task (by testing issue ownership theory) and by gender (by testing role congruency theory).

### Design

A  $3 \times 2$  within-subjects design was used to test the hypothesized relations in a survey experiment. The respondents were asked to imagine that they contacted a local street-level bureaucrat and asked to rate the bureaucrat based on their first impression. Based on the results of the first experiment, the experimental manipulation for core task included three (instead of two) sets of bureaucrats (regulation oriented, service oriented, and both). From each cluster, three bureaucrats were chosen at random to increase generalizability across bureaucrats (see table 3). To minimize differences in characteristics across groups, respondents were randomly allocated to one of three bureaucrats within each treatment group. The second experimental manipulation, for gender, included varying the name of the street-level bureaucrat. Jan indicated male and Anna female, which are typical Dutch names (for similar manipulations, see Baekgaard and George 2018; Riccucci, Van Ryzin, and Li 2016).

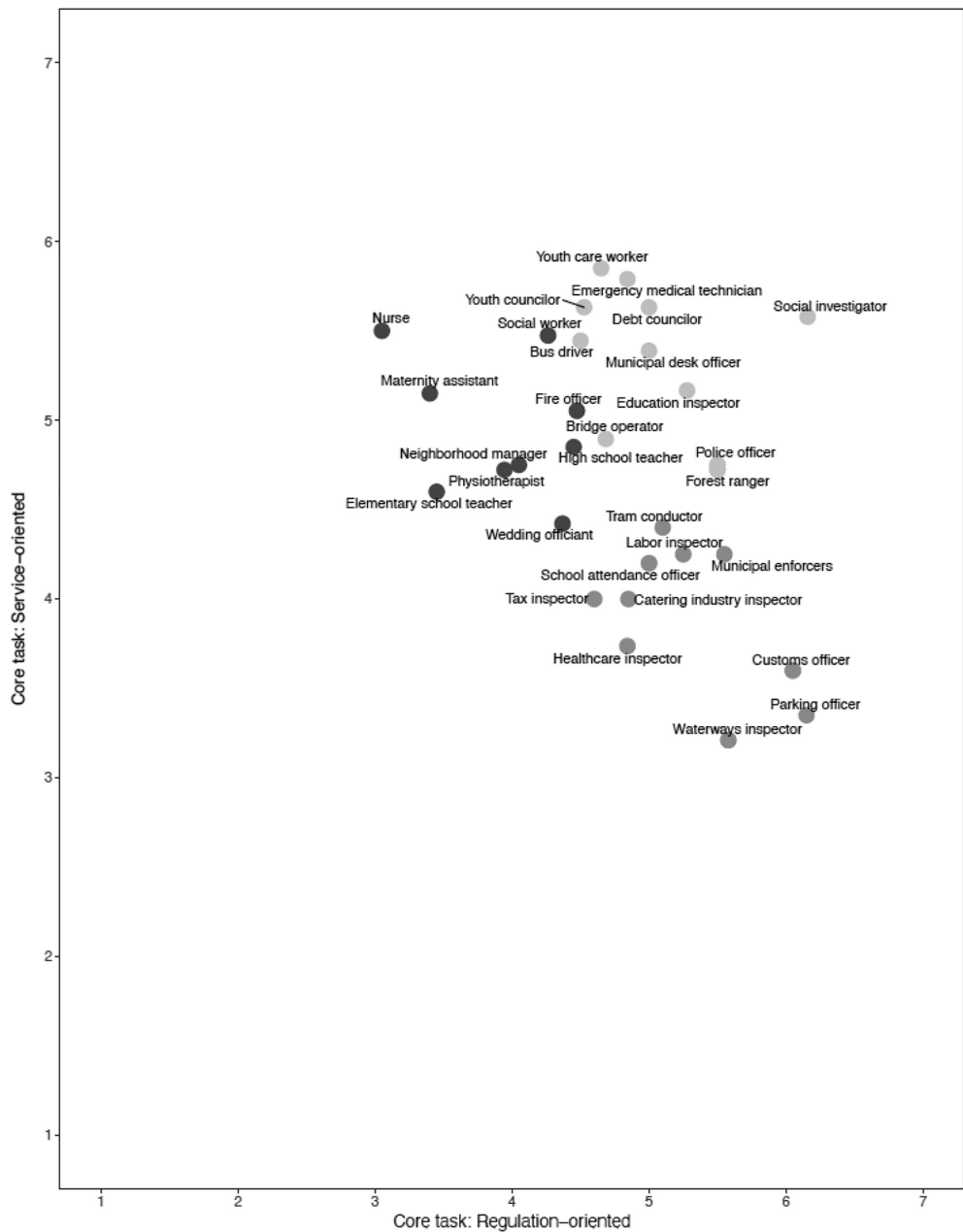


Figure 1 Three-Cluster Solution

Two potential covariates were controlled for. First, the need for the bureaucrat-citizen interaction (sentence 1 in the vignette) and the type of contact (sentence 2 in the vignette) were constant because we did not want respondents to draw their own conclusions about the public encounter. Despite potential threats to external validity, a deliberate choice was made to use a low-cue setting (telephone) instead of a face-to-face interaction, where, arguably, more cues are present to boost realism. Second, one sentence on the outcome of the contact with the street-level bureaucrat was included as a way to control for the potential influence of this covariate to ensure that respondent did not make inferences about possible outcomes. In this way, respondents were forced to rely only on the cues that *were* varied and also boosted realism.

The resulting vignette was presented as follows:

Imagine that you have a question for a [type of bureaucrat based on core task] to which you cannot find the answer yourself. You, therefore, contact a [type of bureaucrat based on core task] by phone. Your request is respondent by [gender]. Your question is answered within the average completion time of 15 minutes.

Procedure

Respondents were first asked to fill in some demographic questions, followed by the introduction to the experiment. Respondents then were randomly allocated to one of the six treatment groups. Within each treatment group, respondents were randomly shown one of three possible street-level bureaucrats (see table 3). After the treatment, all respondents filled in an identical post-experiment questionnaire about the dependent variables. Following the well-established trait assessment literature, the dependent variable in this study, trait assessment, was measured using validated items of its two dimensions: perceived competence and warmth (see Fiske et al., 1999; 2002; Funk 1996; Goren 2002). Respondents were asked, “How well or badly do the following words describe [X]?” X was replaced by the street-level bureaucrat in their treatment. The words *competent, confident, capable, efficient, intelligent, and qualified* were used to measure competence. The words *likeable, good-natured, friendly, warm, sincere, and caring* measured warmth. All six items were measured on a 7-point scale ranging from very bad (1) to very well (7). The items form reliable scales with  $\omega = .89$  (competence) and  $\omega = .83$  (warmth).

Respondents

The survey experiment was distributed via Qualtrics panels in June 2019 and its software among a sample of the Dutch population ( $n = 1,602$ ) (see table 4 for sample characteristics and the Supporting Information for a detailed discussion on the panel and sample). Notably, in line with in study 1, citizens with a high education are overrepresented and those with no to little education are underrepresented compared with the general Dutch population. Moreover, respondents are also slightly older.

Balance Checks

A series of (logistic) regressions were used to test whether the six experimental groups were balanced in terms of demographics (age, gender, education level, and ethnicity), prior experience, and political preferences (see appendix B). Results indicate that

Table 3 Treatment Groups’ Core Tasks

Core Task	Street-Level Bureaucrats Included
Regulation oriented	1. Municipal enforcer 2. Customs officer 3. Health care inspector
Service oriented	1. Nurse 2. Maternity assistant 3. Wedding officiant
Both	1. Emergency medical technician 2. Youth councilor 3. Debt councilor

Table 4. Sample Characteristics

		Study 2	Dutch Population <sup>a</sup>
Age (M)		48.8	40.1
Gender (%)	Female	45.4	49.4
	Male	54.5	50.6
	Other	0.1	1.0 <sup>b</sup>
Education level (%)	None	1.0	9.0
	Elementary	2.1	60.0 <sup>c</sup>
	High school	27.0	
	MBO	32.0	19.0
	Bachelor’s	19.1	11.0 <sup>d</sup>
	Master’s	16.4	
	MBA	0.8	
	PhD	1.6	24.0
Ethnic minority (%)	Yes	11.2	76.0
	No	88.8	

<sup>a</sup>Based on 2019 data from the Dutch Central Bureau for Statistics (<http://www.cbs.nl>).

<sup>b</sup>Includes unknown.

<sup>c</sup>Includes both high school and MBO.

<sup>d</sup>Includes master’s, MBA and PhD.

randomization was successful and that the sample is balanced across groups, as only one difference is statistically significant, which is comparable with results due to chance (for similar testing and results, see Baekgaard and George 2018).

Results

**Hypothesis 1.** Hypothesis 1 predicted a main effect—specifically, that bureaucrats with a regulation-oriented core task would be assessed as (1) more competent and (2) less warm than bureaucrats with a service-oriented core task. Based on study 1, an additional main effect was tested for which no hypothesis was formulated, namely, the effect of bureaucrats having both a regulation- and service-oriented core task on competence and warmth ratings by participants. Analysis of variance indicates that the main effects of core task (i.e., regulation oriented, service oriented, or both) on both competence ( $F[1,1,599] = 6.24, p = .002$ ) and warmth ratings of bureaucrats by participants ( $F[1,1,599] = 10.61, p = .000$ ) are statistically significant.<sup>4</sup>

However, for competence, post hoc analysis shows that hypothesis 1a is rejected. Contrary to expectations, table 5 shows that competence ratings are not highest but lowest among regulation-oriented bureaucrats ( $M = 4.87, SD = 1.03$ ). Competence ratings are highest among bureaucrats with a core task that is both regulation and service oriented ( $M = 5.08, SD = 1.05$ ), followed by service-oriented bureaucrats ( $M = 5.03, SD = 0.95$ ). More specifically, the mean difference in competence

**Table 5.** Descriptive Statistics Competence and Warmth Ratings

		N	Mean	SD	Min.	Max.
Competence	Regulation oriented	534	4.87	1.03	1	7
	Service oriented	534	5.03	0.95	1	7
	Both	534	5.08	1.05	1	7
	<i>Total</i>	1,602	5.00	1.02	1	7
Warmth	Regulation oriented	534	4.88	1.00	1	7
	Service oriented	534	5.10	0.96	1	7
	Both	534	5.15	1.06	1	7
	<i>Total</i>	1,602	5.04	1.01	1	7

**Table 6.** Main Effect on Competence and Warmth Ratings

Dependent Variable	Group A	Group B	Mean diff. (A-B)	SE	Cohen's <i>d</i>
Competence	Service	Regulation	0.16** (0.023)	0.06	0.16
	Both	Regulation	0.21** (0.002)	0.06	0.20
	Both	Service	0.04 (0.751)	0.06	—
Warmth	Service	Regulation	0.22** (0.001)	0.06	0.22
	Both	Regulation	0.27*** (0.000)	0.06	0.26
	Both	Service	0.05 (0.72)	0.06	—

Notes: Post hoc using Tukey comparisons, *p*-value between brackets, \*\*\* *p* < .001, \*\* *p* < .05. Robustness checks found similar results while checking the three types of street-level bureaucrat in each of the three groups independently.

ratings between participants exposed to the bureaucrats with a service-oriented core task is 0.16 higher than those exposed to the regulation-oriented bureaucrats. Likewise, participants exposed to the bureaucrats with both a regulation- and a service-oriented core task rated the competence of that bureaucrat 0.21 higher than those exposed to the regulation-oriented bureaucrats. Both differences are statistically significant (with *p* = .023 and *p* = .002, respectively) and small effects (with *d* = 0.16 and *d* = 0.20, respectively). Notably, the difference between service-oriented bureaucrats and bureaucrats with both core tasks is not statistically significant (*p* = .751).

For warmth, post hoc analysis reveals that hypothesis 1b is confirmed. Table 5 confirms that warmth ratings of bureaucrats by participants are lowest among regulation-oriented bureaucrats (*M* = 4.88, *SD* = 1.00). Moreover, the mean difference between participants' ratings of bureaucrats' warmth exposed to the service-oriented bureaucrats is 0.22 higher than those treated to the regulation-oriented bureaucrat. Regulation-oriented bureaucrats are also rated 0.27 lower on warmth than bureaucrats with both core tasks. The effects are statistically significant (*p* = .001, *p* = .000) and small (*d* = 0.22, *d* = 0.26) (see table 6). In line with the mean differences for competence, no statistically significant difference is found between warmth ratings for bureaucrats with both core tasks and solely the service-oriented core task.

In sum, the post hoc analysis for competence ratings indicates that although there is a statistically significant effect of bureaucrats' core task on participants' trait ratings of bureaucrats, hypothesis 1a is rejected because the effect occurs in the opposite direction. To put it differently, bureaucrats with a regulation-oriented core task are assessed as *less* competent than bureaucrats with a service-oriented core task. In addition, they are rated as *less* competent than bureaucrats with both a regulation-oriented and a service-oriented core task. Thus, bureaucrats with a regulation-oriented core task are rated significantly less competent than bureaucrats with other core tasks. Hypothesis 1b is confirmed. In line with expectations, regulation-oriented bureaucrats rated less warm than service-

**Table 7** Descriptive Statistics Interaction Effect Gender on Competence and Warmth Ratings

			N	Mean	SD	Min.	Max.
Competence	Regulation oriented	Male	267	4.92	1.00	1	7
		Female	267	4.82	1.06	1	7
	Service oriented	Male	267	4.95	0.98	1	7
		Female	267	5.12	0.93	1	7
	Both	Male	267	5.10	0.96	1	7
		Female	267	5.06	1.13	1	7
Warmth	Regulation oriented	Male	801	4.99	0.98	1	7
		Female	801	5.00	1.05	1	7
	Service oriented	Male	267	4.89	0.98	1	7
		Female	267	4.88	1.03	1	7
	Both	Male	267	5.02	0.94	1	7
		Female	267	5.18	0.98	1	7
	<i>Total</i>	Male	801	5.10	1.05	1	7
		Female	801	5.04	1.02	1	7

oriented bureaucrats. Notably, all effects have a Cohen's *d* around 0.20, meaning that the difference in trait assessments is small.

**Hypothesis 2.** Hypothesis 2 expected that female bureaucrats would be assessed as *more* warm and *less* competent than male bureaucrats. Independent two-group *t*-test analysis shows that the effect of gender on competence is not statistically significant (*t*[1593] = -0.275, *p* = .784). The mean scores on competence of male (*M* = 4.99, *SD* = 0.98) and female bureaucrats (*M* = 5.00, *SD* = 1.05) are almost identical. There is a statistically significant difference in mean scores on warmth (*t*[1591] = -2.10, *p* = .036). Female bureaucrats are assessed as warmer (*M* = 5.10, *SD* = 1.05) than male bureaucrats (*M* = 4.99, *SD* = 0.98). The effect is small with Cohen's *d* of 0.11. Hypothesis 2 is thus only partly confirmed. Female bureaucrats are assessed as warmer, as predicted, but there is no difference in competence assessment between male and female bureaucrats.

**Hypothesis 3.** Hypothesis 3 predicted that when issue ownership and gender are congruent (service oriented \* female and regulation oriented \* male), the effect on trait assessments would be larger than when issue ownership and gender are incongruent (regulation oriented \* male and service oriented \* female). ANOVA analyses reveal no statistically significant interaction effects for either competence (*F*[2,1,596] = 2.83, *p* = .06) or warmth (*F*[2,1,596] = 1.43, *p* = 0.240). Table 7 shows that the means differences across treatment groups are, indeed, highly similar. Hypothesis 3 is rejected.

## Conclusion and Discussion

This article has investigated (1) how citizens assess street-level bureaucrats and (2) whether core task and gender alter citizens' assessments of bureaucrats' competence and warmth by drawing on issue ownership (Hayes 2005, 2010) and role congruency theory (Eagly and Karau 2002). The results of this research contribute to the literature in four ways.

First and foremost, this article shows that, in the eyes of citizens, one uniform street-level bureaucrat does not exist. A typology of three groups of bureaucrats was found based on their perceived core task, namely, bureaucrats that are (1) regulation oriented, (2) service

oriented, and (3) both. These findings contribute to the street-level bureaucracy literature because they clearly speak to Maynard-Moody and Musheno's (2000, 2003) classic distinction between citizen- and state-agents. Maynard-Moody and Musheno (2000, 2003) reveal two dominant narratives among bureaucrats about how they make decisions. The state-agents stick to rules, regulations, and procedures to make consistent frontline decisions. The state is central, not the citizen. The citizen-agents put their judgments of citizens at the center and make decisions that are moral and ethical for the client. Our findings show that citizens largely mirror this analytical distinction of role perceptions of bureaucrats based on bureaucrats' core task.

Street-level bureaucracy work has focused predominantly on the commonalities among different types of bureaucrats, such as their discretion (e.g., Harrits 2019; Jilke and Tummers 2018; Lipsky 2010; Maynard-Moody and Musheno 2000, 2003; Pedersen, Stritch, and Thuesen 2018). However, there are also notable differences between bureaucrats. Zacka (2017, 23) illustrates that "unlike teachers, police officers carry guns and sometimes make life-or-death decisions; unlike welfare workers, these officers interact not just with individuals seeking services but with the population at large; and unlike social workers, who have repeated encounters with clients through which a personal relationship can develop, our encounters with police officers are often episodic and happen on a one-time basis." This study has shown that focusing on differences between types of bureaucrats, such as their core task, may be a fruitful avenue for future research. Future research could also investigate whether issue ownership trickles down to how citizens assess the traits of street-level managers (see Gassner and Gofen 2018).

Second, this article found that street-level bureaucrats' competence and warmth are assessed differently based on their core task (i.e., issue ownership theory). Bureaucrats with a primarily regulation-oriented task are assessed as *least* warm compared with bureaucrats executing a service-oriented core task as well as bureaucrats executing both core tasks. Contrary to expectations, regulation-oriented bureaucrats are also assessed as *least* competent. This could indicate that the reputation of regulation-oriented public organizations among citizens is least favorable compared with public organizations, and this, in turn, trickles down to perceptions of street-level bureaucrats. This finding could help explain variations in organizational reputation and help draft reputation management strategies of public organizations (Busuioc and Lodge 2016; Carpenter and Krause 2012; Maor, Gilad, and Bloom 2013).

In addition, the less competent and warm traits associated with regulation-oriented bureaucrats advance insights on citizens' evaluations of the public sector and the potential presence of negativity bias. Scholarship is mixed on whether public organizations can steer citizens toward more favorable interpretations (e.g., Baekgaard and Serritzlew 2016; James 2010; Olsen 2015, 2017a; Piotrowski, Grimmelikhuijsen, and Deat 2017). This study reveals that a potential explanation for whether public organizations can steer citizens' evaluations is their possession of (un)favorable traits (i.e., [not] warm, [not] competent). Moreover, the findings could indicate that negative perceptions of public organizations cannot just be clustered by subgroups of citizens

(Hvidman 2018) but also in subgroups of public organizations or their street-level agents. Future research could investigate whether stereotypical notions of street-level bureaucrats can also trickle back up to citizens' assessments of public organizations at large. Experimental methods would be particularly suitable.

Third, bureaucrats' competence and warmth are assessed differently based on the bureaucrat's gender. No differences were found in competence assessments between male and female bureaucrats. However, female bureaucrats were assessed as slightly warmer than males. This finding contributes to the literature on representative bureaucracy. There are mixed findings on whether a bureaucrats' gender affects citizens (e.g., Doornkamp, Van den Bekerom, and Groeneveld 2019; Guul 2018; Piatak and Mohr 2019). Only Doornkamp, Van den Bekerom, and Groeneveld (2019) measure stereotypical beliefs but do so directly, which, as the authors themselves already indicate, is a limitation. This article does dissect stereotypical gender beliefs in terms of traits. Our findings indicate that differences in warmth perceptions could explain the mixed findings of the effect of a bureaucrats' gender on citizens' biases. In other words, because male and female bureaucrats are perceived differently in terms of warmth, that could explain why citizens sometimes are and sometimes are not affected by bureaucrats' gender. Notably, it would also be fruitful to connect our findings to the work by scholars using social identity theory to study stereotypical beliefs (e.g., Abrams and Hogg 1990; Stets and Burke 2000). Future research could explore whether gender stereotypes are affected when citizens identify as belonging to the same group as the street-level bureaucrat they encounter.

Fourth, gender and core task did not strengthen or weaken each other's effects. This result contributes to the literature on stereotyping and discrimination (e.g., Harrits 2019; Jilke and Tummers 2018; Pedersen, Stritch, and Thuesen 2018; Raaphorst, Groeneveld, and Van de Walle 2018; Thomann and Rapp 2018; Yang 2005). It could be that other cues that were not included in the experimental design are (in)congruent and, in turn, affect a bureaucrats' trait assessment. Cues such as seniority or physical appearance could be cues to incorporate. Future research could further explore the numerous cues citizens can draw on during bureaucrat-citizen encounters or investigate how core task and a bureaucrats' gender affect other aspects, such as bureaucrats' trustworthiness.

A note of caution is necessary when interpreting the findings since all the effects are small. This is not surprising as the isolated effects of only two cues were tested. In real life, citizens draw on multiple cues to assess the bureaucrats across the table. These small effects are also in line with stereotyping research, showing that cues work in subtle ways (Raaphorst, Groeneveld, and Van de Walle 2018). However, even small effects can have profound implications for the day-to-day interactions of citizens with bureaucrats and, potentially, how trustworthy and cooperative they are during those encounters. Regardless, this article shows that bureaucrats, much like citizens, are stereotyped in terms of traits based on core task and gender. Future research is needed to assess the implications of the variations in trait assessment identified in this article.

As with any research, there are limitations to this study. First, a limitation of survey experiments is the low external validity because



simplistic treatments are induced that are less complex than real-world settings. This limitation should be taken into account. On the one hand, citizens can be expected to use cues to assess the traits of bureaucrats based on more than one heuristic in real-life encounters. On the other hand, reliance on heuristics may be even more prone in real-life settings since multiple stereotypical cues will be available to the citizen regarding the same characteristic of the street-level bureaucrat. Core task, for instance, could be induced by function, clothing and type of office. In addition, some contact between citizens and bureaucrats may be more likely to occur in real-life settings. For instance, the majority of citizens ask more questions of nurses in their lifetime than waterway inspectors. Nevertheless, our effects are robust across 30 different bureaucrats. Hence, though our findings likely generalize to other low-information settings, experimental replications of our findings are needed to assess the implications for the external validity. Replications using field experiments and thus real-life interactions will be particularly fruitful (for an overview, see Hansen and Tummers 2020).

Second, in the experimental design, the outcome of the bureaucrat-citizen contact was controlled for by stating that the question the citizen had *was* answered. It could be that stereotypical beliefs change when the question is *not* answered or that stereotypical beliefs are mitigated. Future research could consider theorizing the effect of the outcome of the public encounter on stereotypical beliefs and including it as a treatment. Third, though names were used to manipulate sex, which is a common practice in the trait literature, it could be that there are other attributions connected to those names, such as socioeconomic status. Future research could include a manipulation check about the names being varied. Last but not least, we varied the core task in study 2 using different occupations. Though this variation was empirically grounded in our empirical results of study 1, we acknowledge that this is a proxy. Future studies should incorporate more extensive treatment checks in pre-studies to ensure no other related associations are made by respondents when presented with our treatment.

This article has made the important first step in understanding how citizens assess street-level bureaucrats' traits. If we want to understand the implications of variations in assessments, it is essential for future research to dissect the effects of trait assessment on the way bureaucrat-citizen encounters unfold and how citizens behave.

## Acknowledgments

I want to thank Asmus Olsen for his invaluable feedback and advice during my stay at Copenhagen University. He was instrumental in the development of the theoretical foundation of this article. I also want to thank Erik Hans Klijn and Jasper Eshuis for providing the funds to conduct this study.

## Notes

1. Including Euclidean distance, dendrogram, cluster membership, silhouette plot (Hair et al. 1995; Rencher 2003).
2. This process resembles the interpretation of a scree plot and its eigenvalues in factor analysis.
3. Following Rencher (2003), both complete and average linkages were compared. The results were similar to a substantial extent, indicating that natural clusters were identified in the data.
4. A multivariate analysis of variance generated similar significant effects.

## References

- Abrams, Dominic, and Michael A. Hogg. 1990. *Social Identity Theory: Constructive and Critical Advances*. New York: Springer-Verlag.
- Baekgaard, Martin, and Bert George. 2018. Equal Access to the Top? Representative Bureaucracy and Politicians' Recruitment Preferences for Top Administrative Staff. *Journal of Public Administration Research and Theory* 28(4): 535–50.
- Baekgaard, Martin, and Søren Serritzlew. 2016. Interpreting Performance Information: Motivated Reasoning or Unbiased Comprehension. *Public Administration Review* 76(1): 73–82.
- Bauer, Nichole M. 2013. Rethinking Stereotype Reliance: Understanding the Connection between Female Candidates and Gender Stereotypes. *Politics and the Life Sciences* 32(1): 22–42.
- . 2015. Who Stereotypes Female Candidates? Identifying Individual Differences in Feminine Stereotype Reliance. *Politics, Groups, and Identities* 3(1): 94–110.
- Black, Donald. 2010. *The Behavior of Law*, Special ed. Bingley, UK: Emerald Group.
- Boyne, George, Patricia Day, and Richard Walker. 2002. The Evaluation of Public Service Inspection: A Theoretical Framework. *Urban Studies* 39(7): 1197–212.
- Burn, Shawn Meghan. 1996. *The Social Psychology of Gender*. New York: McGraw-Hill.
- Busuioc, E. Madalina, and Martin Lodge. 2016. The Reputational Basis of Public Accountability. *Governance* 29(2): 247–63.
- Carpenter, Daniel P., and George A. Krause. 2012. Reputation and Public Administration. *Public Administration Review* 72(1): 26–32.
- de Boer, Noortje. 2020. The (Un)intended Effects of Street-Level Bureaucrats' Enforcement Style: Do Citizens Shame or Obey Bureaucrats? *Public Policy and Administration*. Published online February 26. <https://doi.org/10.1177/0952076720905005>.
- de Boer, Noortje, and Jasper Eshuis. 2018. A Street-Level Perspective on Government Transparency and Regulatory Performance: Does Relational Distance Matter? *Public Administration* 96(3): 452–67.
- Doornkamp, Laura, Petra Van den Bekerom, and Sandra Groeneveld. 2019. The Individual Level Effect of Symbolic Representation: An Experimental Study on Teacher-Student Gender Congruence and Students' Perceived Abilities in Math. *Journal of Behavioral Public Administration* 2(2). <https://doi.org/10.30636/jbpa.22.64>.
- Dubois, Vincent. 2010. *The Bureaucrat and the Poor: Encounters in French Welfare Offices*. New York: Routledge.
- Eagly, Alice H. 1997. Sex Differences in Social Behavior: Comparing Social Role Theory and Evolutionary Psychology. *American Psychologist* 52(12): 1380–3.
- Eagly, Alice H., and Karau Steven J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological review* 109(3): 573.
- Fiske, Susan T., Xu, J., Amy J.C. Cuddy, and Peter Glick. 1999. (Dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of social issues* 55(3): 473–89.
- Fiske, Susan T., Amy J.C. Cuddy, Peter Glick, and Xu, J. 2002. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology* 82(6): 878.
- Fiske, Susan T., Amy J.C. Cuddy, and Peter Glick. 2007. Universal Dimensions of Social Cognition: Warmth and Competence. *Trends in Cognitive Sciences* 11(2): 77–83.
- Funk, Carolyn L. 1996. The impact of scandal on candidate evaluations: An experimental test of the role of candidate traits. *Political Behavior*, 18(1): 1–24.
- Funk, Carolyn L. 1999. Bringing the Candidate into Models of Candidate Evaluation. *Journal of Politics* 61(3): 700–20.
- Gassner, Drorit, and Anat Gofen. 2018. Street-Level Management: A Clientele-Agent Perspective on Implementation. *Journal of Public Administration Research and Theory* 28(4): 551–68.
- Gofen, Anat. 2015. Citizens' Entrepreneurial Role in Public Service Provision. *Public Management Review* 17(3): 404–24.

- Goren, Paul. 2002. Character weakness, partisan bias, and presidential evaluation *American Journal of Political Science* 627–41.
- Goren, Paul. 2007. Character Weakness, Partisan Bias, and Presidential Evaluation: Modifications and Extensions. *Political Behavior* 29(3): 305–25.
- Gul, Thorbjørn Sejr. 2018. The Individual-Level Effect of Gender Matching in Representative Bureaucracy. *Public Administration Review* 78(3): 398–408.
- Hansen, Jesper Asring, and Lars Tummers. 2020. A Systematic Review of Field Experiments in Public Administration. *Public Administration Review*. Published online March 23. <https://doi.org/10.1111/puar.13181>.
- Harrits, Gitte Sommer. 2019. Stereotypes in Context: How and When Do Street-Level Bureaucrats Use Class Stereotypes? *Public Administration Review* 79(1): 93–103.
- Hayes, Danny. 2005. Candidate Qualities through a Partisan Lens: A Theory of Trait Ownership. *American Journal of Political Science* 49(4): 908–23.
- . 2010. Trait Voting in US Senate Elections. *American Politics Research* 38(6): 1102–29.
- Herrnson, Paul S., J. Celeste Lay, and Atiya Kai Stokes. 2003. Women Running “as Women”: Candidate Gender, Campaign Issues, and Voter-Targeting Strategies. *Journal of Politics* 65(1): 244–55.
- Huddy, Leonie, and Nayda Terkildsen. 1993. Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science* 37(1): 119–47.
- Hvidman, Ulrik. 2018. Citizens’ Evaluations of the Public Sector: Evidence from Two Large-Scale Experiments. *Journal of Public Administration Research and Theory* 29(2): 255–67.
- James, Oliver. 2010. Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments. *Journal of Public Administration Research and Theory* 21(3): 399–418.
- Jensen, Didde Cramer. 2018. Does Core Task Matter for Decision-Making? A Comparative Case Study on Whether Differences in Job Characteristics Affect Discretionary Street-Level Decision-Making. *Administration & Society* 50(8): 1125–47.
- Jensen, Didde Cramer, and Line Bjørnskov Pedersen. 2017. The Impact of Empathy—Explaining Diversity in Street-Level Decision-Making. *Journal of Public Administration Research and Theory* 27(3): 433–49.
- Jilke, Sebastian, and Lars Tummers. 2018. Which Clients Are Deserving of Help? A Theoretical Model and Experimental Test. *Journal of Public Administration Research and Theory* 28(2): 226–38.
- Keiser, Lael R. 2010. Understanding Street-Level Bureaucrats’ Decision Making: Determining Eligibility in the Social Security Disability Program. *Public Administration Review* 70(2): 247–57.
- Konrad, Alison M., and Kathy Cannings. 1997. The Effects of Gender Role Congruence and Statistical Discrimination on Managerial Advancement. *Human Relations* 50(10): 1305–28.
- Kunda, Ziva, and Steven J. Spencer. 2003. When Do Stereotypes Come to Mind and When Do They Color Judgment? A Goal-Based Theoretical Framework for Stereotype Activation and Application. *Psychological Bulletin* 129(4): 522–44.
- Laustsen, Lasse, and Alexander Bor. 2017. The Relative Weight of Character Traits in Political Candidate Evaluations: Warmth Is More Important than Competence, Leadership and Integrity. *Electoral Studies* 49: 96–107.
- Lawless, Jennifer L. 2004. Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era. *Political Research Quarterly* 57(3): 479–90.
- Lipsky, Michael. 2010. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*, 30th anniversary ed. New York: Russell Sage Foundation.
- Loyens, Kim, and Jeroen Maesschalck. 2010. Toward a Theoretical Framework for Ethical Decision Making of Street-Level Bureaucracy: Existing Models Reconsidered. *Administration & Society* 42(1): 66–100.
- Maor, Moshe, Sharon Gilad, and Pazit Ben-Nun Bloom. 2013. Organizational Reputation, Regulatory Talk, and Strategic Silence. *Journal of Public Administration Research and Theory* 23(3): 581–608.
- 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.
- Meier, Kenneth J. 2018. Theoretical Frontiers in Representative Bureaucracy: New Directions for Research. *Perspectives on Public Management and Governance* 2(1): 39–56.
- Olivola, Christopher Y., and Alexander Todorov. 2010. Elected in 100 Milliseconds: Appearance-Based Trait Inferences and Voting. *Journal of Nonverbal Behavior* 34(2): 83–110.
- Olsen, Asmus Leth. 2015. Citizen (Dis)satisfaction: An Experimental Equivalence Framing Study. *Public Administration Review* 75(3): 469–78.
- . 2017a. Compared to What? How Social and Historical Reference Points Affect Citizens’ Performance Evaluations. *Journal of Public Administration Research and Theory* 27(4): 562–80.
- . 2017b. Human Interest or Hard Numbers? Experiments on Citizens’ Selection, Exposure, and Recall of Performance Information. *Public Administration Review* 77(3): 408–20.
- Pautz, Michelle C., and Carolyn Slott Wamsley. 2012. Pursuing Trust in Environmental Regulatory Interactions: The Significance of Inspectors’ Interactions with the Regulated Community. *Administration & Society* 44(7): 853–84.
- Pedersen, Mogens Jin, Justin M. Stritch, and Frederik Thuesen. 2018. Punishment on the Frontlines of Public Service Delivery: Client Ethnicity and Caseworker Sanctioning Decisions in a Scandinavian Welfare State. *Journal of Public Administration Research and Theory* 28(3): 339–54.
- Petrocik, John R. 1996. Issue Ownership in Presidential Elections, with a 1980 Case Study. *American Journal of Political Science* 40(3): 825–50.
- Piatak, Jaclyn, and Zachary Mohr. 2019. More Gender Bias in Academia? Examining the Influence of Gender and Formalization on Student Worker Rule Following. *Journal of Behavioral Public Administration* 2(2). <https://doi.org/10.30636/jbpa.22.76>.
- Piotrowski, Suzanne, Stephan Grimmelikhuijsen, and Felix Deat. 2017. Numbers over Narratives? How Government Message Strategies Affect Citizens’ Attitudes. *Public Performance and Management Review* 42(5): 1–24.
- Raaphorst, Nadine, Sandra Groeneveld, and Steven Van de Walle. 2018. Do Tax Officials Use Double Standards in Evaluating Citizen-Clients? A Policy-Capturing Study among Dutch Frontline Tax Officials. *Public Administration* 96(1): 134–53.
- Rencher, Alvin C. 2003. *Methods of Multivariate Analysis*. Hoboken: Wiley.
- 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.
- Ritter, Barbara A., and Janice D. Yoder. 2004. Gender Differences in Leader Emergence Persist Even for Dominant Women: An Updated Confirmation of Role Congruity Theory. *Psychology of Women Quarterly* 28(3): 187–93.
- Sparrow, Malcolm K. 2000. *The Regulatory Craft: Controlling Risks, Solving Problems, and Managing Compliance*. Washington, DC: Brookings Institution Press.
- Stets, Jan E., and Peter J. Burke. 2000. Identity Theory and Social Identity Theory. *Social Psychology Quarterly* 63(3): 224–37.
- Teodoro, Manuel P., and Seung-Ho An. 2018. Citizen-Based Brand Equity: A Model and Experimental Evaluation. *Journal of Public Administration Research and Theory* 28(3): 321–38.
- Thomann, Eva, and Carolin Rapp. 2018. Who Deserves Solidarity? Unequal Treatment of Immigrants in Swiss Welfare Policy Delivery. *Policy Studies Journal* 46(3): 531–52.
- Tummers, Lars L.G., Victor Bekkers, Evelien Vink, and Michael Musheno. 2015. Coping during Public Service Delivery: A Conceptualization and Systematic

- Review of the Literature. *Journal of Public Administration Research and Theory* 25(4): 1099–126.
- Wæraas, Arild, and Haldor Byrkjeflot. 2012. Public Sector Organizations and Reputation Management: Five Problems. *International Public Management Journal* 15(2): 186–206.
- Willis, Janine, and Alexander Todorov. 2006. First Impressions: Making Up Your Mind after a 100-Ms Exposure to a Face. *Psychological Science* 17(7): 592–8.
- Wojciszke, Bogdan, Roza Bazinska, and Marcin Jaworski. 1998. On the Dominance of Moral Categories in Impression Formation. *Personality and Social Psychology Bulletin* 24(12): 1251–63.
- Yang, Kaifeng. 2005. Public Administrators' Trust in Citizens: A Missing Link in Citizen Involvement Efforts. *Public Administration Review* 65(3): 273–85.
- Zacka, Bernardo. 2017. When the state meets the street. Harvard university press.

## Appendix A

**Appendix A.** Decks of Vignettes

Deck #	Vignettes	Domain <sup>b</sup>
1	Police officer <sup>a</sup>	1
	Neighborhood manager	2
	School attendance officer <sup>a</sup>	3
	Nurse	5
	Tax inspector <sup>a</sup>	6
2	Municipal enforcers <sup>a</sup>	1
	Elementary school teacher	3
	Labor inspector <sup>a</sup>	5
	Maternity assistant	5
	Catering industry inspector <sup>a</sup>	6
3	Parking officer <sup>a</sup>	1
	High school teacher	3
	Tram conductor <sup>a</sup>	4
	Youth care worker	5
	Customs officer <sup>a</sup>	6
4	Municipal desk officer	1
	Forest ranger <sup>a</sup>	2
	Bus driver	4
	Education inspector <sup>a</sup>	3
	Physiotherapist	5
5	Emergency medical technician	1
	Bridge operator <sup>a</sup>	2
	Debt counselor	5
	Social investigator <sup>a</sup>	5
	Youth counselor	5
6	Fire officer	1
	Waterways inspector <sup>a</sup>	2
	Social worker	5
	Health care inspector <sup>a</sup>	5
	Wedding officiant	6

<sup>a</sup>Street-level bureaucrats with investigative authority (*buitengewoon opsporingsambtenaar*).

<sup>b</sup>Domains: (1) public space; (2) environment, well-being, and infrastructure; (3) education; (4) public transport; (5) work, income, and health care; (6) general investigation.

## Appendix B

**Appendix B.** Balance Test

	Sex	Ethnicity	Education	Age	Prior Experience	Political Preference
<i>z</i> / <i>t</i> -value	0.90	1.16	1.06	−0.20	2.55	−0.32
B	0.03	0.01	0.02	−0.05	0.05	−0.01
SE	0.03	0.01	0.02	0.27	0.02	0.02
<i>p</i> -value	.36	.25	.29	.84	.01**	.75

Note: A series of (logistic) regressions were used in which sample characteristics (i.e., sex, ethnicity, education, age, prior experience, and political preference) were predicted by the experimental conditions.

## Supporting Information

A supplementary appendix may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1111/puar.13217/full>.