

Which Clients are Deserving of Help? A Theoretical Model and Experimental Test

Final Manuscript for the *Journal of Public Administration, Research & Theory*

Alphabetical author order based on equal contributions

Sebastian Jilke & Lars Tummers

Contact details:

Dr. Sebastian Jilke

School of Public Affairs and Administration, Rutgers University-Newark

The United States

sebastian.jilke@rutgers.edu

Dr. Lars Tummers

School of Governance, Utrecht University, Bijlhouwerstraat 6, 3511 ZC Utrecht

The Netherlands

L.G.Tummers@uu.nl (Corresponding author)

Abstract

Street-level bureaucrats have to cope with high workloads, role conflicts and limited resources. An important way in which they cope with this is by prioritizing some clients, while disregarding others. When deciding on whom to prioritize, street-level bureaucrats often assess whether a client is deserving of help. However, to date the notion of the deserving client is in a black box as it is largely unclear which client attributes activate the prevailing social/professional category of deservingness. This article therefore proposes a theoretical model of three deservingness cues that street-level bureaucrats employ to determine whom to help: earned deservingness (i.e., the client is deserving because (s)he earned it: “*the hard-working client*”), needed deservingness (i.e., the client is deserving because (s)he needs help: “*the needy client*”), and resource deservingness (i.e., the client is deserving as (s)he is probably successful according to bureaucratic success criteria: “*the successful client*”). We test the effectiveness of these deservingness cues via an experimental conjoint design among a nationwide sample of US teachers. Our results suggest that needed deservingness is the most effective cue in determining which students to help, as teachers especially intend to prioritize students with low academic performance and members of minority groups. Earned deservingness was also an effective cue, but to a lesser extent. Resource deservingness, in contrast, did not affect teachers’ decisions whom to help. The theoretical and practical implications of our findings for discretionary biases in citizen-state interactions are discussed.

Keywords

Street-level Bureaucracy, Deservingness, Conjoint experiment, Behavioral Public Administration, Citizen-State Interactions

Introduction

Street-level bureaucrats – such as police officers, social workers and teachers - have a great deal of freedom in making decisions for their clients. They can grant clients a claim, give them a fine, provide them with additional information, or simply offer some extra help. While street-level bureaucrats in their daily interactions with citizens should exercise bureaucratic impartiality, a large literature highlights discretionary biases in interactions between street-level bureaucrats and their clients (Lipsky 1980; Keiser, Mueser and Choi 2004; Schram et al. 2009; Maynard-Moody and Musheno 2003).

Street-level bureaucrats often base their decisions of which clients to prioritize on client attributes that tap the prevailing social/professional category of deservingness (Hasenfeld and Steinmetz 1981; Maynard-Moody and Musheno 2003; Keiser et al. 2004; Lipsky 1980; Tummers 2017). A teacher, for example, would view hard working students - compared to the ones she regards as 'lazier' - as more deserving of help and thus will focus greater attention on them. Similarly, less well performing students may be seen as needy and hence deserving extra help.

A problem with studying the impact of such deservingness cues is that they are hard to disentangle empirically. For instance, a low performing student may be regarded as making only little effort, while this is not necessarily the case. In addition, minority students are often expected to not perform well academically, especially by White teachers (Gershenson, Holt and Papageorge 2016). This shows that it is difficult to decompose the effect of a single client attribute that taps different notions of deservingness. Partly because of these empirical problems, theoretical progress on deservingness cues has stalled. Public administration scholarship discusses the notion of client deservingness (or worthiness) in quite general terms (for instance Ellis 2011).

In this article, we develop a theoretical model that highlights three deservingness cues that activate street-level bureaucrat's notions of deservingness: earned deservingness, needed deservingness and resource deservingness. First, street-level bureaucrats can view clients as deserving because they have earned it through exhibiting high levels of effort. Hence, these so-called '*hard-working clients*' are perceived deserving of help. Second, clients can be seen as deserving because they are in need of help. For instance, clients who are not performing well in terms of bureaucratic success criteria, such as students with bad grades. Street-level bureaucrats regard these clients as deserving of help; these are the '*needy clients*'. The third deservingness cue is related to the resources used to help clients. Already high-performing clients have the highest chance to succeed in terms of bureaucratic success criteria. Such "*successful clients*" may not be 'deserving' in terms of their effort or

need, but deserving in terms of being worthy of investment in a situation of scarce time and limited resources.

However, these different cues of deservingness and their empirical manifestations are often conflated. It is therefore hard to disentangle the impact of deservingness cues and their respective empirical manifestations, such as minority status, work effort or client performance. To appropriately deal with this problem, we implement a conjoint experiment. The conjoint experimental design has been developed in mathematical psychology (Luce and Tukey 1964) and has been applied extensively in marketing (e.g., Churchill and Iacobucci 2006; Green and Rao 1971), and lately also in political science (for example Hainmueller, Hopkins and Yamamoto 2014). A conjoint design allows us to independently randomize multiple client attributes, thereby capturing the causal effect of each client attribute separately.

We designed a unique conjoint experiment to test the impact of three deservingness cues and their empirical manifestations among a nationwide sample of US teachers. We presented teachers various pairs of student profiles and subsequently asked them each time to select one of the profiles they would prioritize in giving extra help. Our results show that teachers intend to prioritize students they perceive as *hardworking*, *low performing*, or being *a member of a minority group*. These cues activate the notions of earned and needed deservingness and hence lead street-level bureaucrats to prioritize these types of clients. Resource deservingness, however, was not influential. In addition, the importance of certain deservingness cues is – to some degree - contingent on personal characteristics of the street-level bureaucrat, particularly gender and race.

This study is structured as follows. First, we discuss the theoretical background on cues of deservingness, client attributes and street-level bureaucrats' prioritization behavior. Based hereon, we develop a theoretical model of three cues of deservingness. We then outline the method for testing the effectiveness of these deservingness cues. We conclude with a discussion and theoretical reflection of the main findings, and develop future research directions for the study of discretionary biases of street-level bureaucrats.

Literature Review

Street-level bureaucracy and discretion

This study focuses on client attributes and how street-level bureaucrats respond to these attributes. This topic is deeply rooted in the street-level bureaucracy literature (e.g., Hill and Hupe 2009; Meyers and Vorsanger 2007). In his book *Street-level bureaucracy: Dilemmas of the individual in public services*, Lipsky (1980) analyzed the behavior of frontline staff in

policy delivery agencies. He refers to these frontline workers as ‘street-level bureaucrats’; public employees who interact directly with citizens and have substantial discretion in the execution of their work. Examples are teachers, police officers, general practitioners, and social workers. Street-level bureaucrats deliver services to citizens, such as providing education, processing welfare benefits or maintaining safety. Related terms include public professionals (Noordegraaf 2015), public service providers (Jensen and Vestergaard 2017) and frontline workers (Maynard-Moody and Musheno 2003).

When delivering services to citizens, street-level bureaucrats have a certain degree of discretion in their work (Brodkin 1997; Brodtkin 2011; Jensen and Pedersen 2017; Hill and Hupe 2009; Olson 2017; Sandfort 2000). Lipsky (1980) sees discretion as the freedom that street-level bureaucrats have in determining the sort, quantity and quality of sanctions and rewards during policy implementation.

To understand street-level bureaucrats’ use of their discretion, we can analyze what types of decisions street-level bureaucrats make when working with clients. In a recent overview article, Tummers et al. (2015) used the notion of ‘coping during public service delivery’ to analyze the decision-making processes of street-level bureaucrats towards clients. They identified nine ways of coping, which resemble different types of behaviors street-level bureaucrats can use towards clients, such as prioritizing some clients over others, routinizing work or working overtime. In this study, we analyze one important way of coping street-level bureaucrats can use: prioritizing clients. Prioritizing during public service delivery is defined as “giving certain clients more time, resources, or energy”. Hence, street-level bureaucrats decide to focus on certain clients, while disregarding others. Prioritizing is an important way of coping in service delivery that happens regularly, among else in education (Maynard-Moody and Musheno 2003; Tummers 2017; Hagen and Owens-Manley 2002). However, prioritizing some clients over others can threaten equal treatment; It can be beneficial for certain clients or client groups, but others may be worse off.

The impact of client attributes on prioritization intentions

How do street-level bureaucrats choose which clients to prioritize? In this study we develop a theoretical model for this by linking classical street-level bureaucracy literature with insights from psychology (see also Grimmelikhuijsen et al. 2017). Street-level bureaucrats can think about clients in terms of belonging to certain social groups. They can use many different categories. For instance, in large disasters physicians prioritize patients who benefit the most from their help. This is referred as triage. In less extreme situations, other criteria may be used, such as whether a client is friendly versus hostile (Sandfort 2000), expected to perform well (Baviskar 2013), or whether clients are also putting in effort themselves (Križ

and Skivenes 2014). However, clients might also be categorized on more general attributes, such as gender or race (Marvel and Resh 2015; Wilkins and Williams 2008).

As discussed in the introduction, we identify three deservingness cues: earned deservingness ('the hard-working client'), needed deservingness ('the needy client') and resource deservingness ('the successful client'). These three cues activate social/professional categories that street-level bureaucrats have about their clients. In other words, these client attributes can act as information cues that tap into the characteristics of a *deserving* client. This idea is closely related to the social psychological concept of social categorization, which refers to the natural cognitive processes by which people place others into social groups (Allport 1954; Tajfel 2010). Gordon Allport in '*The Nature of Prejudice*' writes that "The human mind must think with the aid of categories. [...] Once formed, categories are the basis for normal prejudgements. We cannot possibly avoid this process. Orderly living depends upon it." (1954: 20).

Social categorization is derived from cognitive flexible processing between our two memory systems (Macrae and Bodenhausen 2000; McClelland, McNaughton and O'Reilly 1995). On the one side, humans have general beliefs about the world which are very stable, the so-called neocortical (i.e., slow-learning) system (McClelland et al. 1995). These social beliefs, norms and expectations about the world provide a foundation against which the experienced social world and its stimuli are interpreted. On the other side, there is the so-called hippocampal (i.e., fast-learning) system, which allows new experiences without automatically modifying schematic knowledge (i.e., the neocortical system). When people think about or interact with other people, they commonly use new information to think about them by means of their general beliefs. In other words, people simplify complex processes by using their general beliefs to categorize new people.

Hence, we classify people on the predominant social categories they belong to, and for which we have a range of long-term memory. Attributes that represent such social categories may include race, gender, or wealth, among many others. People use the contents of their activated long-term knowledge about these social categories to form expectations about people, including stereotypical judgements (Macrae and Bodenhausen 2000). Street-level bureaucrats also classify clients in social or professional categories. This helps them to navigate through the complexities of their work (Lavy and Sand 2015; Maynard-Moody and Musheno 2003). Once these classifications are triggered, the contents of the associations they have with this particular group of clients are activated. While people – in our case, students - may pose a multitude of attributes, it is important to figure out which ones trigger (all other things being equal) the activation of the social/professional category of *deservingness*.

Street-level bureaucrats may employ cues to activate the category of a 'deserving client'. Triggering the category of a deserving client and its associated contents may result in being more likely to prioritize (i.e., help) a particular group of clients. This line of reasoning is in line with key tenets of the 'deservingness heuristic' in political psychology (Van Oorschot 2000; Petersen 2015). For example, Petersen et al. (2010) show that an aged man who has been on the labor market all his life is almost three times more likely to get public support to receive social welfare, when compared to a young man without a lot of work experience. This can be seen as *earned deservingness*. In a subsequent experiment, Petersen et al. (2010) show that these effects are automatically activated and indeed caused by deservingness perceptions. Related to this, in a recent study on attitudes of Europeans towards refugees, Bansak et al. (2016) show that humanitarian concerns with regard to asylum seekers' deservingness shape European attitudes toward asylum seekers. For instance, they show that refugees who have been the victim of torture are around 11% more likely to be accepted than are those with no special vulnerabilities. This can be related to *needed deservingness*.

The notion of deservingness is also present in the public administration literature, although it is often framed in terms of worthiness. This is mainly based on the seminal work of Maynard-Moody and Musheno (2003; see also Baviskar and Winter, 2017; Ellis et al. 1999; Harrits and Møller, 2014; Marrow, 2009; Scott 1997). One of the main findings of Maynard-Moody and Musheno (2003) is that street-level bureaucrats – or frontline workers in their terminology – especially help those clients they view are worthy of investment. Related to this, Maynard-Moody and Leland (2000) provide examples of Vocational Rehabilitation counsellors, noting that clients that are seen as worthy by their counsellors receive extraordinary services. Counsellors cut through red tape for them, keep their cases open for longer and even work overtime, for instance coming in on the weekend to help a client move to a new home.

Based on the notion of the deserving client, we argue that street-level prioritization intentions are systematically related to those client characteristics that activate the social/professional category of deservingness in street-level bureaucrats' minds. As a consequence they will more likely aim to prioritize them. We focus on four client attributes which – based on the literature – can be expected to influence street-level bureaucrat's prioritization intentions of clients: the clients' work effort, performance, gender and race. These client attributes can be classified into three groups of deservingness cues: earned, needed and resource deservingness. Clients' work effort is directly related to earned deservingness, because street-level bureaucrats perceive hard-working clients as making an effort themselves (Maynard-Moody and Leland 2000; Tummers 2017). Low performing clients can be related to needed deservingness, as they are typically the ones most in need

of support. High performing clients on the other hand, can be seen as deserving as they have a high chance to succeed in terms of bureaucratic success criteria and investing in them is a good use of scarce resources (resource deservingness)¹. Lastly, we study the impact of client race and gender, both attributes being related to needed deservingness because racial minorities and females are often perceived of being structurally disadvantaged in many areas in life (Lewis 1992; Parker et al. 1997; Connell 2014; Feather 1996). By focusing on three deservingness cues and linking client attributes to these cues, we provide a first understanding of the importance of different types of deservingness cues on street-level bureaucrats' prioritization intentions. Below we lay out theoretical predictions linking client attributes and the prioritization intentions of street-level bureaucrats.

Client effort

First, we expect that street-level bureaucrats aim to help those clients who are making an effort themselves. Studies indicate that the behavior of street-level bureaucrats is strongly influenced by their perceptions of whether clients are making an effort themselves, being 'hard working' (Berglind and Gerner 2002). For instance, Hagen and Owens-Manley concluded that most street-level bureaucrats place "tremendous emphasis on the clients' efforts to help themselves" (2002: 175). This showed workers that these clients wanted to develop themselves. As a result, street-level bureaucrats especially helped such clients. Similarly, Maynard-Moody and Leland (2000: 117) noted that "motivated clients are good clients; they are the truly deserving poor". In addition, Anagnostopoulos (2003: 305) quotes teachers who state that they are "helping students who help themselves" and are "not wasting energy on kids who don't care." Helping motivated clients can also be beneficial for street-level bureaucrats. Tummers (2017) showed that street-level bureaucrats who prioritize motivated clients receive higher performance ratings from their supervisors. Hence, street-level bureaucrats take into account the effort of their clients when determining whom to help. A crucial characteristic is whether these clients are – according to the street-level bureaucrat – 'hard working'.

In a similar vein, psychological research suggests that deservingness perceptions of needy people depend on perceptions that people do not try to free-ride on the work of others (Cosmides and Tooby 1992). Political psychologists also show that effort is a crucial cue of deservingness (e.g., Petersen et al. 2010). Based hereon, we expect that high work effort of clients will lead street-level bureaucrats wanting to prioritize them over those clients who

¹ The concept of resource deservingness is closely related to the notion of client cream-skimming.

make less of an effort. The first hypothesis therefore analyses the impact of the earned deservingness examining the impact of client effort.

H1: Street-level bureaucrats will more likely intend to help clients who show a high effort (i.e., high earned deservingness), than those who show a low effort (i.e., low earned deservingness)

Client performance

The second client attribute we study is the actual performance of clients. Performance and effort are often positively related, but not necessarily. For instance, a bright student can perform well without making much of an effort. It is therefore important to disentangle these two attributes. There are two competing views on whether client performance has a positive or negative effect on client prioritization intentions. Therefore, in the following we will develop two competing hypotheses to predict the effects of client performance on whether street-level bureaucrats intend to prioritize them or not.

The literature on the deservingness bias predicts that street-level bureaucrats will especially help clients who are underperforming. These are the 'needy' clients and therefore street-level bureaucrats intend to especially help such clients. Van Oorschot (2000) found that the public believes that people in higher need (for example those without versus with a job, single income versus double income) should be helped more. In the street-level bureaucracy literature, a similar line of argumentation exists. Goodsell (1981) provides clear examples of street-level bureaucrats helping 'underperforming' clients, noting that bureaucrats often engage in 'positive discrimination'. Social workers, for example, prioritize clients that are definitely in need of help, such as the elderly, disabled, abused children, and refugees. Related to this, in a study of Australian welfare workers, McDonald and Marston (2006) note that such workers often prioritize people who are underperforming and in need of help. In addition, Kelly (1994) reports about a teacher who is giving a poorly performing student special attention and even higher grades, than the student should get based on his academic achievements.

On this basis we argue that low performing clients will be regarded as deserving by street-level bureaucrats. They are perceived as needy clients (i.e., needed deservingness). Street-level bureaucrats' social/professional category of deservingness, we argue, will be triggered by this particular client attribute, making them more likely aim to prioritize low performing clients.

H2a: Street-level bureaucrats will more likely intend to help clients who show low performance (high needed deservingness), than those who show high performance (low needed deservingness).

On the other side, the literature on ‘cream-skimming’ or ‘creaming’ argues that street-level bureaucrats would especially help clients who are *performing well*, or have the potential to perform well. Although often not framed by using concepts like deservingness or worthiness, it can be argued that such clients are deserving as they are worthy of investment because they have the highest chance to succeed (Hasenfeld 1985; Winter 2005). Building upon the works of Lipsky, Winter (2002: 3) notes that “‘creaming’ [is] a concept implying that “street-level bureaucrats often choose (or skim off the top) those clients who seem most likely to succeed in terms of bureaucratic success criteria” (Lipsky 1980), but who might not be the most needy ones.” Creaming occurs when street-level bureaucrats are confronted with more clients than they can optimally help. They will, therefore, tend to choose those clients who are most likely to succeed in terms of bureaucratic success criteria. These are often the high performing clients. Lipsky (1980: 108) notes that “if teachers were assessed by the rate of progress their students made compared to predicted criteria, then the high achievement students will not necessarily be highly valued. But in reality, teachers are judged implicitly by the status and accomplishment of their students and thus seek to teach high achievement classes”. In a similar vein, Koning and Heinrich (2013) examine access to social welfare services that have experienced an increase in performance management practices. They find that service providers were less likely to admit unemployed and disabled workers into performance-based programs. In their recent book on policy implementation, Sandfort and Moulton (2015) reiterate this point and show that performance management systems often result in street-level bureaucrats focusing their effort on high performing clients, from teachers helping the best students to unemployment officers moving easy-to-employ people into employment. Such clients may not be ‘deserving’ in terms of their need, but ‘deserving’ in terms of being worthy of investment in a situation of scarce time and limited resources (i.e., resource deservingness). On this basis, we suggest a competing hypothesis:

H2b: Street-level bureaucrats will more likely intend to help clients who show high performance (high resource deservingness), than those who show low performance (low resource deservingness).

Client race

We also expect that more general attributes like race and gender can serve as empirical manifestations of deservingness cues. Starting with race/ethnicity, we argue that racial minority status (being Hispanic or African-American as opposed to being Caucasian) may activate street-level bureaucrats' social/professional category of deservingness. Racial minorities are argued to be structurally disadvantaged in the allocation of opportunities and resources, especially in the educational system of the United States (Bonilla-Silva 2006). Street-level bureaucrats may therefore regard these minority groups as deserving (i.e., needed deservingness). In line with works on the deservingness heuristic, race is a factor outside the control of clients, thereby making them deserving of help. This hypothesis gets further support from research in the field of educational psychology that puts forward the theory of a positive feedback bias (Harber 1998). It argues that White teachers would be more likely to give more positive feedback to out-group members, which are believed to be structurally disadvantaged. In other words, White teachers would give, for example, African-American students more positive feedback than they would give to otherwise similar White students who submit the same quality work. Experimental research has found evidence for such a positive feedback bias among university students (Harber 1998). Subsequent studies of among else public school teachers (Harber et al. 2012) have replicated the positive feedback bias. This is done to for instance boost student's self-esteem (Thomaes et al. 2013) or to show lack of prejudice (Croft and Schmader 2012). Based hereon we hypothesize that racial minority status serves as a cues of needed deservingness to street-level bureaucrats. Consequently, they will intend to prioritize this particular group of clients.

H3: Street-level bureaucrats will more likely intend to help racial minority clients (African American and Hispanic, high needed deservingness), than White (Caucasian) clients (low needed deservingness).

Client gender

Finally, we examine gender as an empirical manifestation of a deservingness cue. A number of scholars have related gender to client deservingness (Scott 1997; Kullberg 2005; Maynard-Moody & Musheno 2003; Scourfield 2006). Generally, it was found that female clients were seen as more deserving than male clients. For instance, Kullberg (2005) used a vignette study to show that Swedish social workers viewed single fathers with custody of their children as less deserving and more responsible for their situation than single mothers. Related to this, recent studies in education show that female students are often disadvantaged (Arliss and Borisoff 2001), because they are perceived as less able than their

male counterparts, especially in ‘hard’ science courses like math or physics (Shepardson and Pizzini 1992). Furthermore, more general studies on deservingness and gender show that males are often seen as less deserving, as they are more dominant in work settings (Lewis 1992; Parker et al. 1997) and society more in general (Connell 2014; Feather 1996). On this basis we argue that female clients will be regarded as more deserving by street-level bureaucrats as they are perceived to be in more need of help. Street-level bureaucrats’ social/professional category of needed deservingness will be triggered by this particular client attribute, making them more likely intend to prioritize female clients. We therefore formulate the following hypothesis:

H4: Street-level bureaucrats will be more likely intend to help female clients (high needed deservingness) than male clients (low needed deservingness).

Table 1 shows the three deservingness cues, the client attributes they relate to and respective hypotheses.

Table 1 Three deservingness cues

| Deservingness cue | Description | Example | Hypothesis |
|---|---|---|-------------------|
| Earned deservingness (‘the hard-working clients’) | Clients are seen as worthy of investing time and resources in because they have shown high effort. | A student who is seen as deserving of help as she is working very hard every day to learn as much as possible. | H1 |
| Needed deservingness (‘the ‘needy clients’) | Clients are seen as worthy of investing time and resources in because they are perceived to be need of help. | A working and disabled single mother who is applying for extra financial assistance is seen as deserving as she is in need of help. | H2a, H3, H4 |
| Resource deservingness (‘the successful clients’) | Clients are seen as worthy of investing time and resources in because they are perceived to be successful in terms bureaucratic success criteria. | A person who has graduated in engineering and is looking for a first job is seen as deserving of help as this person will likely succeed and therefore the time invested is well spent. | H2b |

Method

Experimental setting

We have chosen to focus on education because prioritizing among students has been documented to be particularly important (Keiser et al. 2004; Maynard-Moody and Musheno

2003; Tummers et al. 2015). Studying teachers' intended prioritization behavior is at the core of bureaucratic discretion, and teachers face considerable freedom in their decisions whether to provide additional help to their students, for example, after class. Not only can teachers decide if they invest extra time and effort in their students, they typically have to prioritize with *whom* they do so. Or as Lipsky (1980: 110) puts it: "Choosing among students who are thought to be more worthy of teacher's time is a way of solving the dilemma of discovering limits to a theoretically unlimited, but practically limited dedication". Indeed, prioritizing can lead to discriminatory practices favoring a group of students over another. Therefore, this is not only a problem of theoretical concern, but also one of great practical importance for the social equity of a nation's educational system. In addition, in times of limited resources discretionary biases are most likely to come into being. Education is a field that has experienced an increase in high-powered performance systems where demands on teachers' resources, including their time, are steadily increasing (Hursh 2013), which is comparable to developments in other service sectors like healthcare, or social welfare (Tummers 2011; Soss, Fording and Schram 2011).

Subject recruitment

The data for this study was collected using an online survey of United States teachers. This was part of larger data collection effort. In February 2015, we have sent an email to a random subset of 7,500 teachers who are members of a major US association of educators, of which only 114 responded. In April 2015, we included a survey link in the association's newsletter. Based hereon, an additional 333 teachers responded. No exclusion criteria for subjects were employed. The total number of survey respondents was therefore 447. From these 447 respondents 349 completed our conjoint experiment, possibly because the preceding survey instrument was relatively lengthy. The subject characteristics of our sample can be found in table 2. No systematic differences between responds from the intimal sample, and those who responded to the newsletter were detected.

Table 2. Subject characteristics ($n=349$)

| | Frequency | Percent | Mean | SE | Min/ Max |
|-------------------------|-----------|---------|------|------|----------|
| Gender | | | | | |
| <i>Male</i> | 58 | 16.9 | | | |
| <i>Female</i> | 286 | 83.1 | | | |
| Race | | | | | |
| <i>White</i> | 285 | 81.7 | | | |
| <i>African-American</i> | 25 | 7.2 | | | |
| <i>Hispanic</i> | 17 | 4.9 | | | |
| <i>Other</i> | 22 | 6.3 | | | |
| Age | NA | NA | 49.4 | 10.1 | 27/ 77 |
| Tenure | | | | | |
| <i>Less than 1 year</i> | 6 | 1.7 | | | |
| <i>1-4 years</i> | 21 | 6.1 | | | |
| <i>5-9 years</i> | 25 | 7.3 | | | |
| <i>10-14 years</i> | 46 | 12.5 | | | |
| <i>15-19 years</i> | 74 | 21.5 | | | |
| <i>20-24 years</i> | 57 | 16.5 | | | |
| <i>25 years or more</i> | 119 | 34.5 | | | |

We studied to what extent our sample resembles basic characteristics of the US national population of teachers. In general, it is fairly similar. The gender composition of the respondent group was 83% female, which is consistent with the American average of teachers (84%) (Feistritzer, Griffin and Linnajarvi 2011). The respondents' average age was somewhat higher than that of the average US teachers (49 against 41). The respondents' race was in line with national averages: 82% Caucasian, 7% African American, 5% Hispanic in the sample versus 84% Caucasian, 7% African American, 6% Hispanic in national averages (Feistritzer et al. 2011). While not perfectly representative to the US population, our sample is sufficiently diverse in terms of teacher characteristics, comparable to the population of US teachers.

While a total of 349 respondents completed the conjoint experiment, our design permits us to treat a single respondent as 6 observations because each respondent had to rate a total of 6 student profiles (of course taking into account the nested structure of the data)² (see also Hainmueller et al. 2014). This yields an analytical sample of 2,088 observations ($6 \times 349 - [4 + 2]$) providing sufficient statistical power to detect any medium-sized effect.

² From the 349 respondents, one respondent dropped-out after responding to the first student pair, and a second dropped-out after responding to two pairs, therefore they represent only 2 and 4 observations respectively.

Experimental design

To test our hypotheses regarding the effects of deservingness cues on prioritization intentions, we showed teachers multiple pairs of student profiles and asked each time to pick the one they would prioritize. Multidimensional choices like prioritization intentions are particularly well suited for conjoint experiments because of their ability to assess multiple causal hypotheses simultaneously. In contrast, conventional survey experiments typically focus on one factor at a time (e.g., the client's gender, or performance). This might be problematic in the context of multidimensional decision-making process in which the causal effect of a single factor may be composed of multiple components – a problem referred to as composite treatment effect (Hainmueller et al. 2014; see also Jilke and Van Ryzin 2017). In the case of street-level bureaucrats' prioritization intentions, for instance, we would need to independently randomize work effort and achievement because they are arguably correlated with each other: a student may be hardworking, but not automatically receiving good grades, and vice versa. Examining this simple relationship in the context of gender and different racial groups may complicate things even further. Without independently randomizing these factors simultaneously, they may mask each other and hence scholars will not be able to distinguish the independent effects of different client attributes from each other. To deal with the problem of composite treatment effects in studying prioritization intentions, we independently randomize all client attributes.

The conjoint experiment we have implemented involves the presentation of multiple pairs of students (see Appendix for the detailed setup). Teachers then had to choose the particular student they would intend to prioritize in giving extra help. Doing so, we asked teachers to respond to a hypothetical but realistic scenario, noting that:

“In this last part of the survey we are interested in your behavior in a hypothetical situation with students. We ask you to carefully read a text about the situation and answer some questions about it. Please try to be honest in answering the questions. Describe what you would really do if a similar situation occurs in your working live. We will show you a total of three brief scenarios. Remember that your answers to these questions will be kept completely confidential.

Please read the descriptions of two students below. They both want your help as they are struggling with a certain question.”

Teachers were then presented with two student profiles (three times in total). The students had various attributes, which were randomized independently. In other words, each respondent had to choose one student for each of three different student pairs. The order of student attributes was randomized across respondents to avoid order effects, but fixed within respondents to avoid confusion (see also Hainmueller et al. 2014).

Four client attributes were independently randomized. For effort, we chose three options: “Hard working student” (high effort), “Not a hard worker, but also not lazy” (average effort) and “A somewhat lazy student” (low effort). For performance, we developed another three options, ranging from very high to very low performance: “Top 5 % of the class”, “Bottom 5% of the class” and “scoring in the middle of the class”. Race was operationalized as follows: “Caucasian/White” (non-minority), “African American/Black” (minority), and “Hispanic/Latino” (minority). We also stated the student’s gender: “Male” or “Female”. To provide greater realism, we also stated the nature of the question of the student, next to these four client attributes. It could be a question about “An important school assignment” or a question about “Feeling alone in class”. There were no substantial differences in the way the respondents answered, therefore we averaged respondents across both question types. In total, we had two options for gender, three options for effort, three for academic achievement, three for race, and two for the type of question. This yields 13 independent criteria and 108 unique combinations. Below we present an example of two student profiles:

| | Student 1 | Student 2 |
|--------------------------|---------------------------------------|--------------------------------|
| Gender | <i>Female</i> | <i>Male</i> |
| Effort | <i>Hard working student</i> | <i>A somewhat lazy student</i> |
| Achievement | <i>Top 5 % of the class</i> | <i>Bottom 5 % of the class</i> |
| Ethnic background | <i>Hispanic</i> | <i>Caucasian/ White</i> |
| Question about | <i>An important school assignment</i> | <i>Feeling alone in class</i> |

To measure our dependent variable (prioritizing intentions), we then asked teachers which of the students they would help: Student 1 or 2. Such a choice-based outcome is commonly used in conjoint analyses. Indeed, recent empirical work has shown that discrete choice specifications like this closely resemble real-world decisions (Hainmueller et al. 2015). More specifically, we asked teachers:

Suppose you can only help one of these students (for instance because you really have a high workload). If you had to choose between them, which of these two students would you help?

Student 1

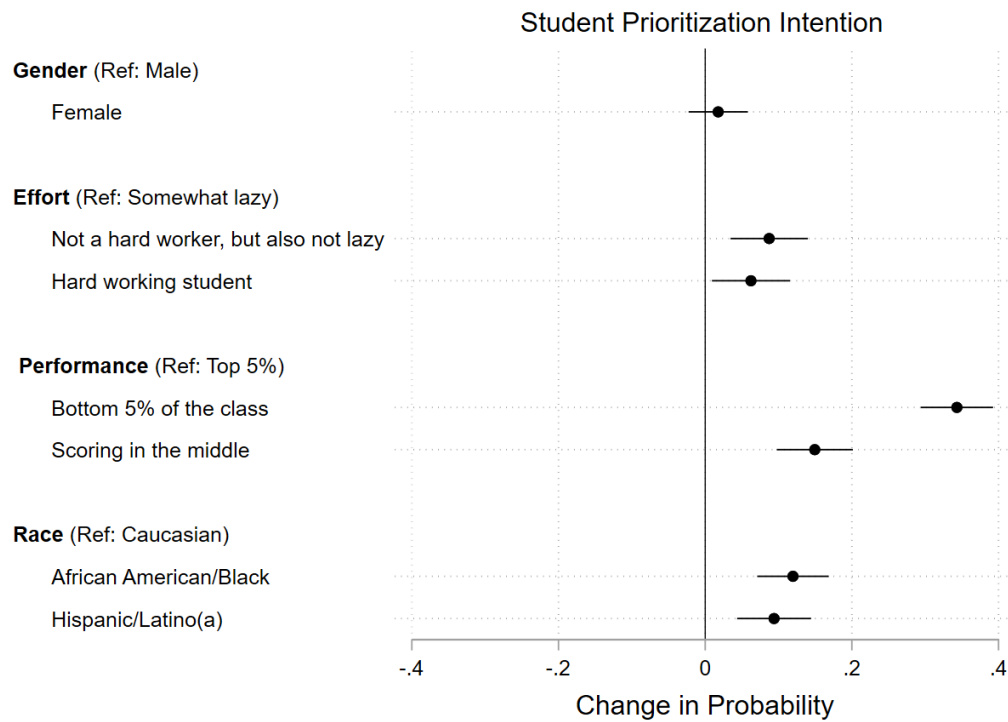
Student 2

Results

Recently, conjoint analyses have been integrated with the so-called potential outcomes framework for causal inference (Hainmueller et al. 2014). Hainmueller et al. formally show that individually randomized attributes within a conjoint specification can be nonparametrically identified by means of a conventional linear regression model. The resulting causal quantity of interest is the Average Marginal Component Effect (AMCE). The AMCE is defined as the marginal effect of each randomized client attribute, averaged over the joint distribution of all attributes. Hence, we can identify the AMCE of each conjoint attribute simultaneously (because they are orthogonal to each other) by estimating a linear regression model with student prioritization intentions as outcome variable. Because each respondent evaluated six students (three pairs consisting of two students each), we need to account for the clustered nature of our data (individuals nested in student-observations). Therefore, standard errors were clustered by respondents. Estimating hierarchical regression models instead yields no substantial differences.

The results of our analysis are shown in figure 1. The dots in the figure represent the regression coefficients (i.e., the AMCEs) from table 1a in the Appendix (including their 95% confidence intervals). The AMCE can be interpreted as the probability of intending to prioritize a student that exhibits a particular characteristic. When the dot plots do not include the black vertical line, they are statistically significant at a 95% confidence level.

Figure 1. Overall results (95% confidence intervals; $n=2,088$)



Hypothesis 1 states that street-level bureaucrats will more likely intend to help clients who exhibit a high effort. Our results support this expectation; teachers are 6% more likely ($p<0.05$) intending to prioritize hard working students, when compared to students that are described as “somewhat lazy”. Similarly, students that are depicted as “not hard working, but also not lazy” are about 9% more likely ($p<0.01$) to be prioritized, when compared to “somewhat lazy students”. This exemplifies that work effort is indeed an important empirical manifestation of earned deservingness, and that street-level bureaucrats take it into account when prioritizing clients.

With regard to performance, we tested two contradicting hypotheses: H2a states that street-level bureaucrats will more likely intend to help clients who show low performance because they regard them as in need of help (i.e., needed deservingness). H2b, in contrast, argues that street-level bureaucrats will cream-skim those clients who perform well in terms of bureaucratic success criteria (i.e., resource deservingness). Our results lead us to reject the resource deservingness hypothesis (H2b). In particular, we find that students who range in the bottom 5% of the class are 34% more likely ($p<0.01$) to be said to be prioritized than those students who are said to range in the top 5%. This is a sizeable effect. Similarly, students who are described as scoring in the middle of the class, are still 15% more likely ($p<0.01$) to be prioritized than those students who are said to range in the top 5%. We can therefore conclude that teachers seem to be especially stating that they would help low performing students.

We also have examined whether teacher's prioritization decisions are influenced by a student's race. Street-level bureaucrats, we argue, perceive racial minority clients as in need. Hence, H3 states that street-level bureaucrats will intend to especially help minority groups (African American and Hispanic) as opposed to White (Caucasian) clients. In particular, teachers that were presented an African American student profile were about 12% more likely ($p < 0.01$) to prioritize this student, when compared to a Caucasian student. Similarly, teachers were 9% more likely ($p < 0.01$) to prioritize a student that was of Hispanic origin.

Finally, we explored the impact of gender. Our results indicate that female students were only about 2% more likely to be prioritized by teachers. This effect was neither statistically significant nor of substantial magnitude. Therefore, we find no evidence that gender serves as a cue of needed deservingness which disproportionally leads teachers to state that they would help female students more often. In other words, we have to reject hypothesis 4.

Additional exploratory analysis

In addition to examining the main hypotheses, we performed additional exploratory analyses to assess the heterogeneity of our results. In particular, we looked at whether teachers' propensity to use certain empirical manifestations of deservingness cues to determine which students to prioritize, is contingent on personal characteristics of the teacher. We therefore tested whether the effects of client attributes vary across the following sub-populations: 1) gender (male versus female teacher), and 2) race (Caucasian versus African-American versus Hispanic teacher). Before turning to the result, a note of caution is warranted given the small sample sizes for some subpopulations, ranging from 102 (Hispanic teachers only) to 1,712 (female teachers only) observations. Results are shown in Table 3. Indeed, all subgroups (except Hispanic teachers, but here we have to keep the relatively small sample size in mind) use client attributes to choose which students to prioritize. In addition, low performance (representing needed deservingness) was the most consistently used attribute, with effect sizes varying between 25% to sizeable 43%.

We first examined whether the use of client attributes varies across gender. We find that men ($n=346$) mainly focus on low performance, while women ($n=1,712$) use the full range of deservingness cues provided to them. Neither of them, however, consider the student's gender. Hence, we did not find support for the assumption that teachers would be more likely to state that they would help students more frequently if they share the same gender.

Next, we looked at teacher's self-reported race. Most interestingly here, we find that African-American teachers ($n=144$) are twice as likely to intend to prioritize minority students, when compared to White teachers ($n=1,656$). Yet, White teachers still state that they would prioritize minority students at about 6 to 11 percent. Hispanic teachers ($n=102$), however, did only take performance and effort into account. In conclusion, both minority and non-minority teachers aim to prioritize minority students, although African-American teachers aim to prioritize minority students to a greater extent than White teachers.

Table 3. Heterogeneous treatment effects

| | Gender of teacher | | Race of teacher | | |
|-------------------------------|--------------------|---------------------|--------------------|--------------------|--------------------|
| | Male | Female | White | African-American | Hispanic |
| Gender (Ref: Male) | | | | | |
| <i>Female</i> | 0.040 (0.047) | 0.011 (0.023) | 0.012 (0.023) | -0.029 (0.066) | -0.037 (0.092) |
| Effort (Ref: Somewhat lazy) | | | | | |
| <i>Hard working</i> | 0.021 (0.070) | 0.073* (0.030) | 0.065* (0.031) | 0.095 (0.104) | 0.217* (0.106) |
| <i>In between</i> | 0.062 (0.065) | 0.0909** (0.030) | 0.100** (0.029) | 0.031 (0.123) | 0.151 (0.117) |
| Performance (Ref: Top 5%) | | | | | |
| <i>Bottom 5% of the class</i> | 0.317** (0.070) | 0.353** (0.027) | 0.338** (0.029) | 0.433** (0.070) | 0.414** (0.127) |
| <i>Middle of the class</i> | 0.137* (0.062) | 0.157** (0.030) | 0.135** (0.030) | 0.146+ (0.079) | 0.236 (0.142) |
| Race (Ref: Caucasian) | | | | | |
| <i>African-American</i> | 0.064 (0.061) | 0.130** (0.028) | 0.111** (0.028) | 0.293** (0.103) | 0.077 (0.138) |
| <i>Hispanic</i> | 0.036 (0.070) | 0.106** (0.028) | 0.064* (0.028) | 0.334** (0.092) | 0.066 (0.102) |
| Constant | 0.274 (0.051) | 0.190 (0.031) | 0.221 (0.031) | 0.087 (0.097) | 0.128 (0.116) |
| R-squared | 0.072 | 0.101 | 0.092 | 0.193 | 0.134 |
| <i>n</i> (observations) | 346 | 1,712 | 1,656 | 144 | 102 |
| <i>n</i> (individuals) | 58 | 286 | 277 | 24 | 17 |

Robust standard errors in parentheses; ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Discussion and conclusions

When delivering services to citizens, street-level bureaucrats have a certain degree of discretion in their work. They can decide to which clients they give extra time, resources or energy. Prioritizing among clients is an important way of coping in service delivery that happens regularly, especially in education (Maynard-Moody and Musheno 2003; Hagen and Owens-Manley 2002). However, prioritizing some clients over others can threaten equal treatment; it can be beneficial for certain clients or client groups, but others may be worse off.

The primary contribution of this study is that it develops a model to analyze on what grounds street-level bureaucrats decide who is deserving of their help. We identified three main deservingness cues: earned, needed and resource deservingness. The deservingness cues are empirically represented by client attributes such as effort, performance, or race. Next to this, we contribute methodologically by using a conjoint design which allows us to disentangle conflated client attributes and show which attributes are effective in impacting prioritization intentions of street-level bureaucrats.

Our empirical study shows that especially needed and earned deservingness play an important role in affecting prioritization intentions, while resource deservingness was not influential. Teachers were significantly more likely to state that they would prioritize students they perceive as hardworking, low performing, or being a member of a minority group. The most important attribute is low performance (i.e., need). Indeed, the importance of aiming to help low performing, hardworking and minority status aligns with prior studies that show that street-level bureaucrats and the general public believe that especially 'needy' and 'hard-working' clients should be helped and that street-level bureaucrats often engage in positive discrimination (e.g., Van Oorschot 2000; Goodsell 1981; McDonald and Marston 2006; Kelly 1994; Harber et al. 2012; Maynard-Moody and Leland 2000; Petersen et al. 2010). It partly contradicts studies that show that street-level bureaucrats 'cream'; helping clients who are performing well, or are expected to perform well (Hasenfeld 1985; Winter 2005).

One of the benefits of our research design is that we were able to disentangle client effort and performance from each other empirically, which are usually strongly correlated. This could – in parts - explain the discrepancy of our results with studies on creaming. However, regarding race, the results are less easily resolved. Indeed, we have to be cautious when interpreting whether the outcomes of our study translate into real-world behavior. In our experimental design subjects do not act within a real classroom, but rather make claims about how they would behave. Or in other words, who they deem deserving of help. Recently, Hainmueller et al. (2014) explicitly studied to what extent conjoint experiments relate to real-world behavior. They find that "effects estimated from the surveys

match the effects of the same attributes in the behavioral benchmark remarkably well” (p. 2395). Specifically, the paired conjoint – the design we used for this article – comes closest to real world behavior (p. 2400) (see also Shamir & Shamir (1995). Hence, our design is generally less prone to social desirability bias as compared to other experimental designs. However, our design still involves a hypothetical situation. It could be that the stated preferences (indicating to be willing to help minority group students more) are not fully aligned with revealed preferences (truly helping minority group students more in real life). Therefore, we would argue that a valuable future research direction is an external validation test of our results using field experiments. Another approach would be to include various ex-ante and ex-post validation checks in surveys (Louviere, Hensher and Swait 2000). Related to this, scholars can conduct non-participant observation where the behavior of street-level bureaucrats is observed in the real world.

More generally, a future research agenda that explores discretionary biases in citizen-state interactions would involve testing the relationship between client attributes and prioritizing of street-level bureaucrats across sectors and countries, as it is to be expected that there will be differences. For instance, in education in the United States race is a very prominent issue with a significant racial achievement gap in student performance. It might be that this attribute is less influential in other countries, or in other sectors. Interesting follow up work would involve theorizing about such differences and test them empirically across different countries/ service sectors.

To conclude, this study provides important insights that help to understand which types of clients street-level bureaucrats intend to prioritize. We aimed to disentangle different client attributes to develop insight into the characteristics of the ‘deserving’ client. We found that street-level bureaucrats especially intend to help *low performing clients*, those who are *hardworking* and *from a minority group* (African-American or Hispanic). These attributes represent notions of earned and needed deservingness. Embracing and further researching this notion of the deserving client using a variety of methods and empirical settings should prove to be a timely and productive endeavor for both researchers and practitioners alike.

NOTES

Author's note

Author order is alphabetical. The authors contributed equally to this article.

Acknowledgement

This paper was presented at the Public Governance and Management Research Day at Utrecht University, School of Governance. We would like to thank the participants for their constructive feedback. We would also like to thank Editor Brad Wright and the three anonymous reviewers. They all provided us with constructive and insightful comments, significantly improving the article.

Funding

Lars Tummers acknowledges funding of NWO grant VENI-451-14-004.

References

- Allport, Gordon W. 1954. *The nature of prejudice*. New York, NY: Doubleday.
- Anagnostopoulos, Dorothea. 2003. The new accountability, student failure, and teachers' work in "urban high schools. *Educational Policy*, 17(3):291-316.
- Arliss, Laurie P., and Borisoff, Deborah. 2001. *Women and men communicating*. Prospect Heights, IL: Waveland.
- Bansak, Kirk., Hainmueller, Jens., and Hangartner, Dominik. 2016. How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science*, 354(6309), 217-222.
- Baviskar, Siddhartha. 2013. Who creams? Explaining the cream-skimming behavior of public school teachers in Denmark from a street-level bureaucracy perspective. SFI Advisory Research Board, Copenhagen, Denmark, 6–7 June.
- Baviskar, Siddhartha., & Winter, Søren C. 2017. Street-Level Bureaucrats as Individual Policymakers: The Relationship between Attitudes and Coping Behavior toward Vulnerable Children and Youth. *International Public Management Journal*, 1-38.
- Berglind, Hans., and Gerner, Kristian. 2002. Motivation and return to work among the long-term sicklisted: an action theory perspective. *Disability and Rehabilitation*, 24(14):719-726.
- Bonilla-Silva, Eduardo. 2006. *Racism without racists: Color-blind racism and the persistence of racial inequality in the United States*. Maryland: Rowman & Littlefield Publishers.
- Brodkin, Evelyn Z. (2011). Policy work: Street-level organizations under new managerialism. *Journal of Public Administration Research and Theory*, 21(suppl 2), i253-i277.
- Brodkin, Evelyn Z. 1997. Inside the welfare contract: Discretion and accountability in state welfare administration. *The Social Service Review*, 71(1):1-33.
- Churchill, Gilbert A., and Iacobucci, Dawn. 2006. *Marketing research: methodological foundations*. New York: Dryden Press.
- Connell, R. W. (2014). *Gender and power: Society, the person and sexual politics*. London: John Wiley & Sons.
- Cosmides, Leda., and Tooby, John. 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.
- Croft, Alissa., and Schmader, Toni. 2012. The feedback withholding bias: Minority students do not receive critical feedback from evaluators concerned about appearing racist. *Journal of Experimental Social Psychology*, 48(5), 1139-1144.
- Ellis, Kathryn. 2011. 'Street-level Bureaucracy' Revisited: The Changing Face of Frontline Discretion in Adult Social Care in England. *Social Policy & Administration*, 45(3), 221-244.

- Ellis, Kathryn, Davis, Ann and Kirstein Rummery. 1999. Needs Assessment, Street-level Bureaucracy and the New Community Care. *Social Policy & Administration* 33: 262–280.
- Feather, N. T. (1996). Domestic violence, gender, and perceptions of justice. *Sex Roles*, 35(7), 507-519.
- Feistritzer, Emily C., Griffin, Steven., and Linnajarvi, Alice. 2011. *Profile of teachers in the US, 2011*. Washington, DC: National Center for Education Information.
- Gershenson, Seth., Holt, Stephen B., and Papageorge, Nicholas W. 2016. Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, 52, 209-224.
- Goodsell, Charles T. 1981. *The public encounter: where state and citizen meet*. Bloomington: Indiana University Press.
- Green, Paul E., and Rao, Vithala R. 1971. Conjoint measurement for quantifying judgmental data. *Journal of Marketing research*, 355-363.
- Grimmelikhuisen, Stephan., Jilke, Sebastian., Olsen, Asmus L., and Tummers, Lars. 2017. Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, 77(1), 45-56.
- Hagen, Jan L., and Owens-Manley, Judith. 2002. Issues in implementing TANF in New York: The perspective of street-level bureaucrats. *Social Work*, 47(2):171-182.
- Hainmueller, Jens, Hangartner, Dominik., and Yamamoto, Teppei. 2015. Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112(8), 2395-2400.
- Hainmueller, Jens, Hopkins, Daniel J., and Yamamoto, Teppei. 2014. Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis*, 22(1), 1-30.
- Harber, Kent D. 1998. Feedback to minorities: evidence of a positive bias. *Journal of personality and social psychology*, 74(3), 622.
- Harber, Kent D., Gorman, Jamie L., Gengaro, Frank P., Butisingh, Samantha, Tsang, William, and Ouellette, Rebecca. 2012. Students' race and teachers' social support affect the positive feedback bias in public schools. *Journal of Educational Psychology*, 104(4), 1149.
- Harrits, Gitte S., and Møller, Marie Ø. 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):447-480.
- Hasenfeld, Yeheskel. 1985. Citizens' encounters with welfare state bureaucracies. *The Social Service Review*, 622-635.

- Hasenfeld, Yeheskel., and Steinmetz, Daniel. 1981. Client-official encounters in social service agencies. In: *The public encounter: Where states and citizen meet*, Charles T. Goodsell (editor), 83, 101.
- Hill, Michael., and Hupe, Peter. 2009. *Implementing public policy: An introduction to the study of operational governance*. London: Sage.
- Hursh, David. 2013. Raising the stakes: High-stakes testing and the attack on public education in New York. *Journal of Education Policy*, 28(5), 574-588.
- Jensen, Didde C., and Pedersen, Line B. 2017. The Impact of Empathy—Explaining Diversity in Street-Level Decision-Making. *Journal of Public Administration Research and Theory*, 27(3), 433-449.
- Jensen, Ulrich T., and Vestergaard, Christian F. 2017. Public Service Motivation and Public Service Behaviors: Testing the Moderating Effect of Tenure. *Journal of Public Administration Research and Theory*, 27(1), 52-67.
- Jilke, Sebastian., and Van Ryzin, Gregg. 2017. Survey experiments for Public Management Research. In James, Oliver, Jilke, Sebastian. and Van Ryzin, Gregg. (eds.). *Experiments In Public Management Research: Challenges and Contributions*. Cambridge: Cambridge University Press.
- Keiser, Lael R., Mueser, Peter R., and Choi, Seung-Whan. 2004. Race, bureaucratic discretion, and the implementation of welfare reform. *American Journal of Political Science*, 48(2), 314-327.
- Kelly, Marissa. 1994. Theories of justice and street-level discretion. *Journal of Public Administration Research and Theory*, 4(2), 119-140.
- Koning, Pierre., and Heinrich, Carolyn. 2013. Cream-Skimming, Parking and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts. *Journal of Policy Analysis and Management*, 32(3): 461-483.
- Križ, Katrin., and Skivenes, Marit. 2014. Street-level policy aims of child welfare workers in England, Norway and the United States: An exploratory study. *Children and Youth Services Review*, 4:71-78.
- Lavy, Victor., and Sand, Edith. 2015. On the origins of gender human capital gaps: short and long term consequences of teachers' stereotypical biases (No. w20909). *National Bureau of Economic Research*.
- Lewis, J. (1992). Gender and the development of welfare regimes. *Journal of European Social Policy*, 2(3), 159-173
- Lipsky, Michael. 1980. *Street-level bureaucracy: Dilemmas of the individual in public services*. New York: Sage.
- Louviere, Jordan J., Hensher, David A., and Swait, Joffre D. 2000. *Stated choice methods: analysis and applications*. Cambridge University Press.

- Luce, Duncan R. and Tukey, John W. 1964. Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1(1): 1-27.
- Macrae, Neil C., and Bodenhausen, Galen V. 2000. Social cognition: Thinking categorically about others. *Annual Review of Psychology*, 51(1): 93-120.
- Marrow, H. B. (2009). Immigrant bureaucratic incorporation: The dual roles of professional missions and government policies. *American Sociological Review*, 74(5), 756-776.
- Marvel, John D., and Resh, William G. 2015. Bureaucratic Discretion, Client Demographics, and Representative Bureaucracy. *The American Review of Public Administration*, 45(3), 281-310.
- Maynard-Moody, Steven., and Leland, Suzanne. 2000. Stories from the front lines of public management. Ed: Brudney, J.F., O'Toole, L. and Rainey, H. *Advancing Public Management*, pp. 109-23. Washington: Georgetown University Press.
- Maynard-Moody, Steven., and Musheno, Michael C. 2003. *Cops, teachers, counselors: Stories from the front lines of public service*. Michigan: University of Michigan Press.
- McClelland, James L., McNaughton, Bruce L., and O'Reilly, Randall C. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419.
- McDonald, Catherine and Marston, Greg. 2006. Room to move? Professional discretion at the frontline of welfare-to-work. *Australian journal of social issues*, 41(2), 171.
- Meyers, Marcia K., and Vorsanger, Susan. 2007. *Street-level bureaucrats and the implementation of public policy*. *The handbook of public administration*, 153-163. London: Sage.
- Noordegraaf, Mirko. 2015. *Public Management: Performance, Professionalism and Politics*. London: Palgrave Macmillan.
- Olson, Jeremiah C. 2016. Race and punishment in American prisons. *Journal of Public Administration Research and Theory*, 26(4), 758-768.
- Parker, C. P., Baltes, B. B., & Christiansen, N. D. (1997). Support for affirmative action, justice perceptions, and work attitudes: A study of gender and racial-ethnic group differences. *Journal of Applied Psychology*, 82(3), 376-389.
- Petersen, Michael B. 2015. Evolutionary Political Psychology: On the Origin and Structure of Heuristics and Biases in Politics. *Political Psychology*, 36(S1): 45-78.
- Petersen, Michael B., Sloothuus, Rune., Stuberger, Rune., and Togeby, Lise. 2010. Deservingness versus values in public opinion on welfare: The automaticity of the deservingness heuristic. *European Journal of Political Research*, 50: 24-52.

- Sandfort, Jodi R. 2000. Moving beyond discretion and outcomes: Examining public management from the front lines of the welfare system. *Journal of Public Administration Research and Theory*, 10(4):729–756.
- Sandfort, Jodi R., Moulton, Stephanie. (2015). *Effective Implementation in Practice: Integrating Public Policy and Management*. Jossey-Bass: San Fransisco
- Schram, Sanford F., Soss, Joe., Fording, Richard C., and Houser, Linda. 2009. Deciding to discipline: Race, choice, and punishment at the frontlines of welfare reform. *American Sociological Review*, 74(3), 398-422.
- Scott, Patrick. G. 1997. Assessing determinants of bureaucratic discretion: An experiment in street-level decision making. *Journal of Public Administration Research and Theory*, 7(1), 35-58.
- Scourfield, J. (2006). The challenge of engaging fathers in the child protection process. *Critical Social Policy*, 26(2), 440-449.
- Shamir, Michal., and Shamir, Jacob. 1995. Competing values in public opinion: A conjoint analysis. *Political Behavior*, 17(1), 107-133.
- Shepardson, Daniel P., and Pizzini, Edward L. 1992. Gender bias in female elementary teachers' perceptions of the scientific ability of students. *Science Education*, 76(2), 147-153.
- Soss, Joe, Fording, Richard and Sanford Schram. 2011. *Disciplining the Poor: Neoliberal Paternalism and the Persistent Power of Race*. Chicago: University of Chicago Press.
- Tajfel, Henri. 2010. *Social identity and intergroup relations*. Cambridge: Cambridge University Press.
- Thomaes, Sander., Brummelman, Eddie., Reijntjes, Albert., and Bushman, Brad. J. 2013. When Narcissus was a boy: Origins , nature, and consequences of childhood narcissism. *Child Development Perspectives*, 7(1), 22-26.
- Tummers, Lars, Bekkers. Victor J.J.M., Vink, Evelien and Musheno, Michael. 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.
- Tummers, Lars. 2011. Explaining the willingness of public professionals to implement new policies: A policy alienation framework. *International Review of Administrative Sciences*, 77(3), 555-581.
- Tummers, Lars. 2017. The Relationship Between Coping and Job Performance. *Journal of Public Administration Research and Theory* 27(1): 150-162.
- Van Oorschot, Wim. 2000. Who should get what, and why? On deservingness criteria and the conditionality of solidarity among the public. *Policy & Politics*, 28(1): 33–48.
- Winter, Søren C. 2002. Explaining street-level bureaucratic behavior in social and regulatory policies. *APSA*, Boston, MA, August 29–September 1, 2002.

Winter, Søren C. 2005. "Effects of Casework: The Relation between Implementation and Social Effects in Danish Integration Policy." Paper presented at the 2005 Research Conference of the Association for Public Policy and Management, Washington, DC, November 3–6.

Online Appendix

Online Appendix for Jilke, S. & Tummers, L.G. (2018). Which Clients are Deserving of Help? A Theoretical Model and Experimental Test, *Journal of Public Administration, Research & Theory*.

Overall results

Table 1. Overall results (Average Marginal Component-specific Effects)

| | Model 1 |
|---|--------------------|
| <i>Gender (Reference: Male)</i> | |
| Female | 0.018 (0.021) |
| <i>Effort (Reference: A somewhat lazy student)</i> | |
| Hard working | 0.062* (0.027) |
| In between | 0.087** (0.027) |
| <i>Performance (Reference: Top 5% of the class)</i> | |
| Bottom 5% of the class | 0.343** (0.025) |
| Middle of the class | 0.149** (0.026) |
| <i>Race (Reference: Caucasian/ White)</i> | |
| African-American | 0.120** (0.025) |
| Hispanic | 0.094** (0.026) |
| Constant | 0.206** (0.027) |
| R-squared | 0.094 |
| N (observations) | 2,088 |
| N (individuals) | 349 |

Robust standard errors in parentheses; ** p<0.01, * p<0.05

Conjoint experiment

Below, we show an example of the setup of the conjoint analysis as used in this study. Note that the client attributes are independently randomized. The order of student attributes was randomized across respondents to avoid order effects, but fixed within respondents to avoid confusion.

Helping

In this last part of the survey we are interested in your behavior in a hypothetical situation with students. We ask you to carefully read a text about the situation and answer some questions about it. Please try to be honest in answering the questions. Describe what you would *really* do if a similar situation occurs in your working live. We will show you a total of three brief scenarios. Remember that your answers to these questions will be kept completely confidential.

Please read the descriptions of two students below. They both want *your help* as they are struggling with a certain question.

| | Student 1 | Student 2 |
|--------------------------|--------------------------------|-------------------------|
| Gender | Female | Male |
| Effort | Hard working student | A somewhat lazy student |
| Achievement | Top 5 % of the class | Bottom 5 % of the class |
| Ethnic background | Hispanic | Caucasian/ White |
| Question about | An important school assignment | Feeling alone in class |

Suppose you can only help *one* of these students (for instance because you really have a high workload). If you had to choose between them, which of these two students would you help?

- Student 1
- Student 2