To be bashed or not to be bashed

A research towards the consequences of meaningless bureaucrat bashing on the (self-reported) performance of public servants.

Source image front page Ten Bos (2016), illustration by Cliff van Thillo.
From left to right: the bureaucrat as the devil, the megalomaniac, the angel, the planner and the muddler (see introduction).

Name: Julia Wesdorp
Student number: 6403174
First supervisor: Barbara Vis
Second supervisor: Sheeling Neo
Date: 22 June 2022

Table of contents

1. Introduction

> *Bureaucrats are devils and angels. Hopelessly megalomaniac, but at the same time planners and muddlers, whom we simply cannot do without.*
>
> (Ten Bos, 2016)

How would citizens look at the bureaucracy in 2022? The way Ten Bos describes it, is probably not the way Max Weber, one of the founding fathers of the bureaucracy theory, had envisioned it. Weber's ideal-type bureaucracy is, amongst other things, characterized by impersonality, written rules, ethical codes and hierarchical divisions to create efficiency and effectiveness (Höpfl, 2006, pp. 10-11; Olsen, 2005, p. 2; Sager & Rosser, 2009, p. 6). However, Weber acknowledged that this ideal-type is not all sunshine and rainbows (Weber, 1951, pp. 190-205; in Höpfl, 2006). The tensions and conflicting demands bureaucrats could face, are not publicly apparent. Furthermore, it goes largely unnoticed and without appreciation within society when the bureaucracy achieves a high degree of efficiency and effectiveness. Indeed, today's researchers endorse Weber's acknowledgement and even describe a negative trend towards bureaucrats within the society, which is not by definition based on objective experiences (Del Pino et al., 2016, p. 1).

Actually, being or called a bureaucrat has a negative connotation these days. An example is the following definition of 'bureaucrat' provided by Oxford Languages (n.d.)*:* '…an official in a government department (…) with <u>procedural correctness</u> at the <u>expense of people's needs</u>' (emphases added). Collins Dictionary (n.d.) describes the bureaucrat as a label of someone: '…if you <u>disapprove</u> them because they seem to follow rules and procedures <u>too strictly</u>' (emphases added). Besides the neutral part of the definitions (e.g. an official in a department), these descriptions give voice to a more negative view of the term bureaucrat or bureaucracy.

1.1 Public servants: the focus of this research

Bureaucrats are thus sometimes negatively portrayed, but who is this bureaucrat? The bureaucrat is, amongst others, a public servant who is occupied with policy formulation, preparation or monitoring at different governmental levels (Borst et al., 2014, p. 28). Examples are municipality clerks and servants at the national ministries. Based on literature insights, we can expect that these public servants are subject to more negative views from the public in contradiction to other types of bureaucrats, such as a nurses (street-level bureaucrat). These negative views are visible within the mass media and research (Zacka, 2017, p. 1). The mass media often display these negative views of the bureaucrat: the lazy and insensitive bureaucrat is frequently part of cartoons and tv-series. This positioning could even strengthen the negative perceptions regarding bureaucrats by enlarging characteristics like inefficacy (Pautz &

Warnement, 2013, pp. 571, 577).[1] Also, researchers have identified a broad range of negative views regarding bureaucrats: because there is no daily interaction and thus a more considerable perceived distance between the public servant and the citizen, the image of the public servant remains vague and can preserve negative views, such as lazy, greedy, corruptive and cold-hearted (Goodsell, 1994; in Van de Walle, 2004, p.3; Willems, 2020, pp. 807-810; De Boer, 2020, p. 539; Del Pino et al., 2016, p. 2. Since I am interested in these negative views regarding public servants, it is valuable to minimalise variation in the direction of the perceptions (either positive or negative) and to focus on the **public servant, whose work centres around policy formulation, preparation or monitoring** (Borst et al., 2014, p. 28). In other words, this research focuses on the most likely case instead of the other groups of bureaucrats (e.g. nurses, teachers) who may receive less bashing. As the empirical literature on this topic is scarce, starting with a most likely case is a good starting point.

In the upcoming chapters, I sometimes use the general term 'bureaucrat', and mention it explicitly when discussing the described public servant. Furthermore, the research context is in the Netherlands, thus the used Dutch term is 'ambtenaar', which covers the broader term bureaucrat.[2]

1.2 Bashing the public servant

What are these negative perceptions towards public servants exactly? Looking at the literature, the concept of *stereotype* comes into play. A stereotype is a shortcut in the human cognitive process and helps to make sense of a new situation, person, or experience, either positively or negatively. This shortcut provides an overview of the most representative characteristics of certain people, such as public servants, and enables a fast decision making process (Willems, 2020, pp. 808-811). This research focuses on the negative views, or in other words the negative stereotypes, public servants experience. More specifically, the concept of bureaucrat bashing is relevant, which refers to the frequent negative framing of bureaucrats and can strengthen or develop the negative profile of the bureaucrats or the bureaucracy; this **bureaucrat bashing** is the focus of this research (Szydlowski et al., 2022, p. 10). There are two reasons for this. First, there are indications that these negative views have substantial negative effects on public servants; for example, studied public managers experience bureaucrat bashing as 'devastating' and feel ashamed for working at the government (Garrett et al., 2006, p. 232). Second, besides the effect on public servants, bureaucrat bashing also has long-lasting memory effects on

[1] See for example the following videos to get an impression of how the negative reputation of bureaucrat is sketched in family and adult movies:
Asterix & Obelix – insensitivity, rigidity, coldness: https://www.youtube.com/watch?v=ZHRGjfEQpy4
Zootopia – lazy, slow, frustration with clients: https://www.youtube.com/watch?v=rR2ZhR-GVEU
Jupiter Ascending – rigidity, rule-worship: https://www.youtube.com/watch?v=VveTsyjFlNA
[2] The term 'ambtenaar' has within Dutch language probably a more neutral undertone than the English term 'bureaucrat'.

citizens, which are triggered when sequential bashing signals are initiated (Caillier, 2018). Because of these possible long-term and intensive effects, it is crucial to take a closer look at this phenomenon.

Because of the frequent use of these stereotypes (or bureaucrat bashing) in daily life, several scholars urge scientific research to look further into this phenomenon (Bertram et al., 2022; Goodsell, 2004, in: Bertram et al., 2022; Marvel, 2016, p. 156). Stereotypes are not understudied, but this is often done from the perspective of the citizen: which stereotypes do the citizens have (e.g. Baldwin, 1990, p. 19; Chen & Bozeman, 2014, p. 556; Hvidman & Andersen, 2015, p. 117; Marvel, 2016, p. 150)?

Nonetheless, it is also essential to study these stereotypes from the public servant's perspective. There are few empirical demonstrations about the consequences of the deep-seated stereotypes (or bureaucrat bashing) for public servants. However, there are indications that the consequences do not turn out well for public servants: the stereotypes can lower or damage job motivation, job satisfaction, working relationships, policy implementation and emotional well-being (Chen & Bozeman, 2014, p. 553; Garrett et al., 2006, p. 237; Zacka, 2017, p. 1). The results of these studies show that bashing towards public servants may have negative consequences in the workplace, but how does bureaucrat bashing impact the *performance* of the public servants? I define performance as a behavioral effort of the public servant within a public organization on job-specific tasks and within the psychological and societal environment (based on Sonnentag et al., 2008, p. 428; Boselie et al., 2021, p. 484).[3] Research shows the importance of motivation and job satisfaction for performance; yet precisely these factors are under pressure when the public servant encounters bashing (Jalagat, 2016, p. 39) This research studies this **relationship** between the experienced bureaucrat bashing and performance.

1.3 Coping with bureaucrat bashing

Based on the above, the bashing of bureaucrats is as an external perceived element for public servants (because the bashing comes from the 'outside', e.g. from media, citizens or politicians) and has possibly considerable consequences on performance. Tummers et al. (2015, pp. 1102-1103) describe that one can experience this external element as a demand or conflict, which can raise the servant's stress-level. To deal with this stress, the public servant copes and use coping mechanisms; behavioral and cognitive efforts the public servants can employ during their work to master, tolerate or reduce these external demands (Tummers et al., 2015, pp. 1102-1103). How does a public servant cope with the harmful effects of bureaucrat bashing, and how does this influence the servants' performance?

[3] To illustrate:

- Performance in the psychological environment -> one could think of the well-being of the public servant, for example his/her job satisfaction or stress-level.
- Performance in the societal environment -> one could think of the impact of the public servant or the public organization on the society, for example by contributing to wicked issues.

By answering this question and studying the earlier described relationship, this research contributes to an existing scientific puzzle, which points to two contradicting theories about how public servants can cope with bashing. As often demonstrated within gender and ethnicity stereotypes, the negative stereotypes can lead to a self-fulfilling prophecy, where the public servant handles in accordance to this stereotype (Aronson, 2002, pp. 280-281; Jussim et al., 2005, p. 89). This can result in lower performance of the servant, where he/she copes with the negative impact of the stereotype (or bureaucrat bashing) that has indirect negative effects for the citizen. Second, the literature points to a possible countereffect, as some research indicates that the effects of bureaucrat bashing are not as intense as sometimes is assumed. For example, Szydlowski et al. (2022, p. 8) and Caillier (2020, p. 695) demonstrate that bureaucrat bashing does not affect the behavior of citizens towards public servants. If the bashing does not affect the way citizens express themselves about public servants, coping is possibly not necessary or is done in a way that is indirectly positive for the citizen. Important to note is that both studies look at the effect of bashing on the behavior of citizens towards public servants and do not directly explore the effect of bashing on public servants (which this thesis does). These two views demonstrate the relevance of understanding how the public servant copes with the negative effects of bureaucrat bashing.

Additionally, the literature identifies several specific coping mechanisms. For example, Tummers et al. (2015, p. 8) describe behavioral and cognitive mechanisms that may occur during interactions with citizens and public servants. An example is social support from colleagues (not during interaction; behavioral). In this thesis, the mechanisms **social support** and **distancing** are studied.[4] The reason for this focus is that both mechanisms study coping that does not occur during direct interactions with citizens and cover the behavioral as well as the cognitive coping mechanisms. Besides, literature shows us that both mechanisms are important job resources for different types of workers, when dealing with the cognitive effects of stressful or uncomfortable situations (for example: Schwarzer & Knoll, 2009, Van Veelen et al., 2020; Kim et al., 2020; Rosen et al., 2022; Cole et al., 2007; Derks et al., 2015). To my best knowledge, these are not yet studied with public servants, which makes it an interesting empirical question whether these coping mechanisms and the possible advantages also apply to the servants.

1.4 Research goal and research question

To conclude, the goal of this research is to gain more knowledge about the consequences of bureaucrat bashing towards Dutch public servants, while formulating, preparing, or monitoring a specific policy.

[4] Social support is defined as '…an exchange of resources between two individuals perceived by the provider or the recipient to be intended to enhance the well-being of the recipient' (Shumaker & Brownell, 1984).
Distancing is defined as '…an individual mobility response whereby group members dissociate from their stigmatized ingroup, to avoid the negative experience of being stigmatized, to reap benefits from being less associated with the ingroup, or to better fit in with a high-status outgroup' (Van Veelen et al., 2020, p. 1090).

Specifically, the research identifies whether bureaucrat bashing affects the public servant's performance and how the public servant copes with these possible impacts of bureaucrat bashing. Based on this research goal, the following research question is formulated:

To which extent does bureaucrat bashing influence the performance of public servants, whose task is focused on policy formulation, policy preparation or policy monitoring and to which extent is this relationship moderated by coping mechanisms?

To answer the research question, the following theoretical and empirical sub questions are developed:

- Theoretical sub questions:
  1. What is theoretically the influence of bureaucrat bashing on the performance of public servants?
  2. What is theoretically the role of the coping mechanisms *distancing* and *social support* in the relationship between bureaucrat bashing and performance?
- Empirical sub questions:
  3. What is the influence of bureaucrat bashing on the performance of public servants?
  4. What is the role of the coping mechanisms *distancing* and *social support* in the relationship between bureaucrat bashing and job performance?

1.5 Relevance

This research is in societal and scientific terms relevant After formulating the research and sub questions, it is also important to look at the relevance of this research, both in societal and scientific terms.

1.5.1 Societal relevance

Researching the effects of bureaucrat bashing is societal relevant. First, this research contributes to the knowledge of stereotypes and its consequences for the public servant. A long-lasting negative view of the public regarding the(ir) government could have considerable effects on the bureaucrat's work environment, especially when these demands, including bureaucrat bashing, exceed the positive aspects of the work environment. As Garrett et al. (2006, p. 232) show, meaningless bashing has serious emotional effects and creates a work environment that hampers high-quality, effective and efficient policymaking or implementation. The individual, emotional impacts are thus closely connected to the public servants' work, meaning that bureaucrat bashing has profound and diverse consequences. Knowing how bureaucrat bashing impacts the work environment also gives us information about how to create an environment that empowers public servants to perform better. For example, when social support helps public servants dealing with the negative effects of bureaucrat bashing, this indicates how the work environment of the public servant is ameliorated by creating, for example, a weekly 'check-in' hour (where employees can chat with each other about their well-being and work) or improve the visibility of the organization's confidant.

Second, researchers state that the negative attitude towards the government is part of a country's culture or political discourse (Van der Walle, 2003, pp. 10, 17; Willems, 2020, p. 821). For example, in the Netherlands, some political parties often speak of an 'overgrown bureaucracy' (NOS, 2021). This research shows the impact of this political discourse and culture on public servants and their performance, which is valuable information especially for the ministry of the Interior and Kingdom Relations (Binnenlandse Zaken en Koninkrijksrelaties).[5] Yearly, the ministry investigates the relevant trends and numbers surrounding their public servants. Besides the factual numbers, my research insights can supplement this documentation through adding, for example, a well-being chapter, where the effects of bureaucrat bashing are centralized to give a more accurate view of the bureaucratic work. Besides the added value for the ministry, the insights of the research help create awareness within society about the effects of the bureaucrat bashing and the relevance of high performance within the public sector. As Osborne (2020, p. 1) states, the development and execution of public policy are crucial for the quality of public life. If the experienced negative effects of meaningless bashing hamper job performance, this also indirectly affects the life of the citizens; for example, what if the possible side-effects of a new policy regarding unemployment benefits are not fully explored, so that those, who are the exception to the broader rules, are excluded?

1.5.2 Scientific relevance

In addition to solving the earlier described puzzle, this study is also scientifically relevant because it brings a new perspective to the literature on stereotypes and performance. First, the literature often studies stereotypes or bashing towards public servants from the perspectives of citizens, media or politicians (e.g. De Boer, 2020; Willems, 2020; Marvel, 2015; Doring & Willems, 2021; Del Pino et al., 2016; Garrett et al., 2006). However, the full puzzle is not yet complete by focusing on one part. This research does it the other way around, namely from the perspective of the public servant, which is (as far as is known) not studied as much as the perspective of the citizen. For example, Garrett et al. (2006, p. 228) state that the knowledge about the possible impacts of bureaucrat bashing is limited. Studying especially the impacts or effects of bureaucrat bashing is valuable and not so much how to change them, as it is impossible to require from citizens to change their opinion about the public sector (Bertram, 2022). Therefore this research contributes to a broader understanding of the effect and impact of bureaucrat bashing by adding the perspective of the public servant.

[5] This does not mean, however, that the existing stereotypes are necessarily inaccurate; the goal is to show the consequences of some of the common stereotypes within society and politics.

The next chapter describes the relevant theories and concepts, resulting in several hypotheses. Then, after explaining how the research is done (methods), the results are described, which give some interesting non-significant results of bureaucrat bashing on performance. Finally, I discuss the insights from a broader perspective in the discussion and conclusion.

2. Theoretical framework

This part of the research defines the theoretical concepts within the research question. Besides, this chapter describes the theoretical relationship between the variables, resulting in two hypotheses.

2.1 Working within the public sector

Many different occupations fall under the heading of bureaucrat. Nurses, teachers, policy makers, public managers and police officers: all are public workers. However, some of these workers are grouped by the characteristics of their work. The first group is the street-level bureaucrat. Lipsky (2010, p. 1) describes these bureaucrats as interacting face-to-face with a citizen, while delivering a service or executing a policy and having considerable discretionary room. The literature indicates that the street-level bureaucrat is often seen more positively (De Boer, 2020, p. 538; Willems, 2020, p. 818). Recall that this is not the case for all bureaucrats, especially for the second group of bureaucrats: the public worker focused on policy preparation, formulation or monitoring, who are in terms of Borst et al. (2014, p. 28) public managers. However, Peters (2009, pp. 10-16) sees the public manager as a specific type of public worker and also identifies other roles, such as policy-maker and negotiator; all together called public servants. Because the term 'public manager' has different meanings, the term public servant, next to street-level bureaucrat, is more appropriate. Since this study focuses on the effects of negative views, it is valuable to focus on the group of servants for whom it is most likely to receive these negative stereotypes and minimalise variations within these perceptions (either positive or negative). Therefore, the theoretical framework is narrowed down to the public servant.

2.1.1 The work of the public servant

What characterises the work of the public servant? Here, the literature on policy analysis serves as a basis. One of the ways to look at the policy-process is as a two-phased dichotomy: policy-making and policy implementation. These two phases are sequential and strongly divided, which match to a certain extent with the above described difference. For example, if the public servant finishes the policy-making, the street-level bureaucrat starts with the implementation (Marume et al., 2016, p. 87). However, this two-phased rational view of the policy process does not give a full image of the diverse roles of the public servant (Bovens et al., 2017, p. 113). Think, for example, of the servant who works at the Dutch Central Bureau for Statistics: he/she falls within the policy making phase by delivering the necessary information for the development of the policy, but it is also possible to think of a role within policy implementation, where he/she studies the implementation effects of specific policies. In other words, a strict dichotomy between the making and implementation of policy seems not to work within this research.

Therefore, we should have a more diverse view of the policy process, such as the policy cycle of Harold Lasswell. His cycle distinguishes several phases: agenda setting, formulation, adoption, implementation, evaluation and feedback (Knill & Tosun, 2017, p. 353). Skok (1995, p. 326) describes that this model is useful for explaining the actions and practices of relevant actors (in this case, the public servant). Table 1 provides an overview with the content of the phases and an example of a possible role for the public servant. This is relevant as the table shows how broad the tasks of the public servants are and thus where the bureaucrat bashing may hamper the servant's performance. Note that the role of the public servant is diverse and thus may be bigger or smaller depending on the context; besides, the focus is here on the public servant, not the street-level bureaucrat.

Table 1

Overview of phases, definitions and examples

Phase	Definition	Example – role of public servant
1. Agenda setting	Recognition of problem, which needs action from government.	Public servant of Netherlands Environmental Assessment Agency researches the quality of surface water and finds out that the quality is bad. He/she publishes the research and urges the ministry to intervene.
2. Formulation	Creation of policy: which actions should be taken?	Public servant at ministry Agriculture, Nature and Food Quality looks, by order of governor, at possible actions and prepares a concept policy.
3. Adoption	Selecting the policy alternative; decision making within government institutions.	Public servant at the ministry acting as an adviser to governor: what is the most feasible option?
4. Implementation	Brings the chosen policy into practice.	Most of the time street-level bureaucrat, but in this case it is possible to think of a project manager working at a regional water authority.
5. Evaluation & feedback	Has the implemented policy achieved the previously determined goals, gives the implementation rise to new problems?	Public servant of Netherlands Environmental Assessment Agency researches the quality of surface water again to find out if the quality is improved. He/she develops points for improvement.

Source: own compilation based on Bovens et al., 2017, p. 111; Knill & Tosun, 2017, p. 353; Brommel, 2010, pp. 55-62.

2.1.2 The other side of the work: constraints and difficulties

To give a complete view of the public servant's work, one should also look at the constraints and difficulties servants face when doing their work. This is import, because in this way we get better insights into the current work environment of the public servant and therefore understand in which context the bureaucrat bashing takes place. Public servants experience, for example, decreased budgets, new technological developments, conflicting goals, conflicting values or politicization, which can challenge their work (Needham & Mangan, 2016, p. 265; Hertwig, 2017, p. 152; Suzuki, 2020).

These constraints and/or difficulties are either external or internal demands. To illustrate: the demand to create a high-quality policy with decreased budgets can be an external demand; dealing with conflicting personal and occupational values can be seen as an internal demand (Tummers et al., 2013, p. 4). The conflicts between these demands, the insufficient resources of the servants to deal with these demands and the already existing workload, can add up to an increased stress-level for the public servant. Research indeed shows the existence of a high stress-level among public servants: psychological stress is high, leading to cognitive tiredness (Borst & Blom, 2021, p. 129). Additionally, De Simone et al. (2016, p. 1570) complement the emotional effects with work-related consequences: lowered job satisfaction, higher absence and lower performance.

2.2 Bashing towards the public sector

Now that the work of the public servant is explored, it is important to look at the bureaucrat bashing that also characterizes the work in the public sector. After this, we zoom out to create a broader framework surrounding bureaucrat bashing.

2.2.1 Bureaucrat bashing

Recall that the views towards bureaucrats are either positive or negative. However, within bureaucrat bashing the content is often negative. Marvel (2015, p. 211) describes bashing as ambiguous negative views, which do not serve a function or a useful goal. However, other authors identify multiple forms of bashing, namely substantive and meaningless bashing (Garrett et al., 2006, p. 229; Szydlowski et al., 2022, p. 11). The last one corresponds with the view Marvel (2015) has on bashing, and an example of this type is: ‘Policy makers are lazy and annoying.’ In this example, it is clear that the bashing does not serve a function or goal and is based on subjective arguments (Garrett et al., 2006; Caillier, 2018, p. 206). Besides this functionality-aspect, Jahan & Shahan (2012, p. 275) mention the personality-aspect of meaningless bashing as characteristic: attacking the public servant personally. Considering this, meaningless bureaucrat bashing entails a general statement that does not have a specific goal or function and attacks a public servant personally.

Next to meaningless bashing, there is so-called substantive bashing. This entails negative content that focuses on specific problems and possible solutions (Szydlowski et al., 2022, p. 11). This is a broadly accepted definition (as far as is considered): other authors also demonstrate that substantive bashing entails a specific problem and a way to fix this problem (Caillier, 2018, p. 206; Jahan & Sahan, 2012, p. 275). Looking at these characteristics, we can ask ourselves if substantive bashing is a synonym for just having a critical view of the public sector and therefore if it is appropriate to call it ‘bureaucrat bashing’. Yet, Garrett et al. (2006, pp. 229, 236) demonstrate that both forms are ‘real’ bashing, because they have a deplorable, negative tone and foster hostility and/or demoralization.

To conclude, bureaucrat bashing is a negative view towards bureaucrats, which is either meaningless or substantive. The results of Garrett (2006, p. 229) show that meaningless bashing has more harmful consequences than substantive bashing. This because meaningless bashing entails a personal attack with no specific solution to fix the problem, which the public servant can negatively appeal to him/herself (in contrast to substantive bashing). Because the effects of meaningless bashing are probably more intense and *'...far more damaging…'* for the public servants, the research focuses on this type of bashing (Garrett et al., 2006, p. 229).

2.2.2 Zooming out: bureaucrat bashing and stereotypes

As Caillier (2018, p. 206) states, bashing the bureaucrat focuses on 'shorthand symbolism', where pre-established images, or stereotypes, play an essential role (Van der Walle, 2003, p. 2). Stereotypes can cause bashing, but the bashing can also strengthen the stereotypes (Cailler, 2018). Because stereotypes are a crucial factor within meaningless bashing, it is important to take a closer look at this concept.

So, what is a stereotype? As Willems (2020, p. 808) puts it, a stereotype is a summary within someone's cognition to process information. Spencer (2016, p. 281) affirms this and sees a stereotype as a cognitive view to think in an easy way about people. However, these definitions do not say something about the extent of accuracy of a stereotype: Allport (1958, in: Van de Walle, 2003) describes stereotypes as an '…overgeneralized belief…'. Although this type of belief bases itself on either reality or fiction, Bordalo et al. (2016) assume that stereotypes always have some sort of reality by pointing to the *kernel-of-truth* hypothesis, which makes them valuable. An example is 'Dutch are tall', which is an overstated generalization, but to an extent it describes a real characteristic. Of course, not all Dutch citizens are tall, but on average Dutch citizens are the tallest citizens in the world (BBC News, 2015). In this way some stereotypes have aspects that accurately reflect certain elements of the empirical reality, but still need a critical view as they also contain elements that do not reflect the reality (Bordalo et al., 2016). However, many experts (e.g. sociologists) criticise this kernel-of-truth hypothesis and its applicability should also be critically reviewed within the context of the public sector. Especially regarding bureaucrat bashing, research shows that the stereotypes towards the public sector do not provide a complete and reliable image of the public sector (Garrett et al., 2006; Goodsell, 2014). Because of this questionable applicability, this research does not further use this characteristic. Connected to this, Inzlicht & Schmader (2012, pp. 4-6) point to the impact a stereotype can have: even when the stereotype is invalid, but the stereotyped person believes it is the reality of the sender, it could create the fear of acting according to this stereotype (stereotype threat).

Taken this together, it is possible to identify the following characteristics of stereotypes:

- stereotypes serve as a way, a source or a summary within the human cognition;

- it is a short and fast way to process information about groups of people within society, i.e. it is socially constructed;
- the images within the cognition can be neutral, positive or negative;
- stereotypes impact the receiver, especially if he/she believes their actions or behavior are judged according to this stereotype by others.

The definition of Bos et al. (2018, p. 2) fits to a great extent to these identified characteristics. It allows for different emphasises and contexts: 'Stereotypes are a set of beliefs a person holds about the personal attributes of a group of people. They are descriptive (what group members are like), prescriptive (what group members should be like) and proscriptive (what group members should not be like).' This research uses this definition. This is because this definition captures the diverse characteristics of stereotypes: it describes the underlying process of stereotypes, the socially constructed side (by holding a set of beliefs, which a person constructs) and shows how citizens use them (e.g. to show how the public servant should be like). Also, the definition leaves space to describe the impact of the stereotypes: by emphasising what group members should not be like, one may create a situation where the public servant is afraid of confirming certain stereotypes.

2.3 Understanding individual decision making and behavior

So far, the chapter has explored bureaucrat bashing and stereotypes, but why do we use stereotypes? To answer this question, this theoretical framework takes a small side trip to describe the mechanisms behind individual decision making and behavior. The origin of stereotypes is not the focus of this thesis, but it is important to understand these underlying mechanisms to get a better grip on the used theories and hypothesized relationships.

2.3.1 Rational choice theory and bounded rationality

The rational choice framework describes individual decision making and assumes that individuals are fully rational. This means that one has complete knowledge of the context and alternatives, where the individual rationally chooses the alternative that reaches the goal in the best manner (Bellé et al., 2018, p. 828; Shannon et al., 2019, p. 3). However, several authors criticise this perspective; for example Shannon et al. (2019, p. 3) describe the impossible character of some assumptions: being fully rational demands extremely high effort. Next to theoretical difficulties, the theory also has empirical difficulties: within different real-life contexts, the rational choice is not convincing (Jones, 1999, p. 297).

Instead of assuming full rationality, Herbert Simon stated that individuals are bounded rational already during the fifties (Bellé et al., 2018, pp. 828-829). Because of our external and internal environment, individuals are bounded in their decision making and deviate from what a fully rational person would do (Jones, 1999). Think for example of time pressure within the external environment (i.e. coming from outside) and the individual's limited attention within the internal environment (i.e. originating from the individual) (Shannon et al., 2019, p. 8; Simon, 2000, p. 25; Shannon et al., 2019, p. 6). These factors are not unknown for public servants; for example, Shaxson (2005, p. 105) describes that policy-makers often face pressures to deliver results in the short-term.

How do people make decisions with these limitations? The answer is 'heuristics', and we arrive almost at the stereotypes. Several researchers describe a heuristic as a rule of thumb, which can help make sense of a problem and search for a solution (Del Campo et al., 2016; Shannon et al., 2019). As these researchers mainly focus on what a heuristic can do, Gigerenzer (2010, p. 535) looks at the trade-off that occurs: one uses a heuristic to save a lot of effort or time, but this comes at the expense of accuracy (accuracy-effort trade-off). For example, using the first available information is a time-saving way to collect information, but impedes a thorough analysis. Summarizing, the literature differs in its focus on heuristics, but it agrees on the time-dimension: it is a short-cut, which can save time.

2.3.2 Heuristics and stereotypes

The stereotyping of public servants can also be seen as a decision heuristic and resembles to a great extent the representativeness heuristic (Willems, 2020, p. 890; Stolwijk, 2019, p. 1).

The founding fathers of the representativeness heuristic, Daniel Kahneman and Amos Tversky, describe this as a short-cut where individuals review the probability regarding the extent to which X is similar to Y and thus the extent to which X is a representation of Y. As shown by Gigerenzer (2010, p. 535), the heuristic enables fast decision making, but often comes at the expense of accuracy: if an individual concludes that X represents Y, this does not necessarily mean that X and Y are indeed the same (Stolwijk, 2019, pp. 1-3). This is also the case with stereotypes towards public servants. If someone must assess new information about the capabilities of a public servant, it is, given the bounded rationality, impossible to analyse all the information. Therefore, the citizen evaluates the extent to which this new information is similar to the existing stereotypes. As a result, the chance that a public servant resembles all the other lazy bureaucrats is high, although it is not necessarily an accurate judgement (Stolwijk, 2019, pp. 6-9).

Up to this point, the chapter has unravelled the independent side of the research question; however, to understand the different mechanisms within this research fully, the focus moves to the dependent variable, performance.

2.4 Performing on the job

The performance on the job of public servants is frequently studied, but, remarkably, many researchers do not define performance (such as Kumari & Pandey, 2011; Wright & Davis, 2001; Leisink & Steijn, 2010). This makes it difficult to get a grip on their perspectives of the concept performance and to replicate the study. Therefore, it is crucial to take a closer look at this concept to create a definition, which creates transparency for other researchers (replicability) and readers.

Motowidlo & Kell (2003, pp. 92-93) define job performance as an expected value for the organization of individual behavior. Expected value, however, remains somewhat vague: how does the organization calculate this value and how is the individual behavior translated into a value? Following this reasoning, the definition of Viswesvaran & Ones (2000, p. 216) is more practical: the actions or behavior of workers to perform on organizational goals. The literature frequently points to this behavioral aspect of performance. For example, Koopmans et al. (2011) use different types of behavioral performance to construct a performance-framework. Despite these differences, both definitions see performance as something purely for the organization's advantage. This contrasts with the view on performance of Boselie et al. (2021, p. 484), who state that performance can also pertain to the well-being of the individual servant and the society.

This brief literature insight shows that job performance contains diverse elements, and it is important that the definition covers this diversity. One possible way of defining job performance is to integrate the distinction of Sonnentag et al. (2008, p. 428) between task (organizational performance) and contextual (contributes to the organizational, social and psychological environment) performance. The following definition is formed by myself, based on Sonnentag et al. (2008, p. 428) and Boselie et al. (2021, p. 484): 'Job performance is individual behavior within a certain public organization, which covers the performance on job-specific tasks and the performance within the psychological and societal environment.'[6] This combination covers the important behavioral focus and sees performance not something that is purely achieved within the organization (not only job-specific tasks, but also regarding the society for example), which fits with the public sector. Therefore this research uses this definition.

2.5 Coping with the perceptions of citizens

Until now, the framework has elaborated on the independent (meaningless bashing) and dependent (performance) variable; now, the focus is on the possible moderating variable: coping.

One of the first ways of defining coping is done by Folkman & Lazarus (1988, p. 466). The authors describe coping as a way to deal with emotional threats, which subsequently creates a behavioral reaction. For example, a public servant, who feels that he/she is stereotyped as slow and rigid, can complain to the manager (behavioral reaction) about his/her experienced emotional pressures to deal

[6] To illustrate, performance in the…

- …psychological environment -> e.g. the well-being of the public servant, for example job satisfaction.
- …societal environment -> e.g. the impact of the public servant or the public organization on the society, for example by contributing to wicked issues.

with them. Even though the article originates from 1988, this way of behavioral coping is still relevant within the literature. For example, Steenhuisen & Van Eeten (2012, p. 1135) see coping as a behavioral response towards an external stressor, which can also be an emotional threat. Lazarus & Folkman (1984, in: Baker & Berenbaum, 2007) identify this as 'emotion-focused coping', which focuses on diminishing or mitigating the effects of emotional threats and exists next to problem-focused coping (strategies to tackle, solve or modify the problem). However, these definitions are somehow limited as they mainly focus on preventing a negative outcome (because the coping mechanisms can also create positive outcomes, such as compassion) and external demands (because internal demands, such as perfectionism, may also need management) (Nielsen, 2006, p. 863; Steenhuisen & Van Eeten, 2012). Additionally, this behavioral way of coping is often supplemented with a cognitive way of coping, which also enables the individual to manage, regulate or control the threats one faces (Garnefski et al., 2002, p. 603). Examples are self-blaming, cynicism, and emotional distancing (Tummers et al., 2015, p. 8).

A few characteristics of coping are now evident. First, coping is a behavioral/cognitive response to certain external or internal demands/conflicts. Second, these behavioral or cognitive responses are often identified as 'ways of coping' or 'coping mechanisms'. With these mechanisms, the individual tries to manage, reduce, regulate, or control the demands or conflicts. Third, according to Tummers et al. (2015, p. 1106), it is possible to make another distinction within coping: it may take place during client-worker interactions. This research focuses on public servants, where we assume daily interactions with citizens or clients are not self-evident. This does not mean that they do not have any interaction (e.g. interactive governance is nowadays a frequently used), but it is not part of the regular daily tasks of public servants and therefore, this research focus on ways of coping *not* during interactions. A suitable definition that has the abovementioned characteristics is the definition of Tummers et al. (2015). However, it is slightly adapted to the context of this research, which results in the following definition of coping (in italics changed, respectively: added cognitive and changed 'frontline workers employ when interacting with clients'): 'behavioral and *cognitive* efforts *public servants employ during their work*, in order to master, tolerate or reduce external and internal demands and conflicts they face on an everyday basis.' This research uses the following definition, because it shows that coping is a behavioral or cognitive mechanism, and it describes for which purposes coping can be used; all elements that are relevant if we look in the literature.

2.6 Relationship between bureaucrat bashing, performance and coping

So far, this chapter has described bureaucrat bashing, performance and coping individually, but how can we theoretically describe the relationship between these variables? This paragraph further investigates this question.

2.6.1 Bureaucrat bashing and performance

How can we theoretically connect meaningless bashing towards public servants with their performance? The concept *stereotype threat* is, amongst other things, the answer to this question. Spencer et al. (2016, p. 416) describe the stereotype threat in its consequences: a negative stereotype about a person, creates extra pressure for this stereotyped person. This pressure originates from the fear of being seen or being handled negatively, the fear of acting according to stereotypes (own-reputation threat) or the fear of confirming the stereotype surrounding the whole public sector (group-reputation threat) (Shapiro, 2011, pp. 465-467).

Earlier, in paragraph 2.3, the cognitive process behind the stereotypes are discussed, but what are the underlying cognitive or psychological mechanisms behind this stereotype threat? In this case, the concept of feedback is relevant, especially constructive or non-constructive feedback. Constructive feedback is either positive or negative and characterises itself by agreed behavioral or performance norms. This type of feedback can motivate one to reach the set norms or go beyond these norms (Bee & Bee, 1998). This matches with substantive bashing: by showing the problem and giving a solution that points out how the performance norm is reached, one can see this type of bashing as a constructive form of feedback. Besides, the opposite, non-constructive or destructive feedback, also has some parallels with meaningless bashing. As Bee & Bee (1998, p. 3) state, this type of feedback is often based on generalization with a focus on mistakes made by an individual, and thus behavior or performance is not judged according to agreed norms, as is also the case with meaningless bashing.

This non-constructive form of feedback can also elicit a stereotype threat and has several negative impacts on the individual. It could, for example, strengthen the concern of being judged according to certain stereotypes (Inzlicht & Schmader, 2011; Richeson & Shelton, 2011, p. 235; Cohen et al., 2011, pp. 282-286). Therefore the non-constructive feedback itself is not that troublesome, but the psychological reaction to this feedback is. This is because the psychological reaction uses different cognitive assets to deal with the threat, which the individual cannot use for the his/her performance. Furthermore, the threat can limit or disturb the influence of certain incentives within the individual or the work environment (e.g. possibility to excel), which normally boost the performance (Cohen et al., 2011, p. 286). This all makes the stereotyped person more vulnerable for confirming a stereotype, known as the self-fulfilling prophecy: a situation where a stereotyped expectation of the source creates a behavioral reaction with the target, that confirms this stereotype (Jussim et al., 2005, p. 87). Specifically, it is in this case about the possible (un)conscious *fear* of fulfilling the stereotype: what if people think that I am lazy (own-reputation threat) or what if my behavior negatively stereotypes the behavior of all my colleagues (group-reputation threat)? In other words, the threat, that the stereotype or meaningless bashing creates, may arouse the displacement of critical cognitive resources, making it more challenging to perform (Spencer et al., 2016, p. 417). For clarification, see figure 1.

Figure 1

Illustration of working of stereotype threat in individual's cognitive process

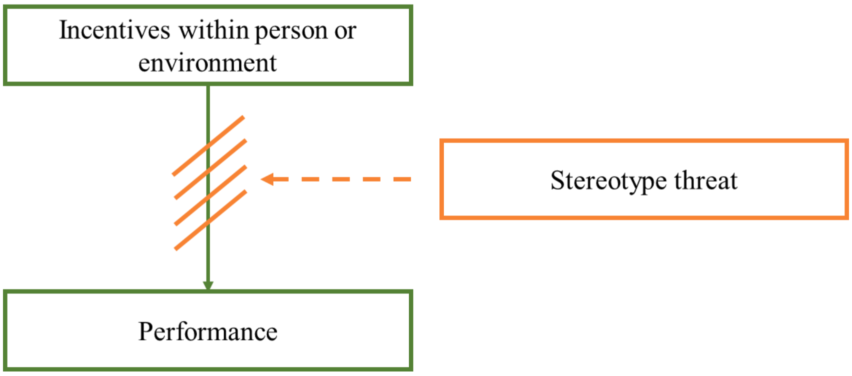


Illustration based on Cohen et al. (2011, p. 282)

Empirical research also identifies these displacing effects on performance. For example, Aronson (2002, p. 283) describes that through the extra pressure a stereotype threat creates, performance is lowered. Additionally, several authors describe that the stereotype's threatening effects hamper the cognitive and the psychological performance (Nguyen & Ryan, 2008, p. 1326; Hively & El-Alayli, 2014). Important to note is that often gender stereotype threats are the focus of these studies, which may have a different functioning than meaningless bashing. Besides the factual performance, Marvel (2016, p. 155) also concludes that the stereotypes regarding the public sector, lead to lowered expectations regarding performance with citizens.

Thus, taking the above into account, experienced meaningless bureaucrat bashing can arouse a stereotype threat with the public servant, which displaced important cognitive resources (because they are used to cope with the threat). This makes it harder to perform on the job with possible lowered performance as a consequence. This reasoning results in the following hypothesis and also answers sub question 1 (what is theoretically the influence of bureaucrat bashing on the performance of public servants?).

Hypothesis 1: public servants, who are exposed to meaningless bureaucrat bashing, will display lower performance than those who are not exposed to meaningless bureaucrat bashing.

2.6.2 Moderation of the coping mechanism distancing

The question is then which coping mechanisms one uses to deal with the (negative effects of) bureaucrat bashing. According to the literature, *distancing* is a frequently effective mechanism to ease negative effects (Van Veelen et al., 2020, p. 1089). For example, by thinking that the stereotype does not apply to you or by dissociating from the stereotyped ingroup, the public servant can reduce the threat of the stereotype. Also, street-level bureaucrats use this mechanism: by distancing from the 'other' women (queen-bee phenomenon) or by denying that the gender stereotype exists, the female police officer dissociates from the stereotype and therefore experiences the effects of the stereotype less (Derks et al., 2015, pp. 1243-1248). Napier et al. (2020) show the benefits of this, as it could improve the well-being and satisfaction of the person. We can apply this same reasoning to the context of this research: if one dissociates him or herself from the bureaucrat bashing, the possible negative experienced effects of the bureaucrat bashing on the performance are also probably less. For example, by thinking that the lazy-stereotype does not apply to you, one can protect him or herself (and probably his or her performance) from the negative effects of meaningless bashing. Based on the above, distancing is thus defined as '… an individual mobility strategy to cope with a threatened identity' (Veldman et al., 2020, p. 118). This definition fits to a great extent the above stated literature, but it is rather limited regarding what this coping mechanism can do for someone. Therefore, the research uses a similar, but more broader definition of Van Veelen et al. (2020, p. 1090): '…an individual mobility response whereby group members dissociate from their stigmatized ingroup, to avoid the negative experience of being stigmatized, to reap benefits from being less associated with the ingroup, or to better fit in with a high-status outgroup.'

According to the literature above, the expected relationship between bureaucrat bashing and performance is less strong, when the public servants uses the coping mechanism distancing. Recall that distancing is a cognitive way of coping. This is because one distances him/herself with the use of cognitive assets and not by a behavioral reaction to diminish the effect of the stereotype. Based on this, the following hypothesis is formulated and answers partially sub question 2 (what is theoretically the role of the coping mechanisms *distancing* and *social support* in the relationship between bureaucrat bashing and performance?):

Hypothesis 2a: the influence of meaningless bureaucrat bashing on performance is negatively moderated by the coping mechanism distancing.

2.6.3 Moderation of the coping mechanism social support

Next to cognitive distancing, it is possible to identify a behavioral coping mechanism that can function as a moderator. Tummers et al. (2015, p. 1106) mention, for example, *social support* from colleagues. For example, by talking and socializing with other colleagues, the public servant experiences if other colleagues also feel threatened or colleagues can reassure the public servant not to worry too much. This can reduce the possible negative effects of bashing.

What is empirically known about this possible relationship? Bakker & Demerouti (2006, pp. 309-317) show how social support acts as a buffer for experienced demands at work and decreases work-related stress. Social support can thus be an essential job resource when dealing with the negative effects of bureaucrat bashing. Cole et al. (2007, p. 608) also demonstrate this role for social support. They additionally demonstrate that providing social support has positive effects on the performance, when someone experiences a stereotype threat. Also, during high-pressure times, collegial support has positive effects. During the Covid-19 crisis, healthcare workers highly valued the collegial support through special *Resilience Coaches* (i.e. program developed for healthcare workers during the Covid-period to improve the resilience and well-being of employees) and created a strong connection between the workers (Rosen et al., 2022, pp. 84-86). To my best knowledge, the effect of social support is not yet studied with public servants. Therefore, it remains an empirical question if these insights also apply to the public servants, which this research answers. The frequently reported positive effects of social support can work in the same way with public servants and may prevent negative effects from, for example bureaucrat bashing, materializing on the performance. Considering this, a suitable definition of social support, used in this research, is that of Shumaker & Brownell (1984): '…an exchange of resources between two individuals perceived by the provider or the recipient to be intended to enhance the well-being of the recipient.' This definition is appropriate for this research, because it points to the necessary exchange between the stereotyped person and the colleague. It also shows the positive consequences of social support for the individual, such as the psychological well-being.

In other words, when a public servant experiences bureaucrat bashing and seeks as a consequence social support, this weakens the negative effects of bashing on the performance. This results in hypothesis 2b and partially answers sub question 2 (what is theoretically the role of the coping mechanisms *distancing* and *social support* in the relationship between bureaucrat bashing and performance?):

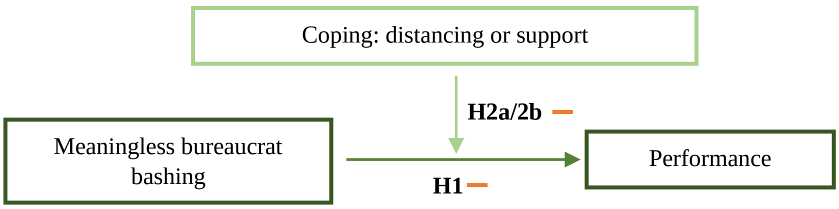
Hypothesis 2b: the influence of meaningless bureaucrat bashing on performance is negatively moderated by the coping mechanism social support.

2.7 Theoretical conceptualization

Figure 2 summarizes the theoretical framework and the formulated hypotheses.

Figure 2

Overview expected relationships – hypotheses 1 (H1) and 2 (H2a/2b)



Note. Orange minus indicates the expected effect, i.e. a negative relationship or weakening effect of moderator.

3. Methods

This part describes the methods used in this research and how the data are analysed. First, this chapters sets out the design and data collection. After this, the operationalisation of the concepts is described. Lastly, the chapter discusses the limitations, reliability and validity of the research. In appendix 1 the control and intervention vignettes are given; the full text of the survey is in appendix 2.

In line with open science practices, the research has been preregistered and is available via https://osf.io/nbm2d.

3.1 Why an experimental design?

This research follows an experimental design, where I intervene within the real world (Jilke et al., 2015, p. 69). There are several reasons for this choice. First, this design decreases the chance of endogeneity (in contrast to a design that observes, e.g. cross-sectional design).[7] By intervening (instead of observing) in the life of the public servant and by manipulating a specific variable (bureaucrat bashing), we can detect if there exist causality between experienced meaningless bashing and performance. In this way, it is possible to understand the direction of the relationship between bureaucrat bashing and performance (simultaneity). Besides, because the research randomly assigns public servants to either the control or the intervention group, the risk of influencing third variables that are not included in this study is to a great extent controlled (self-selection) (Bouwman & Grimmelikhuijsen, 2015, pp. 110-111; Blom-Hansen, 2015). Recall that the research question investigates a possible cause-and-effect relationship, where we want to find out whether there is a causal relationship. As Bouwman & Grimmelikhuijsen (2015, p. 111) state, experimental designs are the designs that are suitable for these kinds of questions.

The experiment is a survey with vignettes (see appendix 1 and 2) and is a between-subjects design, where the control and intervention group receive different, but only one, treatment (Keren, 2014, p. 257). The experiment has a basic design, as it manipulates one independent variable (in contrast to a factorial design with multiple independent variables) (Bouwman & Grimmelikhuijsen, 2015, p. 113).

From a philosophy of science perspective, this research uses a deductive way of thinking. A deductive approach means that one extracts hypotheses and concepts from the theory. The known theory serves here as a guide and provides the input for the research (see chapter 2) (Bryman, 2016, pp. 690-691).[8]

---

[7] Endogeneity occurs when a variable has a connection (such as correlation) with the error term of the tested model (Blom-Hansen et al., 2015). This problem could have several causes, such as: a third variable (confounding variable), which is not included in your test of the hypothesis, influences the dependent and independent variable. This could be the case with self-selection, when individuals themselves can choose whether they want to belong to the control or the intervention group. Here, individual characteristics, e.g. motivation, could play as a third variable influencing your results. Another possibility is that it is possible to indicate causation between the independent and dependent variable, but it is unknown how the causation is constructed (problem of simultaneity), e.g. does X causes Y or Y causes X?

[8] The opposite of deduction is induction, where one tries to create a theory with the results of the study; something that is not done in this research (Bryman, 2016, pp. 690-691).

3.2 Data collection

The upcoming paragraphs describe the sample size, way of sampling and data collection methods.

3.2.1 Sample size

Before getting deeper into the data collection, the focus is first on the sample size. How many respondents are 'enough' respondents? There are several ways to calculate the necessary sample size. When one wants to do a regression analysis, Field (2009, p. 222) uses the following rule of thumb '50 + 8*k* = sample size', where *k* the number of independent and control variables are. This is a rough rule of thumb, but it is also possible to calculate it more precisely by looking at the trio of sample size, probability level and power-level.[9] If we indicate the preferred power-level and probability level, the program GPower (version 3.1) calculates the necessary sample size. This is done with an alpha level of .05, a power of .80, and an effect size of .15 (= medium effect size, Cohen's f^2), which are recommended by Field (2009, p. 58) and described by Cohen (1992, p. 156), and eight predictors (independent variable, two moderators, three controls, two interaction variables). These characteristics make that the minimum sample size is 109 (Field, 2009, p. 58). Also, a 10% oversampling is included (109 * 1.10 ≈ 120).

Noteworthy are the risks connected to a too small sample size. First, as Hackshaw (2008, pp. 1142-1143) argues, a too small sample size can lead to false positives. This happens when the research detects the effect of bashing on performance among, for example, ten public servants, but when the study is replicated in larger samples the effect is gone (type I error). However, a small sample size is more frequently associated with a type II error. Recall that sample size, power and probability level are closely connected. Consequently, if the sample size is small, the power will be low and the chance of a type II error increases (e.g. if the power is .50, this means that the chance of a type II error is .50, which is high) (Columb & Atkinson, 2016). This is the case when the research concludes that bureaucrat bashing does not affect the performance of 10 public servants, while in reality the bashing does have an effect.

3.2.2 How is the data collected?

Recall that this research performs an experimental research, specifically an experimental web survey (Blom-Hansen, 2015). The web survey enables the researcher to reach a bigger group of public servants and is developed within the program *Qualtrics*.

---

[9] Sample size = a small part of the full population, which is used to investigate several concepts to say something about the full population (if this sample is representative)
Probability level = also known as significance level, the chance that a type I error occurs; often $p < .05$
Power = the capacity of a test to identify a relationship with a certain effect; calculated by 1 – *chance of type II error (β)*; often $\beta = .20$, thus power .80.
All definitions are composed by myself and are based on Field (2009), Columb & Atkinson (2016) and Cohen (1992).

First, the survey is spread via the alumni (graduated students) of the Public Governance and Management department of the University of Utrecht (USBO) on a closed LinkedIn group. Second, several government organizations are contacted within the Netherlands and asked about the possibility of setting out the survey (either via mail or telephone). By using the register of *Overheid.nl*, governmental organizations are randomly chosen by giving them a number and letting Excel choose a number. The governmental organizations that are included are municipalities (e.g. Den Bosch), provinces (e.g. Friesland), waterboards (e.g. Rivierenland), umbrella organizations (e.g. VNG), ministeries (e.g. Volksgezondheid, Welzijn en Sport), agencies (e.g. RIVM) and independent administrative bodies (zelfstandig bestuursorgaan, e.g. AFM). In the first round, 50 public organizations are contacted; when an organization did not want to participate, a new organization was selected randomly. In total, 70 public organizations are contacted (of which only seven accepted the request). Third, the survey is also spread via LinkedIn and is shared by several people (all relevant public organizations or public employees).

A clarifying text, where the survey is introduced and the intended research population is described, accompanies the link to the survey. The goal of this is to ensure that the envisaged/targeted public servants of this research indeed get the survey.

3.2.3 Way of sampling

The choice to spread the survey amongst the alumni panel and via LinkedIn is not a way of probability sampling, but convenience sampling. The selected sample is based on nearness or availability (Bryman, 2016, p. 187). This is not an ideal sampling manner, but I believe that a diverse group of public servants is reached in this way. Important to note is that this limits the generalizability of the findings. The sampling of the governmental organizations is done by selecting first the relevant organizations and from this selection, organizations are chosen in a simple random manner (stratified random sampling) (Bryman, 2016, pp. 177-178). To check if the sample is representative of the population, the research performs a Chi-squared test with the variable gender and working organization (49,0% women; 51,0% men; national level = 126.986, provincial level = 11.050, municipal level = 153.668, waterboard level = 10.498; all based on Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2020).

This representativity test shows that for both variables the sample is a representative reflection of the population. Both variables have a *p*-value, which is higher than the used alpha level ($p < .05$) and thus indicates that the observed data in the sample do not significantly differ from the expected data, based on the population (resp. χ^2 (1, 182) = .07, p = .798 ; χ^2 (3, 182) = 2.38, p = .498). [10] To my best knowledge,

[10] Both are based on *Trends & Cijfers 2020* of the ministry of Internal Affairs (https://kennisopenbaarbestuur.nl/rapporten-publicaties/trends-cijfers-2020/).

there is no recent mean available regarding the number of working years, so there is no representative test done for this variable.

3.3 Operationalization of variables

This part of the method chapter operationalizes experienced bureaucrat bashing (independent), self-reported performance (dependent), social support (moderator), distancing (moderator), and the control variables. The independent, dependent and moderator variables are tested in an Exploratory Factor Analysis (EFA). An EFA tests whether the items of a specific construct (e.g. items in bureaucrat bashing scale) all measure the same underlying factor (thus: meaningless bureaucrat bashing), which is particularly important because of the translated questions (as the survey is in Dutch for accessibility reasons). In addition, to check the internal reliability of this research, the constructs are also tested using the Weighted Omega. The upcoming paragraphs describe the results for each variable; the full results, including the reliability scores of the same constructs, are in appendix 3.

3.3.1 Bureaucrat bashing

This research manipulates the independent variable experienced meaningless bashing within vignettes. As seen in part 2.2.1, meaningless bashing should entail a critical statement and personal insults a public servant. Therefore, the vignette varies on the personal approach (no personal attack versus personal attack against the public servant). Both vignettes have some criticism, but differ on the personal aspect of the problems within the organization. Information from the vignettes originates from a factsheet of the European Commission (2017). Looking at the given definition of stereotypes in 2.2.2, the vignettes contain primarily descriptive (e.g. public servants are inefficient and lack knowledge) elements and some prescriptive (e.g. these characteristics are unacceptable and thus it need to change) aspects. The vignettes do not fully cover prescriptive and proscriptive elements to keep the vignette simple, short and manageable for participants.

At the beginning of the survey, every participant gets the same introduction text, which describes the goal of the research and notifies the reader of the upcoming news article (introduction text is based on Kevins & Vis (2021)). Subsequently, the control group receives a news article with no bashing; the intervention group receives a news article with bashing. The full vignettes are in appendix 1 and not displayed here. Also, a pilot check is done with the vignettes. The vignettes are sent to five students within the USG (USBO) department. They are asked in person about their feelings and readability of the vignette, based on this feedback the vignettes were slightly adapted (after the preregistration).

After the vignettes, the survey continues. To measure the extent of experienced bashing, this research uses the construct of Szydlowski et al. (2022, p. 17), which is, to my best knowledge, the first and only scale of this concept. The scale measures the statements on a five-point Likert scale (0 = not at all; 1 =

a little bit; 2 = fairly; 3 = quite a bit; 4 = very much) and consist of five items, which are transformed to the context of public servants, namely:

- the extent to which public servants are criticized;
- the extent to which public servants are attacked;
- the extent to which public servants are blamed;
- the extent to which public servants are generalized;
- the extent to which public servants are criticized in a meaningless way.

Results from the factor analysis show that all the items have the same underlying factor and the construct has a high reliability (Ω = .90), see appendix 3 for the complete results.

3.3.2 Performance

In the literature, we see broadly two ways regarding the measurement of performance. The first is to survey supervisors of the studied employees (see for example Tummers (2015)). The second way is to ask the employee to assess their performance, also known as self-reported performance. This research uses the self-reported performance. However, performance rated by the supervisor may give a more reliable view of the public servant's performance. Yet, this is not achievable within the available time. Besides, this does not fit with the scope of this research, as it would require, for example, a specific intervention in the daily work of the public servant within one specific organization. This is because this research focuses on the individual perceptions of the public servants: how do you perceive the effects of bureaucrat bashing? This individual perspective on the public servant makes it more suitable to use a self-reporting scale. Additionally, if more researchers use the self-reported construct, it enables comparisons between studies (Van Loon et al., 2018, p.142).

Looking at the given definition of performance in paragraph 2.4, the construct must measure a certain behavior and must see performance from a broader perspective. A possible way to measure this concept, is the construct of Carlos & Rodrigues (2015), which measures task and contextual performance. This scale consists of 29 items. The negative side of this scale is that it creates a very long survey, which could lead to lower response rates (Galesic et al., 2009). However, the bright side of this long construct is that we can create an in-depth view of the effect on the public servant's performance. Vandenabeele (2009) has developed a shorter scale, which also measures performance from a more broader perspective. The author mentions that performance is a diverse concept and contains more than only task performance (e.g. performance within psychological environment). Because of the fit with the definition and the scale length, this research uses the items of Vandenabeele (2009). The scale is loosely and partially validated and calibrated by Vandenabeele (resp. χ^2 (2, N = 3506) = 1.01, GFI = 1.000; $RMSEA$ = 0.000, CFI = 1.000 , NFI = 1.000; χ^2 (5, N = 3506) = 3.90, GFI = 0.999; $RMSEA$ = 0.000, *CFI* = 1.000 , *NFI* = 0.999)[11] and is multiple times used in research regarding the public sector (e.g. Sarnacchiaro et al., 2019; Bayram & Zoubi, 2019, both have not tested validity again). My factor analysis shows that all items measure the construct self-reported performance and have good reliability of .79, see appendix 3 for the complete results. An example of an item is *I think I am a good employee*, and the scale measures the items on a five-point Likert scale (1 = Strongly disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly agree), with in total four items.

3.3.3 Coping

The public administration literature often studies coping among street-level bureaucrats. Therefore the developed scales are mainly focused on interactions with citizens, which makes these scales hard to transform to the context of a public servant (Tummers, 2016; Tummers & Musheno, 2015). Another way to measure coping, which is frequently used within psychological and medical research, is the *Ways of Coping Questionnaire* of Folkman & Lazarus (originally from 1985, but updated in 2016). This questionnaire consists of 66 questions. However, it is not necessary to use all the questions as my research focuses only on the coping mechanisms of distancing and social support, which are also covered in the questionnaire and match with the used definitions.

This research uses the items of both constructs, which consist each of six items. The items are adapted to the situation of the public servant. Folkman & Lazarus (2016, p. 3) have tested their scale with a Principal Factor Analysis; I have tested the scale again with an Exploratory Factor Analysis. [12] In the analysis of Folkman & Lazarus one item of distancing had a low loading on distancing and is therefore excluded in this research (*I went along with fate, sometimes I just have bad luck*). However, my factor analysis shows that another item also has a possible different underlying factor than all the other items: *I would try to look on the bright sight of things*. This item reflects more 'positivity' instead of 'distancing'. After deleting this item, the construct of *distancing* has a good reliability of .83. The items of the construct *social distancing* all have the same underlying factor, and the construct item has excellent reliability of .91. An example of one of the items is *I went on as if nothing had happened*, and

[11] N = sample size;
GFI = goodness of fit index, analyses whether the obtained data resembles the data expected from the model, >.95 threshold;
RMSEA = root mean square error of approximation, analyses the resemblance of a hypothesized model and a perfect model, <.06 means a good fit, this scale meets this condition;
CFI = comparative fit index, analyses the fit between hypothesized and the null model, resembles NFI to a certain extent, .90 threshold;
NFI = normed fit index, analyses the difference in the Chi-squared test between the hypothesized and the null model, .90 threshold.
According to Vandenabeele (2008), each of the statistics has its strengths and weaknesses, thus to conclude something about the fit of the model, it is important to look at all these statistics (Xia & Yang, 2018). Based on this information, the scale of Vandenabeele (2009) is validated and provides a good fit.

[12] Results of the principal factor analysis (scale available in appendix 2):
- Distancing – factor loadings of .55, .54, .50, .49, .34, .25.
- Social support – factor loadings of .73, .68, .58, .57, .56, .45.

the scale measures the items on a four-point Likert scale. Both scales have five items (0 = not used; 1 = used somewhat; 2 = used quite a bit; 3 = used a great deal). For the complete results of the factor analysis, see appendix 3.

3.3.4 Control variables

Last, the research incorporates some control variables. One of the control variables is *gender*. This is because we can expect different results in how a stereotype (such as meaningless bashing) is experienced. Gender stereotypes are frequently studied, and the studies often show negative consequences for amongst other things the performance of women (see for example Spencer et al., 2016, p. 417; Ertl et al., 2017). Johnson et al. (2012) even conclude that men perform *better* when they experience the stereotype threat. Although this research uses an experimental design and the risk of a confounding variable is low, it is relevant to incorporate gender to ascertain that this variable does not have any influence (and thus eliminate third influences) (Bryman, 2016, p. 46). Besides, Caillier (2018, p. 210) also incorporates *the years of work experience* to study bureaucrat bashing. This is also valuable within this research, as research on caseworkers shows that the years of work experience positively affect how the employee can cope with conflicts or difficulties at work (Baviskar & Winter, 2017). This mechanism can also function within the relationship between bureaucrat bashing and performance, making it a relevant control variable. A public servant with more years of experience, may also have more experience with the bureaucrat bashing and how to cope with it. Finally, the research also wants to gain insight into *where the respondents work (on which level of the government)* for descriptive use.

When using these control variables, it is important to look at their measurement scale. Interval scales are perfectly suitable for a regression analysis because of their quantitative characteristics. However, gender, working organization and years of working experience (answer possible in four categories) have a lower measurement scale, which is unsuitable for multiple regression analyses. A solution is to create dummies of these variables. Dummies are transformed numerical variables and make it possible to analyse these variables. This research uses the following reference categories (and thus make it possible to incorporate in the regression analysis) (Allison, 1999, pp. 9-10):

- gender: man is the reference category and coded as 0;
- years of working experience: the first category (between 1 and 5 years) is the reference category and coded as 0;
- working organization: ministry is the reference category and coded as 0.

To end, a manipulation check is added. Kane & Barbaras (2019, p. 234) advise using a manipulation check to ensure that the intervention manipulates the respondent. This research does this by incorporating a *factual* manipulation check (see 3.3.4). This a question, which has a correct answer and thus makes it possible to check the extent of attentiveness with the respondent towards the vignette (Kane & Barbaras, 2019, p. 234). The following manipulation check is constructed: *Within which policy* *domain is the governmental organization working?* A multiple-choice answer follows with the following possibilities: social (correct), environmental, education and infrastructure. Another possibility is to use attention checks, specifically instructional manipulation checks. This is a check which asks the participant to fill out a specific answer and thus to check their extent of attentiveness (Hauser et al., 2018). However, because public servants choose themselves if they want to participate in the research, it is assumed that this entails a form of positive self-selection, whereby people are interested and put effort into the survey. Therefore, because of this assumed positive self-selection and to keep the survey manageable, this research only checks the effect of the manipulation with the help of the factual manipulation check.

Table 2 gives an overview of all the definitions and operationalizations of the used variables.

Table 2

Overview of definitions and operationalizations

Concept	Definition (DEF) and operationalization (OP)
Bureaucrat bashing	**DEF:** Negative views towards bureaucrats, which can be meaningless (general statements, which serve no function or goal) or substantive (specific statements, which offer a problem and a possible solution) (own compilation). **OP:** -
Meaningless bureaucrat bashing	**DEF:** A general statement, which does not promote a certain goal or function and attacks a public servant personally (own compilation). **OP:** Scale of 5 items from Szydlowski et al. (2022): • the extent to which public servants are criticized; • the extent to which public servants are attacked; • the extent to which public servants are blamed; • the extent to which public servants are generalized; • the extent to which public servants are criticized in a meaningless way.
Performance	**DEF:** Job performance can be seen as individual behavior within a certain public organization, which covers the performance on job-specific tasks and the performance within the psychological and societal environment (formed by myself and based on Sonnentag et al. (2008) and Boselie et al. (2021). **OP:** Scale of 4 items from Vandenabeele (2009): • in my opinion, I contribute to the success of the organization; • I think I am performing well within this organization; • I think I am a good employee; • on average, I work harder than my colleagues.
Coping	**DEF:** Behavioral and cognitive efforts public servants employ during their work, in order to master, tolerate or reduce external and internal demands and conflicts they face on an everyday basis (Tummers et al, 2015). **OP:** -
Coping mechanism – distancing	**DEF:** An individual mobility response whereby group members dissociate from their stigmatized ingroup, to avoid the negative experience of being stigmatized, to reap benefits from being less associated with the ingroup, or to better fit in with a high-status outgroup (Van Veelen et al., 2020, p. 1090). **OP:** Scale of 5 items from Folkman & Lazarus (2016): • I would made light of the situation and refused to get too serious about it; • I would go on as if nothing had happened; • I would not let it get to me and refused to think much about it; • I would try to forget the whole news article;

		• I would try to look on the bright sight of things.
Coping mechanism – support	**DEF:** An exchange of resources between two individuals perceived by the provider or the recipient to be intended to enhance the well-being of the recipient (Shumaker & Brownell, 1984). **OP:** Scale of 5 items from Folkman & Lazarus (2016): • I would talk to a colleague to find out more about experiences of others; • I would talk to a colleague who could do something concrete for me about the negative effects I experience (e.g. supervisor or confidant); • I would ask a near colleague, which I respect, for advice; • I would talk to a colleague about how I am feeling; • I would seek for understanding from someone.	

3.4 Analysing the data

The collected data is analysed using R (version 4.0.5). Participants who did not answer any questions regarding the (in)dependent variable are excluded from further analysis (137 respondents). The following primary analyses are performed and are further explained in the results chapter: descriptive analyses (e.g. means), factor analysis, independent sample t-test, multiple regression and moderation. Next, the consideration is to omit participants who fill out the survey in a very short amount of time and/or who do not correctly fill out the factual manipulation check. There is much discussion about omitting participants from further analysis and whether this is the right or wrong thing to do (Greszki et al., 2015).[13] Therefore, this research performs three types of analyses: (1) with all respondents, (2) without the respondents who do not correctly answer the factual manipulation check and (3) without respondents who deviate more than two standard deviations from the average response time. The last two analyses serve as a robustness check.

3.5 Reliability and validity of the research

There are several categories of reliability and validity. This part highlights the most important categories where this research performs well or less.

3.5.1 Reliability

Reliability is the uniformity of the constructs one uses to measure certain variables. For this research, the internal reliability is primarily relevant, because the constructs have multiple items. These items should be to a certain extent related to each other to ensure that they measure all the same concept. Recall that this research uses the Weighted Omega to check whether the internal reliability is sufficient (minimum level .70) (Deng & Chan, 2017; Tummers & Knies, 2016, p. 18). The results described earlier show that all the constructs score high on the Weighted Omega and therefore strengthen the internal reliability of this research. An alternative that this research can use is Cronbach's alpha, but is frequently criticised regarding, for example, the appropriate alpha level. Bryman (2016, p. 158) states that an alpha between .70-.80 is good and the higher, the better. However, Schmitt (1996, p. 351) states that using

[13] E.g. https://www.researchgate.net/post/Excluding-cases-based-on-too-short-or-too-long-response-time

this bottom-line is myopic, because the alpha level is dependent on different factors (such as the length of the scale).

3.5.2 Validity

In terms of validity, we can identify internal, construct and external validity. Regarding internal validity, the question whether causality can be determined is relevant. An experiment often performs well on this type of validity. This is because experiments test whether an expected cause indeed has consequences for a different variable, i.e. if X causes Y (time precedence) (Kenny, 2004, p. 3). Besides, because participants are randomly allocated either in the control or the intervention group, it is to a great extent possible to rule out spuriousness[14] (Bryman, 2016, p. 46; Kenny, 2004, pp. 4-5). These two aspects strengthen the internal validity of this research.

A possible threat to this internal validity is the attrition effect, which describes the specific drop out of people. This is not a problem when the drop out is random; for example, when people do not have the time to fulfil the survey or become distracted. However, it is a problem when this drop out is biased. If mainly the manipulated public servants feel, for example, disrespected and stop with the survey, this is troublesome (Nunan et al., 2018). To prevent an attrition effect from happening, the survey emphasises that the news article is fictitious, but this does not rule out the risk. The plan was to perform a balance check of the groups (stay versus drop out as dependent variable) to test if there are significant differences between the groups (Miller & Hollist, 2007, p. 58). However, this is not done because of temporary inaccessibility reasons (which are elaborated later), and the balance check may give a skewed image. Also, some authors also over-sample (about 10-25%) to prevent or decrease the chance of an attrition bias, which this research also does (10% and more) (Hindmarch et al., 2015).

Regarding construct validity, the question is whether the used concepts or measures have sufficient relevance regarding the used theories to describe this concept. This validity is, in my opinion, sufficient in this research. When choosing the concepts' measures, the used definition from the literature is taken as a starting point. Regarding the scales' measured validity, only the self-reported performance scale is validated and calibrated. The other researchers (bureaucrat bashing and distancing/support) do not say something about validation. Recall that, bureaucrat bashing, distancing and support are tested on their factor loadings (Julnes, 2008, p. 528). If the factor loadings of the individual items are high, it shows that these items have a common underlying factor and correspond with the concept they measure. How high this loading should be is frequently discussed, but Peterson (2000, p. 264) concludes that the average factor loading in the studied literature .32 is, which this research uses as a minimum. Also, the

[14] Spuriousness means that within a statistical analysis X and Y are connected, but in reality they are not. This is because there is a third variable, a spurious or confounding variable, that creates changes in X and in Y (Bryman, 2016, p. 344; Kenny, 2004, pp. 4-5).

factor analysis is relevant as I translated the scales, because the research population are Dutch public servants. To check if the translated scales are still reliable and valid, they are also tested on their weighted Omega. See for the full results appendix 3.

External validity focuses on the extent of generalisation towards the population, which goes further than the specific research context (Bryman, 2016, p. 691). According to Bouwman & Grimmelikhuijsen (2015, p. 114), experiments are often criticised for their extent of generalizability. This is because the experiment creates an artificial setting, which limits the generalizability. However, according to Bouwman & Grimmelikhuijsen (2015, pp. 114-115), external validity can also be seen as *theoretical* generalization. This research tests the effects of bureaucrat bashing on performance based on theoretical insights, and with the knowledge of my research we can say something about the theory beyond the sample of this research. Besides, Bouwman & Grimmelikhuijsen (2015, p, 114) also see *mundane realism* as important when considering external validity. This means that the experiment resembles a situation that the respondent can also come across in reality. As shown in the introduction, bureaucrat bashing is often identified within, for example, the mass media or the political discourse. Using a news article simulates a real situation; in this way, the public servant can also encounter bureaucrat bashing in daily life. It is also checked with the participants in the pilot study if the article reflects a real situation, which it did.

3.6 Other considerations: ethics and preregistration

A few other considerations are valuable to mention within this methodology chapter. First, some ethical aspects of the research. Doing harm against participants is unacceptable. In this research, there are no consequences that can be harmful. By emphasising the fake character of the news article and debriefing, the harmful consequences of this experiment are very low or do not exist. The survey debriefs the participant in two ways at the end. By dehoaxing, the specific goal of the research is revealed and it is described that there are two different groups, who both receive a different treatment. Desensitizing is also crucial in the context of this research, because of the individual 'attack' the news article gives on the public servant. Therefore, I want to minimalise the negative consequences and allow participants to contact me if they feel disadvantaged by the experiment (Sommers & Miller, 2013).

Furthermore, regarding informed consent, respondents get as much information as possible about the research goal and how the research uses data. Before starting the survey, I ask the respondent if he or she agrees with the voluntary participation in this research and can choose to agree (survey begins and respondent can always stop) or not to agree (survey ends). Besides, questions regarding personal information are not compulsory. Regarding privacy, this research follows the General Data Protection Regulation (GDPR). Participants within the survey are entirely anonymous, and the data cannot be traced back to an individual or organization. All the data is treated confidentially and saved in Yoda.

Yoda is a safe storage for research data offered by Utrecht University. This data is only available for the researcher and the supervisor (Barbara Vis).

As described, the informed consent informs the participants about the scope and goal of the research. However, Bryman (2016, p. 48) states that this information can create reactive effects. This means that someone is aware of the intervention, influencing how the participants fill out the survey. If this is the case, this negatively affects the generalizability of the sample. This is a limitation of this research. It is difficult to fully tackle this, because the research attaches great importance to informing the respondents about their participation in the research. Therefore, the research tries to find a middle ground: inform the participants as much as necessary to make an informed decision about their participation, but also prevent a reactive effect. Besides, the survey emphasises to read the information carefully and to answer the questions as if the respondent is an employee of the organization (Kevins & Vis, 2021).

In addition, the research is preregistered (https://osf.io/nbm2d), and the complete thesis and R-script are shared on OSF (https://osf.io/w4e7h/). The method chapter shows how the research is broadly conducted, but this does not rule out that there are not any deviations made from the plan that was set up at the beginning of this process. How the research precisely is done or what deviations are made is often something that is not reflected upon at the end of the research, which limits the reproducibility and credibility of the results. Therefore, Nosek et al. (2019, p. 816) strongly advise researchers to preregister their research. The preregistration 'forces' the researcher to precisely formulate his or her plans regarding the research and stick to them. This makes it possible for other researchers to replicate the study and reflect on the research's strengths and weaknesses, which can serve as inspiration for future research. Important to note is that variation (when necessary) from this plan is allowed, but we should describe these variations. Table 3 shows all the deviations from the preregistration in OSF.

Table 3
Deviation from the preregistration on OSF

Deviations from preregistration	
1.	In the first place it was described that the survey would be spread only via governmental organizations and the alumni panel. However, as this did not result in enough respondents, the survey was also spread via LinkedIn (public sector) and relevant individuals (with a big network). Additionally, public servant organizations (e.g. FUTUR Jonge Ambtenaren Netwerk) were approached with the question whether they want to spread the survey.
2.	In the preregistration it was stated that a moderation would be performed with the help of the Baron & Kenny test; however, as there was no main effect, PROCESS of Andrew Hayes was done.

3.	The vignettes were slightly adapted after the study is preregistrered. Specifically, the main text (without the bashing) is shorter and I added some extra words to strengthen the bureaucrat bashing (e.g. incompetency, questionable functioning).

4. Results

This chapter describes the results of the performed analyses. Before the results of every analysis are presented, the chapters provides a short description and the used settings of the analyses. For more information about the analyses, used settings or assumptions, I would like to refer to:

- appendix 3 for information and results on the factor analysis;
- appendix 4 for information and results on the independent sample t-test;
- appendix 5 for information on the multiple regression;
- appendix 6 for information on the moderation;
- appendix 7 for the results of the performed analyses without the participants who took it more or less than 2 standard deviations to finish the survey;
- appendix 8 for the results of the performed analyses without the participants who answered the manipulation check wrong.

The following packages are used in R: tidyverse (Wickham, 2019), grid (R Core Team, 2021), REdaS (Maier, 2015), psych (Revelle, 2022), ggplot2 (Wickham, 2016), car (Fox & Weisberg, 2019), dplyr (Wickham et al., 2022), vtable (Huntington-Klein, 2021), readxl (Wickham & Bryan, 2022), and GPArotation (Bernaards & Jennrich, 2005). Process for R is separately downloaded.

4.1 Descriptive statistics

The total sample size is 321 respondents, however 137 respondents quit the survey after the descriptive variables.[15] A possible reason for this is that participants expect that it takes much time to read the news article (or to fulfil the survey) and therefore quit the survey. Another reason centres around technical problems University Utrecht had with Qualtrics, whereby the vignettes were not visible for several days.[16] If people do not see the fictitious newspaper, the accompanying text and questions that followed would not make sense and come out of the blue. This may result in the decision to stop with the survey. These technical problems occurred in the beginning of the distribution of the survey. The question is whether this disadvantaged specific public servants; as far as I can consider this is not the case, because I spread the survey among several governmental organizations already from the beginning of the distribution. Next, two respondents are also excluded from further analyses because they did either not choose to participate in the research or have not filled out the questions regarding the (in)dependent variable. Four participants did not fill out the questions regarding the coping mechanisms. These participants are included for the multiple regression and not in the moderation analysis. Considering this, the total sample size is 182 (95 control; 87 intervention), which meets the minimal sample size and includes the oversampling of 10%.

[15] It is, in this case, difficult to determine the response rate because of how participants were recruited, e.g. sharing via LinkedIn.
[16] Besides the vignettes, the whole survey was also unavailable for a few days.

4.1.2 Results of descriptives

Looking at the descriptives, a few things stand out. First, the distribution of the variable 'gender' is almost evenly: 89 men and 90 women (3 people preferred not to say their gender). Also, within the two groups, the performed balance check shows that men and women are evenly distributed (about 46% in control and 51% in intervention). Within both groups, most of the respondents work in the ministry or in a municipality, as shown in figure 3. The number of working years differ to a small extent between the control and intervention group and the balance check shows that this variable is evenly distributed among the two groups. In total, 66 participants have between one and five years of experience, and 68 participants have more than 16 years of work experience.

Figure 3

Histogram of working organizations of participants

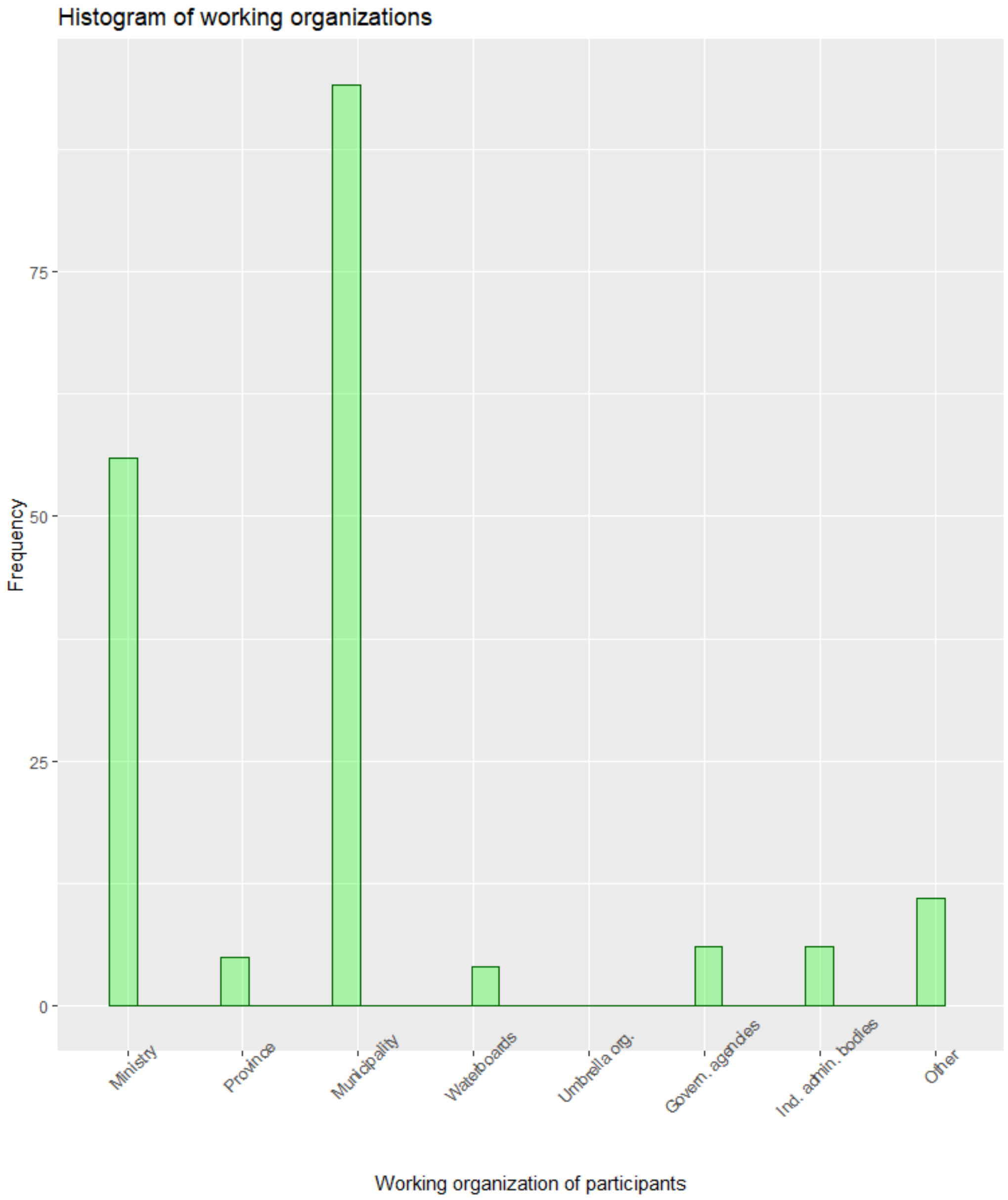


Table 4

Descriptive statistics of control group regarding the independent, dependent and control variables

			Potential range		Real range	
Variable	n	M (SD)	Min	Max	Min	Max
M. bureaucrat bashing	95	2.16 (.62)	1	5	1	5
Performance	95	3.91 (.42)	1	5	1	5
Coping - distancing	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]
Coping - social support	-[a]	-[a]	-[a]	-[a]	-[a]	-[a]

Note. n = sample size. M = mean. SD = standard deviation. Min = minimum. Max = maximum.

[a] The questions regarding coping were not asked to the control group, because there was no bashing present in the vignette.

Table 5

Descriptive statistics of intervention group regarding the independent, dependent and control variables

			Potential range		Real range	
Variable	n	M (SD)	Min	Max	Min	Max
M. bureaucrat bashing	87	3.36 (.81)	1	5	1	5
Performance	87	3.94 (.44)	1	5	2	5
Coping - distancing	87	1.84 (.64)	1	4	0[a]	4
Coping - social support	87	2.26 (.72)	1	4	0[a]	4

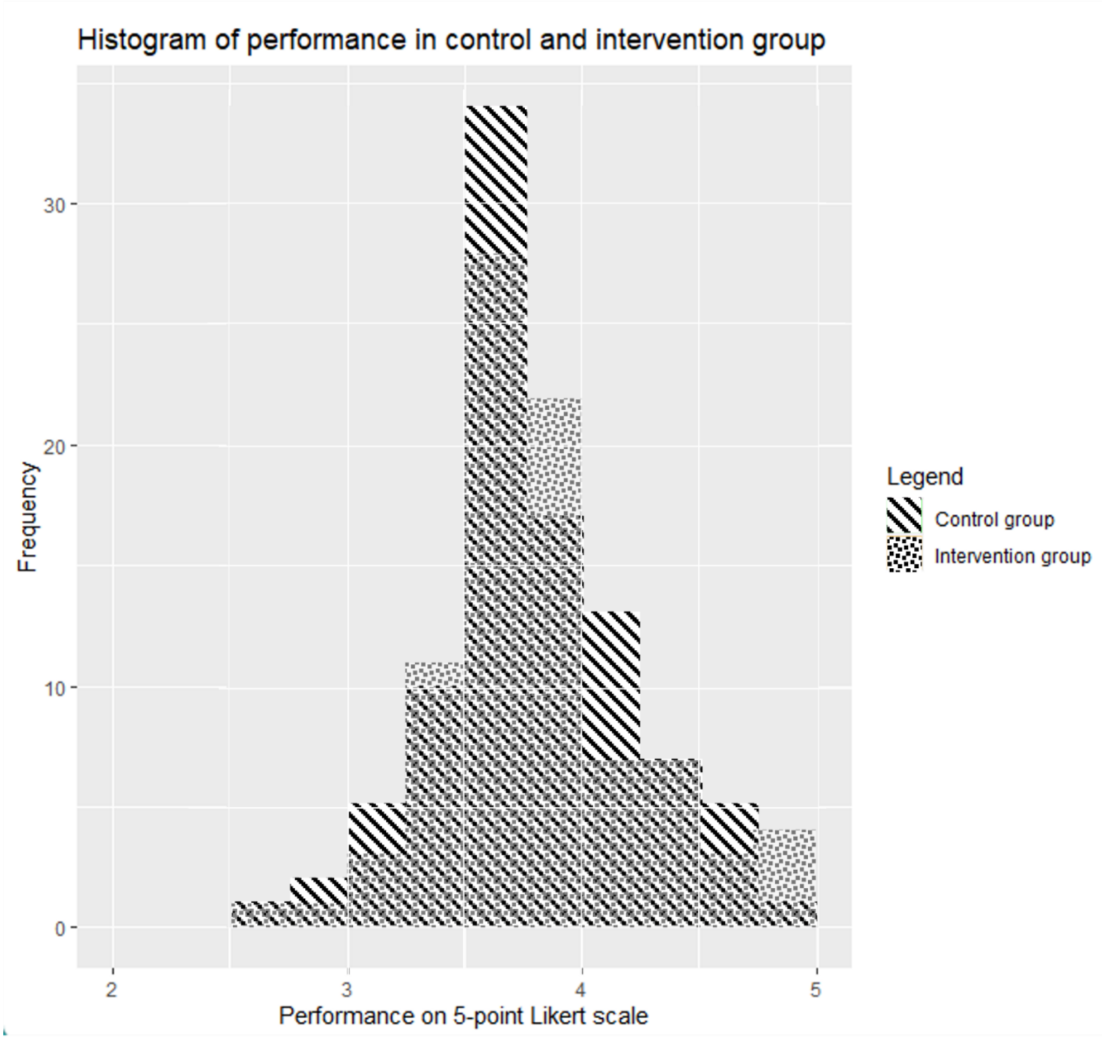
Note. n = sample size. M = mean. SD = standard deviation. Min = minimum. Max = maximum.
[a] Some respondents did not fill in the questions regarding the coping mechanisms, resulting in NA. All the NA's were transformed into 0.

Looking at table 4 and 5, the results show that the mean of bureaucrat bashing in the intervention group (M = 3.36) is higher than in the control group (M = 2.16), showing that meaningless bashing is more experienced in the intervention group and we can carefully conclude that the manipulation worked. Noticeably, both means of performance are very close to each other (control: M = 3.91; intervention: M = 3.94): this can indicate that the bashing intervention has no or a minimal effect on the self-reported performance. This possible working is also visible in the means of distancing and social support, which are low: when bureaucrat bashing does not affect performance, the public servant possibly does not feel the need to use coping mechanisms.

Additionally, the standard deviation of performance in both groups shows a low amount of variation in the given answers, see figure 4 and table 4/5. The figure shows the control group in the diagonal striped bars and the intervention group in the dotted bars. The highest frequency is within both groups situated between the 3 (Neutral) and 4 (Agree). Figure 4 indeed shows that the standard deviation is low: in both groups, the scores of performance centre around the middle, demonstrating the low amount of differences between the reported scores. Also, the histograms of both groups are placed on top of each other. This shows that not only the differences in the scores *in* the group are low, but also *between* the groups, showing the low variation between the control and intervention groups.

Figure 4

Histogram of performance in control and intervention group



4.2 Independent sample t-test

An independent sample t-test tests whether the control and intervention group differ significantly (Field, 2009, pp. 334-342). The data meet the necessary assumptions, making it possible to perform this test. The results show no significant difference between the mean of the control (M = 3.905) and intervention group (M = 3.94) regarding performance, which already can show us that intervention does not affect performance significantly (t (177, 21) = -.497, p = .62). The intervention itself, however, shows a significant difference between the control (M = 2.156) and intervention (M = 3.356) group, showing that bureaucrat bashing is experienced in the intervention group (t (177, 21) = 126.451, p > .001). More information about this analysis is in appendix 4.

4.3 Multiple regression

One uses a multiple regression analysis when the goal is either to predict how the dependent variable changes in reaction to an independent variable or to test if there exists a causal relationship between the independent and the dependent variable, as is also the case within this research (Allison, 1999, pp. 1-2). More information about this analysis is in appendix 5. Next to the 'normal' multiple regression, two robustness analyses (multiple regression) are performed to test whether the results remain stable when the input of the analysis is varied. In this case, the following participants are removed: those who gave a wrong answer on the factual manipulation check and those who took it longer or shorter than two standard deviations to end the survey. See for the results appendix 7 and 8; these results do not show considerable deviations from the results presented here.

4.3.1 Results of the multiple regression

This section tests hypothesis 1, which is: *public servants, who are exposed to meaningless bureaucrat bashing, will display lowered performance than those who are not exposed to meaningless bureaucrat bashing.* Both models meet the necessary assumptions to do a multiple regression, except for the linearity assumption (see table 6). There is no massive deviation from a perfect linear model, but there is some. To deal with this, one either performs a logistic transformation of the dependent variable or continues with multiple regression and accepts the data's limitation. The logistic transformation causes only a tiny change in the model. Because a logistic transformation is more difficult to interpret for the reader and because of the small change it creates, the research performs the original multiple regression. This means, however, that the results of my models are not fully generalizable to the intended population.

Table 6
Overview of (not-)fulfilled assumptions

Model	Assumption
Control group	Assumption of normal distribution
	Assumption of independence
	Assumption of no multicollinearity
	Assumption of linearity
Intervention group	Assumption of normal distribution
	Assumption of independence
	Assumption of no multicollinearity
	Assumption of linearity

The results of the control and intervention group are in table 7. Model 3 shows that the hierarchical addition of meaningless bureaucrat bashing within the control group creates a significant positive effect of bureaucrat bashing on the self-reported performance of public servants: if bureaucrat bashing goes up by 1, the self-reported performance goes up by .19 (β = .19, SE = .07, p < .10). As model 3 is non-significant (F (4, 94) = .57, p > .05), it shows that bureaucrat bashing cannot significantly explain the change in the self-reported performance.

Model 2 and 4 show the results within the intervention group: both models remain significant (in contrast to the models of the control group), where the explained variance slightly increases through the addition of bureaucrat bashing (from approximately 9% to 11%) (F (4, 86) = 2.56, p = .04). However, the variable bureaucrat bashing has a non-significant slightly positive effect (the slope of the regression line is almost close to zero). This observation shows us that bashing has some effect, as the slope is much lower in the intervention group than in the control group. Thus, the absence of bureaucrat bashing (control group) has a positive consequence on the servant's performance, whereas the presence of bureaucrat bashing (intervention group) creates a much lower positive, or almost no, change in the performance. Hereupon, I reflect in the discussion. To conclude, no significant evidence shows that those exposed to meaningless bureaucrat bashing reported lowered performance, meaning that hypothesis 1 is **rejected.**

Table 7

Predictors of self-reported performance in the control and intervention group

Variable	Model 1 (C) B (SE)		Model 2 (I) B (SE)		Model 3 (C) B (SE)		Model 4 (I) B (SE)	
Constant	3.93	***	4.09	***	3.59	***	3.81	***
	(.11)		(.10)		(.16)		(.22)	
Control variables								
Gender	-.11		-.23	*	-.12		-.21	*
	(.09)		(.09)		(.08)		(.09)	
Working organization	.02		.08		-.04		.08	
	(.09)		(.10)		(.09)		(.10)	
Years of work experience	.03		-.15		-.01		-.14	
	(.09)		(.10)		(.09)		(.10)	
Bureaucrat bashing					.19	**	.08	
					(.07)		(.06)	
R^2	.02		.09		.09		.11	
F	.59		2.74	*	2.24		2.56	*

Note. C = control group ; I = intervention group.

* $p < .05$. ** $p < .01$. *** $p < .001$

4.4 Moderation

A moderator changes the strength or the direction of the relationship between the independent and dependent variables (Baron & Kenny, 1986, p. 1174). One way to test this is by using a Baron and Kenny test, where one has to create an interaction variable of the independent variable combined with the moderating variable. However, as the second model of the intervention group does not have a significant main effect between the independent and dependent variables, it is impossible to do a Baron and Kenny test (necessary assumption). However, it is possible to perform a PROCESS-test of Andrew Hayes, where this main effect is not necessary (Hayes & Rockwood, 2016). This analysis uses bootstrapping to calculate the moderation results. Bootstrapping handles the sample data as population data. From this population data, the program draws several bootstrap samples and calculates within each sample the necessary analysis (in this case, a moderation analysis): it compiles all these results to constitute the complete dispersal of the statistic in the sample. In this way, we get further insights into the coefficients of the moderators' slope, the value of the *t*-test and the *p*-value (Field, 2009, p. 163). More information about this analysis is in appendix 6.

Furthermore, the research performs two robustness analyses (moderation) to test whether the results remain stable when the input of the analysis is varied. In this case, the following participants are removed in these robustness analyses: those who gave a wrong answer on the factual manipulation check and those who took it longer or shorter than two standard deviations to end the survey. See for the results appendix 7 and 8; these results do not show considerable deviations from the results presented here.

4.4.1 Results of moderation analysis

In this section hypothesis 2a is tested: *the influence of meaningless bureaucrat bashing on performance is negatively moderated by the coping mechanism distancing.* The results of the moderation analysis of the coping mechanism distancing are not significant, meaning that the third variable distancing does not affect the relationship between bureaucrat bashing and performance ($\beta = .07$, $t(6, 86) = .67$, $p > .05$). Regarding hypothesis 2b (*the influence of meaningless bureaucrat bashing on performance is negatively moderated by the coping mechanism social support),* the same working of the moderator is identified: also here, the moderator social support does not have a significant effect on the relationship between bureaucrat bashing and performance for the public servant ($\beta = -.10$, $t(6, 86) = -1.06$, $p > .05$). However, the model remains significant ($F\ (6, 86) = 2.27$, $p = .04$) and the explained variance increases even to 14.57%. Although the moderator social support is not significant, the negative beta indicates that if public servants cope with the help of social support, the relationship between bureaucrat bashing and performance becomes less strong (but we cannot draw hard conclusions because of the non-significance). These results show that hypotheses 2a and 2b are **rejected**.

4.5 Summary acceptation and rejection of hypotheses

Table 8 shows the accepted and rejected hypotheses.

Table 8
Summary acceptation and rejection of hypotheses

Hypothesis		✓ / ✗
1	Public servants, who are exposed to meaningless bureaucrat bashing, will display lowered performance than those who were not exposed to meaningless bureaucrat bashing.	✗
2a	The influence of meaningless bureaucrat bashing on performance is negatively moderated by the coping mechanism distancing.	✗
2b	The influence of meaningless bureaucrat bashing on performance is negatively moderated by the coping mechanism social support	✗

5. Discussion and conclusion

This chapter answers and discusses deeper the research and sub questions, which are subsequently placed into a broader theoretical and practical context. Also, the chapter discusses the interesting points for future research and limitations of this research. To recap, this research studies the following research question:

> *To which extent does bureaucrat bashing influence the performance of public servants, whose task is focused on policy formulation, policy preparation or policy monitoring and to which extent is this relationship moderated by coping mechanisms?*

5.1 Broader understanding of the results

Each sub question mentions the relevant results again and relates them to discussions in the literature.

5.1.1 Sub question 1: what is the influence of bureaucrat bashing on the performance of public servants?

Reasoned from several theoretical and empirical insights, I argued that experienced meaningless bureaucrat bashing creates a stereotype threat within the individual public servant, disturbing the cognitive and environmental assets (that normally contribute to the individual performance). Therefore it was expected that the experienced bureaucrat bashing lowers the self-reported performance (Cohen et al., 2011, p. 286; Inzlicht & Schmader, 2011). However, as was already derived from the amount of variation between the groups regarding the variable performance, the results show no significant influence of experienced meaningless bureaucrat bashing on the self-reported performance of public servants. This means there is no causal relationship between these two variables, in contrast to the theoretical expectations.

However, interestingly, the descriptives of the variables demonstrate that the mean of the intervention group regarding experienced bureaucrat bashing was higher than in the control group, which can imply that the manipulation has its effect on the participants. Indeed the analyses show some kind of effect created by the manipulation, as the slope of the regression line is much lower in the intervention group than in the control group. More specifically, the absence of bureaucrat bashing has a positive consequence on the servant's performance, whereas the presence of bureaucrat bashing creates almost a neutral change in the performance. How can we explain this different working of the variables bureaucrat bashing and self-reported performance?

First, it is possible that the intervention was not strong enough to create the theoretical expected disturbing effect of meaningless bureaucrat bashing. The used survey experiment provides a methodological contribution to the existing bashing literature, but, as Grimmelikhuijsen & Bouwman (2015, p. 113) note, this kind of hybrid experiment offers a low amount of control: we do not have any control on how carefully the participants read the vignette. The factual manipulation check can tell us more about how attentive the participants were regarding the vignette: in the control group three participants (3.2%) answered the manipulation check wrong; in the intervention group seven participants (8.0%) answered it wrong. This group is only a small part of the whole sample (5.5% of the total sample), showing that we could expect that a big part of the sample carefully read the vignette. Still, there is a chance that participants guessed the correct answer, while not reading the vignette. This is difficult to rule out. Considering the low rates of wrong-answered manipulation questions and the low differences in the robustness check, it is plausible that the non-significant relationship between bureaucrat bashing and performance is not due to the low extent of attentiveness (but this cannot be ruled out entirely).

Besides, this result shows a theoretical contribution: the role bureaucrat bashing has on public servants is probably more complicated than is assumed in the theoretical framework. This observation matches earlier findings, which show that the effects of bashing are mixed. Whereas Caillier (2018) and Garrett et al. (2006) demonstrate the negative direct effect of bashing on public servants and the indirect negative effect of bashing on citizens and resulting attitudes towards public servants, Szydlowski et al. (2022) and Caillier (2020) could not replicate this. Both show that the bashing of public servants does not influence citizens' behavior or views (e.g. compassionate behavior) towards public servants. What can we learn from these mixed insights? A possible lesson from this is that studying bashing is not something dichotomous (e.g. either the presence or absence of bureaucrat bashing), but more a matter of a continuum (e.g. a low, moderate or high extent of bureaucrat bashing). Mendes & Jamieson (2011) connect this to the already existing job demands and resources, when one experiences a stereotype threat. If a high extent of bureaucrat bashing creates a significant aggravation of the demands, whereby they exceed the resources, the expected cognitive and psychological displacement of bashing can materialise on the individual (see figure 5). Nevertheless, when the resources are high, a high extent of bashing can intensify the demands but not outweigh the resources (see figure 6). This can also be the case here, as there is a difference in the experienced bureaucrat bashing between the groups. When the job resources strongly outweigh the job demands in the participant's daily work, the bashing may not yet have the negative displacing effects on performance. This can also be an important explanation for the mixed results of the studies mentioned above. Garrett et al. (2006) describe, for example, that the bashing in political campaigns creates an unpleasant and hostile work environment, which may lead to the disbalance between the demands and resources in favour of the demands. Therefore, for future research, the experienced job demands and resources are thus an important control variables.

Figure 5[17]

Illustration - demands exceed resources

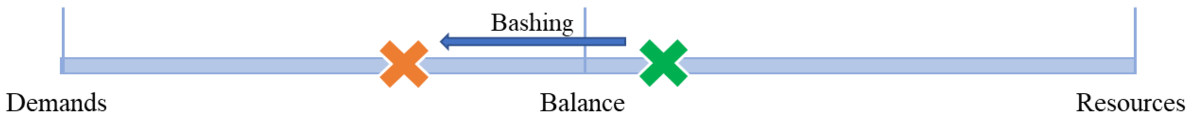


Figure 6

Illustration – demands do not exceed resources

5.1.2 Sub question 2: what is the role of the coping mechanisms distancing and social support in the relationship between bureaucrat bashing and job performance?

Based on the literature insights, it was predicted that public servants either would dissociate (i.e. distance) or search for social support to deal with the negative effects of bureaucrat bashing (Derks et al., 2015; Veldman et al., 2020; Rosen et al., 2022; Cole et al., 2007). Both would thus be essential resources in the daily job of the public servant. The descriptive results show already a difference in the usage of the coping mechanisms: when public servants experience the need to cope, they are less likely to dissociate from the experienced negative demands and more likely to, for example, talk with colleagues about how they feel. However, in this research this is difficult to statistically confirm, as there is no main effect between the variables meaningless bureaucrat bashing and performance (i.e. there is nothing to cope with). However, this research has done PROCESS analyses to gain more insight into the working of the variables social support and distancing. Both analyses show no significant effect of distancing and social support in the relationship between experienced meaningless bureaucrat bashing and job performance. Only social support has a negative coefficient-slope of the regression line, which weakens the relationship between experienced bureaucrat bashing and performance. However, we need to be careful with these observations as these are not significant. To conclude, there is no significant role of the coping mechanisms distancing and social support in the relationship between bureaucrat bashing and job performance.

[17] Figure 5 and 6 show the described balance between the demands and the resources. If the public servant experience more resources than demands, he/she find themselves on the right side of the balance and vice versa. When the public servant experiences the proportion of demands and resources as in figure 5, a high extent of experienced bureaucrat bashing aggravates the demands and exceed the resources. In figure 6, the public servant experiences already higher resources; even when the public servant experiences a high extent of bureaucrat bashing, the resources still prevail over the demands, which possibly make that the servant experiences less negative effects of bureaucrat bashing on, for example, his/her performance.

Again, how can we explain these results? The concept of Public Service Motivation can play an important role in this case. For example, Bakker (2015, p. 730) concludes that because of the endurance, productivity and work engagement caused by this type of motivation to contribute to society, the public servant deals for a great extent with the stressors or demands in daily life. This can explain the low impact of the coping mechanisms such as distancing, as Public Service Motivation may function as an important coping mechanism. Furthermore, the observation that social support has some weakening effect on the relationship between bureaucrat bashing and performance matches with the empirical insights retrieved from the street-level bureaucracy: when dealing with high job demands, social support from especially colleagues, helps alleviate the negative effects of these demands. Yet, it is important to remember that these are only expectations and not observations because of the non-significance of the results, so we need to be careful.

5.2 Future research

The above stated points give interesting points to consider for future research. First, future research can study bureaucrat bashing in a context where there is more control of the manipulation, for example, in a lab experiment. One can study this within a lab environment, where a control group sees, for example, a 'neutral' video, intervention group 1 sees a video with a moderate amount of bashing and intervention group 2 sees a video with a high amount of bashing originating from citizens. Then participants will be asked to perform a simple action to see if the bashing affects on the public servant's performance. This approach makes it possible to monitor the effects of the bureaucrat bashing more closely and in a more controlled manner in contrast to the design used in this research. Second, it is interesting to take a closer look at the working of Public Service Motivation, a differentiating motivation of public sector workers to serve the public interest (Perry, 1997, p. 181). Research from Perry & Wise (1990) demonstrates, for example, that employees are motivated to serve the society, where he/she is prepared to do the most for the citizen, even if this could jeopardize their personal interests. How does this role develops when one gets bashed by a citizen? Does Public Service Motivation have the same working as a coping mechanism, as suggested by Bakker (2015)? It is also interesting to study this from a qualitative aspect to better understand the working of bureaucrat bashing and Public Service Motivation from the perspective of the public servant. For example, by using focus groups we gain information about the role of these two variables, specifically it gives us knowledge about which shared experiences, arguments and observations the public servants may have. This gives the researcher knowledge about the embedding of the topic in the social reality and creates shared understanding of these topics (Boeije & Bleijenbergh, 2019, p 78). Furthermore, this research only focuses on meaningless bureaucrat bashing, however, it is also relevant to look deeper into the other side of bureaucrat bashing, namely substantive bashing: how do public servants react to this type of bashing and does this type also have emotional impacts? Another possibility is to investigate more closely the effects of prescriptive and proscriptive elements of stereotypes (and bashing) instead of mainly descriptive elements. Finally, to my best knowledge, this is the first research on bureaucrat bashing with public servants, working on policy preparation, formulation or monitoring in the Netherlands. Future research is needed to research the working of bureaucrat bashing within this group and other groups of bureaucrats in different contexts.

5.3 Limitations of research

First, a limitation of this research is the generalizability of the findings. Some aspects of the design of this research contribute to this limited generalizability. Recall that the control and intervention group models do not fully fulfil the linearity assumption, meaning that it is more difficult to generalise the findings to the population. In addition, the literature often criticises experimental designs for their limited external validity, which is a limitation of this research (Bouwman & Grimmelikhuijsen, 2016, p. 114). However, by creating a hypothetical news article, which resembles a situation the respondent can also come across in reality, and checking this in a small pilot study, this research tries to contribute to a different dimension of external validity, namely mundane realism (Bouwman & Grimmelikhuijsen, 2016, p. 114). Still, participants knew that the news article was hypothetical, which may limit the extent of mundane realism and the extent to which respondents connect the hypothetical bashing with their actual performance. Future research benefits from using an article that was already published.

Second, the social desirability bias may play an important role in explaining the results. This bias creates the tendency to answer the questions in a way that is socially desirable in the eyes of the respondent. As Krumpal (2011, p. 2026) states, this is often the case with 'sensitive questions'. Both the items of bureaucrat bashing and performance require a review of the public servant's feelings and functioning, which may lead to socially desirable answers (e.g. high performance). By emphasising the anonymity and confidentiality of the respondent's identity and data, by describing that there are no right or wrong answers, and by indicating that the data is only available to myself and my supervisor, I have tried to decrease this social desirability bias to a minimum. As seen from the results, the social desirability bias seems not to interfere in the questions about bureaucrat bashing, as there were clearly differences between the control and intervention group regarding this specific variable.

However, the social desirability bias can still be present in the performance questions (despite the efforts to minimalise it), because the means in the control and intervention group of the performance questions are very close to each other. The performance-related questions may create a *self-representation concern*, and thereby participants did not fully report their real feelings, creating an inaccurate representation of the reality (Krumpal, 2011, pp. 2025-2026). Besides, there is also a possibility of self-selection, where only high-performing public servants choose to participate in the study, thereby making the results biased. Yet, there are low differences, when comparing the mean of performance in this research (M = 3.91) with the means of performance of other researchers. For example, Vandenabeele

(2009, p. 21) and Johari & Yahya (2018, p. 6) study the self-reported performance of public servants; both report a high mean of performance (respectively, M = 3.88 - Likert scale 1-4 and M = 5.523 – Likert scale 1-7). Tummers (2016) has questioned supervisors of street-level bureaucrats regarding the bureaucrats' performance. Performance, measured on two different scales, was also in this research high (M = 3.995 – Likert scale 1-5; M = 4.210 – Likert scale 1-5). The research was conducted in two phases, where bureaucrats first choose to fill out a survey and then the supervisor is asked for an assessment of their participant's performance; this minimalise the chance of self-selection. Although these studies' populations do not fully match the studied population in this research, we do see a low amount of variation in the reported performance. Showing that this self-selection does not play a crucial role in explaining the results. Furthermore, using a randomised experiment decreases the chance of endogeneity and the risks of third, confounding variables. Thus, supposing that social desirability is normally distributed among the sample, the randomised character of the experiment does lessen the worries surrounding the possible confounding effect of the social desirability bias.

Third, a limitation of the research is regarding the possible results. The results tell us something about the causal relationship between meaningless bashing and performance, with no moderating role of coping. The research gets little or no insights into, for example, how public servants interpret the bashing, how often they have experienced this phenomenon and what this means for the public servant (Boeije & Bleijenbergh, 2019, p. 25). At the same time, it is valuable to get more understanding in these more interpretative aspects of bureaucrat bashing. This research does not tackle this limitation; however, it is an important recommendation for future research.

Fourth and lastly, this research uses an experimental design and manipulates the independent variable meaningless bashing. However, as Kane & Barbaras (2019, p. 234) describe, how do we know if the intervention really manipulates the intervention group? Suppose the effect between meaningless bashing and performance does exist, but because of possible less attentiveness with respondents the effect is not measured correctly. This is a limitation of doing survey experiments in contrast with a lab experiment, where the researcher has more control over the experimental setting (Blom-Hansen et al., 2015, pp. 158-161). I have tried to minimalise this limitation by using a manipulation check; recall that only a small part (5.5%) of the sample size answered the factual manipulation check incorrectly.

5.4 Practical implications

Despite that meaningless bureaucrat bashing has not created a significant change in the self-reported performance of public servants, bureaucrat bashing itself was experienced more by public servants in the intervention group. Therefore, it is crucial to give attention to the (possible) effects of bureaucrat bashing, which is the focus of this sub section.

Despite the positive achievements and the complex tasks within the government, the effectiveness and efficiency of the government are frequently put in a negative light, often called the anti-public sector bias (Marvel, 2015, pp. 210-211). Nowadays, this anti-public sector bias is often embedded in several populist contexts, making the topic more relevant as the intensity, frequency and possible effects of bureaucrat bashing may increase for public servants (Moynihan, 2020). Moynihan & Roberts (2021, pp. 152-153) observe, for example, that the bashing of bureaucrats was always present in the political context, but took a rise under the populist Trump administration, who portrayed the bureaucracy as a *swamp*.

Although this research did not empirically find a negative effect of bashing, this does not mean that the effects of bureaucrat bashing on the public servants do not exist; after all, we do see variation in the experienced bureaucrat bashing. Besides, research from Marvel (2015) and Garrett et al. (2006) show that bureaucrat bashing can affect public servants' motivation and attractiveness of the public sector, which is not researched here. Also, Spencer et al. (2016, pp. 419-425) expect that the stereotype threat the bureaucrat bashing could create, affects the performance in different ways, such as memory performance or broader the individual performance on his/her well-being. While the effects on the work performance may be limited, they can be more prominent on other types of performance, making it vital not to lose sight of it. Governmental organizations should therefore start (or keep) giving attention to bureaucrat bashing within their organizations and make the concept discussable among employees. Based on a similar concept, Rosen et al. (2022, p. 85) give an useful example of this. They show that programs, called for example *Battle Buddies*, created by hospitals for their co-workers, where they can talk with each other about daily struggles, helped participants in stressful times (i.e. Covid-19) and created a trustful work environment. Although originating from a different context, this is a valuable point of consideration for governmental organizations. When governmental organizations want to give more attention to the possible impacts of bureaucrat bashing, I would advise setting up a comparable program for their public servants, called, for example, *Talking Buddies*. In this program, public servants are paired based on their daily activities and struggles (amongst other things bureaucrat bashing), where they talk about their experiences. This is all focused on creating and maintaining a healthy and happy workplace.

Besides, it is relevant to openly discuss (the effects of) bureaucrat bashing within society. Negatively displaying bureaucracies, attacking public servants personally and underemphasizing the bureaucracy's positive achievements can also lead less trust of citizens in the bureaucracy and the bureaucrats (Szydlowski et al., 2022). This is mainly important in the face of the already decreasing trust in the Dutch government in general, which was about 30% in September 2021 (Engbersen et al., 2021). Therefore, it is essential to give more attention to the performance and well-founded criticisms of the government, but also show the impact of personally attacking public servants and/or unfounded criticisms. A good example hereof is the media attention towards the research of Bertram and colleagues on stereotypes regarding public servants (RTL Nieuws, 2022; EditieNL, 2022).

5.5 Answer to the research question

To be bashed or not to be bashed, that is the question; a derivative of the famous line of Shakespeare that he used to show the devilish dilemma of Hamlet. In this research, a dilemma is also visible: should the public servant passively accept and undergo the bureaucrat bashing as part of the job, or should we actively discuss bureaucrat bashing and look more closely at the possible effects on public servants? This research shows that meaningless bureaucrat bashing does not significantly influence on the self-reported performance of public servants, whose task focuses on policy formulation, preparation or monitoring. There is no evidence available to show the existence of a causal relationship. Also, the coping mechanisms social support and distancing do not moderate the non-significant relationship between bureaucrat bashing and self-reported performance. However, the results show that the manipulated public servants experience meaningless bureaucrat bashing and that the presence of bureaucrat bashing lowers the self-reported performance in comparison to the absence of bashing (albeit non-significant). Even when bureaucrat bashing does not significantly affect the servant's performance in this research, it does not exclude that there are no effects at all for the public servant. Think of the servant's well-being, motivation or job satisfaction: how are these aspects affected by meaningless bureaucrat bashing? Therefore, as other researchers advise (e.g. Betram et al., 2022) and to give an answer to the dilemma above, we must remain attentive, nuanced and open to discussion regarding meaningless bureaucrat bashing.

Reference list

Allison, P. (1999). What is multiple regression. In P.D. Allison (Ed.), *Multiple regression: A primer* (pp. 1-23). Pine Forge Press.

Aronson, J. (2002). Stereotype threat: Contending and coping with unnerving expectations. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 279–301). Academic Press.

Baker, J.P. & Berenbaum, H. (2007). Emotional approach and problem-focused coping: a comparison of potentially adaptive strategies. *Cognition and Emotion, 21*(1), 95-118.

Bakker, A.B. & Demerouti, E. (2006). The job demands-resources model: state of the art. *Journal of Managerial Pscyhology, 22*(3), 309-328.

Baldwin, J.N. (1990). Public versus private employees: debunking stereotypes. *Review of Public Personnel Administration, 11*(1-2), 1-27.

Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Baviskar, S. & Winter, S.C. (2017). Street-level bureaucrats as individual policymakers: the relationship between attitudes and coping behavior toward vulnerable children and youth. *International Public Management Journal, 20*(2), 316-353.

Bayram, P. & Zoubi, K. (2020). *Management Science Letters, 10*(8), 1771-1776.

BBC News. (2015). *Why are the Dutch so tall? BBC News* [YouTube video]. Retrieved at 29 April 2022, from https://www.youtube.com/watch?v=q_lDTOIJ4xA.

Bee, R. & Bee, F. (1998). *Constructive feedback*. Chartered Institute of Personnel and Development.

Belle, N., Cantarelli, P. & Belardinelli, P. (2018). Prospect theory goes public: experimental evidence on cognitive biases in public policy and management decisions. *Public Administration Review, 78*(6), 828-840.

Bernaards, C.A. & Jennrich, R.I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis [R Package]. *Educational and Psychological Measurement, 65*(5), 676-696.

Bertram, I. Bouwman, R. & Tummers, L. (2022). Socioeconomic status and public sector worker stereotypes: results from a representative survey. *Public Administration Review, 82*(2), 237-255.

Blom-Hansen, J., Morton, R. & Serritzlew, S. (2015). Experiments in public management research. *International Public Management Journal, 18*(2), 151-170.

Boeije, H. & Bleijenbergh, I. (2019). *Analyseren in kwalitatief onderzoek*. Boom Uitgevers Amsterdam.

Boer, N. de (2020). How do citizens assess street-level bureaucrats' warmth and competence? A typology and test. *Public Administration Review, 80*(4), 532-542.

Bordalo, P. Coffman, K., Gennaioli, N. & Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics, 131*(4), 1753-1794.

Borst, R. & Blom. R. (2021). Stressed out public servants? Testing the stress appraisal of psychological and emotional stressors on stress reactions through psychological capital. *International Journal of Public Sector Management, 35*(2), 113-132.

Borst, R., Lako, C. & Vries, M. de. (2014). Wie is die trotste ambtenaar? Een onderzoek naar de determinanten van beroepstrots. *Bestuurswetenschappen*, *68*(3), 24-47.

Bos, A.L. & Schneider, M.C. (2018). Stereotype measurement in political decision making. In *Oxford Research Encyclopedia of Politics* (pp. 1-17). Oxford University Press.

Bos, R. t. (2016, 3 February). Dit zijn de vijf typen bureaucraten die ons land besturen. *De Correspondent.* Retrieved at 14 March 2022, at https://decorrespondent.nl/3982/dit-zijn-de-vijf-typen-bureaucraten-die-ons-land-besturen/1003665274172-6d320a20.

Boselie, P., Harten, J. van & Veld, M. (2021). A human resource management review on public management and public administration research: stop right there… before we go any further… *Public Management Review, 23*(4), 483-500.

Bouwman, R. & Grimmelikhuijsen, S. (2015). Experimental public administration from 1992 to 2014. *International Journal of Public Sector Management*, *29*(2), 110-131.

Bovens, M.A.P., Hart, P. ‘t, Twist, M.J.W. van, Berg, G.F. van den, Steen, M.A. van der & Tummers, L.G. (2017). *Openbaar Bestuur.* Wolters Kluwer.

Brown, J.D. (2009). Choosing the right type of rotation in PCA and EFA [Newsletter]. *JALT Testing & Evaluation, 13*(3), 20-25.

Bryman, A. (2016). *Social Research Methods*. Oxford University Press.

Caillier, J. (2018). The priming effect of corruption and bureaucracy bashing on citizens’ perceptions of an agency’s performance. *Public Performance & Management Review, 41*(2), 201-223.

Caillier, J. (2020). Bureaucratic bashing and praising: what effect does it have on the performance citizens assign agencies? *Public Organization Review, 20*(4), 685-701.

Campo, C. del, Pauser, S., Steiner, E. & Vetschera, R. (2016). Decision making styles and the use of heuristics in decision making. *Journal of Business Economics, 86*(4), 389-412.

Carlos, V.S. & Rodrigues, R.G. (2015). Development and validation of a self-reported measure of job performance. *Social Indicators Research*, *126*(1), 279-307.

Chen, C.A. & Bozeman, B. (2014). Am I public servant or am I pathogen? Public managers’ sector comparison of worker abilities. *Public Administration*, *92*(3), 549-564.

Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin, 103*(2), 276-279.

Cohen, G.L., Purdie-Vaughns, V. & Garcia, J. (2011). An identity threat perspective on intervention. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: theory, process, and application*. Oxford University Press.

Cole, B., Matheson, K. & Anisman, H. (2007). The moderation role of ethnic identity and social support on relations between well-being and academic performance. *Journal of Applied Social Psychology, 37*(3), 592-615.

Collins Dictionary. (n.d.). *Definition of bureaucrat*. Collins. Retrieved at 27 March 2022, from https://www.collinsdictionary.com/dictionary/english/bureaucrat.

Columb, M.O. & Atkinson, M.S. (2016). Statistical analysis: sample size and power estimations. *BJA Education, 16*(5), 159-161.

Dawson, J. (2014). Moderation in management research: what, why, when and how. *Journal of Business and Psychology, 29*(1), 1-19.

Deng, L. & Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educational and Pscyhological Measurement*, *77*(2), 185-203.

Derks, B., Laar, C. van, Ellemers, N. & Groot, K. de. (2015). Gender-bias primes elicit queen-bee responses among senior policewomen. *Psychological Science, 22*(10), 1243-1249.

Döring, M. & Willems, J. (2021). Processing stereotypes: professionalism confirmed or disconfirmed by sector affiliation? *International Public Management Journal*, 1-19.

Engbersen, G., Bochove, M. van, Boom, J. de, Bussemaker, J., Farisi, B. el, Krouwel, A., Lindert, J. van, Rusinovic, K., Snel, E., Heck, L. van, Veen, H. van der & Wensveen, P. van. (2021). *De laag-vertrouwenssamenleving.* Erasmus School of Social and Behavioural Sciences & Kenniswerkplaats Leefbare Wijken.

Ertl, B., Luttenberger, S. & Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with a under-representation of females. *Frontiers in Psychology, 8*(2017).

European Commission. (2017). *Active labour market policies* [Factsheet]. Retrieved at 25 March 2022, from https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/eu-economic-governance-monitoring-prevention-correction/european-semester/thematic-factsheets/labour-markets-and-skills_nl.

Field, A. (2009). *Discovering statistics using R*. Sage Publications.

Field, A. (2012). *Discovering statistics using SPSS*. Sage Publications.

Folkman, S. (2016). *Ways of Coping – updated* [letter including updated questionnaire]. Retrieved on 14 March 2022, at https://prevention.ucsf.edu/sites/prevention.ucsf.edu/files/uploads/tools/surveys/pdf/Ways%20of%20coping.pdf.

Folkman, S. & Lazarus, R.S. (1988). Coping as mediator of emotion. *Journal of Personality and Social Psychology,, 54*(3), 466-475.

Fox, J. & Weisberg, S. (2019). *An R companion to applied regression* [R Package]. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Galesic, M. & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly, 73*(2), 349-360.

Garnefski, N., Legerstee, J., Kraaij, V., Kommer, T. van den & Teerds, J.A.N. (2002). Cognitive coping strategies and symptoms of depression and anxiety: a comparison between adolescents and adults. *Journal of Adolescence, 25*(6), 603-611.

Garrett, R.S., Thurber, J.A., Fritschler, A.L. & Rosenbloom, D.H. (2006). Assessing the impact of bureaucracy bashing by electoral campaigns. *Public Administration Review, 66*(2), 228-240.

Gigerenzer, G. (2010). Moral satisficing: rethinking moral behavior as bounded rationality. *Topics in Cognitive Science, 2*(2010), 528-554.

Goodsell, C.T. (2014). *The New Case for Bureaucracy*. CQ Press.

Gravetter, F.J. & Wallnau, L.B. (2017). *Statistics for the behavioral sciences*. Cengage Learning Inc..

Greszki, R., Meyer, M. & Schoen, H. (2015). Exploring the effects of removing 'too fast' responses and respondents from web surveys. *Public Opinion Quarterly, 79*(2), 471-503.

Hackshaw, A. (2008). Small studies: strengths and limitations. *European Respiratory Journal, 32*(5), 1141-1143.

Hauser, D.J., Ellsworth, P.C. & Gonzalez, R. (2018). Are manipulation checks necessary? *Frontiers in Psychology, 9*(2018), 1-10.

Hayes, A.F. & Rockwood, N.J. (2016). Regression-based statistical mediation and moderation analysis in clinical research: observations, recommendations and implementation. *Behaviour Research and Therapy, 98*(2017), 1-19.

Hertwig, R. (2017). When to consider boosting: some rules for policy-makers. *Behavioural Public Policy, 1*(2), 143-161.

Hindmarch, P., Hawkins, A., McColl, E., Hayes, M., Majsak-Newman, G., Ablewhite, J., Deave, T. & Kendrick, D. (2015). Recruitment and retention strategies and the examination of attrition bias in a randomised controlled trial in children's centres serving families in disadvantages areas of England. *Trials, 16*(79), 1-13.

Hively, K. & El-Alayli, A. (2014). You throw like a girl: the effect of stereotype threat on women's athletic performance and gender stereotypes. *Psychology of Sport and Exercise, 15*(1), 48-55.

Höpfl, H.M. (2006). Post-bureaucracy and Weber's 'modern' bureaucrat. *Journal of Organizational Change Management, 19*(1), 8-21.

Huntington-Klein, N. (2021). *Vtable: variable table for variable documentation* [R Package]. https://CRAN.R-project.org/package=vtable.

Hvidman, U. & Andersen, S.C. (2015). Perceptions of public and private performance: evidence from a survey experiment. *Public Administration Review, 76*(1), 111-120.

Inzlicht, M. & Schmader, T. (2011). *Stereotype threat: theory, process, and application*. Oxford University Press.

Jahan, F. & Shahan, A.M. (2012). Bureau bashing and public service motivation: a case for the civil service in Bangladesh. *International Journal of Public Administration, 35*(4), 272-824.

Jalagat, R. (2016). Job performance, job satisfaction, and motivation: a critical review of their relationship. *International Journal of Advances in Management and Economics, 5*(6), 36-42.

Jilke, S., Walle, S. van de & Kim, S. (2015). Generating usable knowledge through an experimental approach to public administration. *Public Administration Review, 76*(1), 69-72.

Johari, J. & Yahya, K.K. (2018). Organizational structure, work involvement, and job performance of public servants. *International Journal of Public Administration, 42*(8), 654-663.

Johnson, H.J., Barnard-Brak, L., Saxon, T.F. & Johnson, M.K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education, 80*(2), 137-149.

Jones, B.D. (1999). Bounded rationality. *Annual Review of Political Science, 2*(1), 297-321.

Julnes, G. (2008). Identifying dimensions and types in public administration research : introduction to principal component analysis, factor analysis and cluster analysis. In G.J. Miller & K. Yang (Eds.), *Handbook of research methods in public administration* (pp. 515-563). CRC press.

Jussim, L., Harber, K.D., Crawford, J.T., Cain, T.R. & Cohen, F. (2005). Social reality makes the social mind. *Interaction Studies, 6*(1), 85-102.

Kane, J.V. & Barabas, J. (2019). No harm in checking: using factual manipulation checks to assess attentiveness in experiments. *American Journal of Political Science,63*(1), 234-249.

Kenny, D. (2004). Correlational inference. In D. Kenny (Ed.), *Correlation and causation* (pp. 1-14). John Wiley.

Keren, G. (2014). Between- or within-subjects design: a methodological dilemma. In: G. Keren & C. Lewis (Eds.), *A Handbook for Data Analysis in the Behaviorial Sciences: Volume 1* (pp. 257-272). Lawrence Erlbaum Associates.

Kevins, A. & Vis, B. (2021). Do public consultations reduce blame attribution? The impact of consultation characteristics, gender and gender attitudes. *Political Behavior, 2021,* 1-22.

Kim, J., Kim, S., Byun, M. (2020). Emotional distancing in nursing: a concept analysis. *Nursing Forum, 55*(4), 595-602.

Knill, C. & Tosun, J. (2017). Policy-making. In D. Caramani (Ed.), *Comparative Politics* (pp. 349-362). Oxford University Press.

Koopmans, L.., Bernaards, C.M., Hildebrandt, V.H., Schaufeli, W.B., Vet, H. de, Beek A.J. van der. (2011). Conceptual frameworks of individual work performance. Journal of Occupational and *Environmental Medicine, 53*(8), 856-866.

Krumpal, I. (2011). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity, 47*(4), 2025-2047.

Kumari, G. & Pandey, K.M. (2011). Job satisfaction in public sector and private sector: a comparison. *International Journal of Innovation, Management and Technology, 2*(3), 222-228.

Leisink, P. & Steijn, B. (2010). Public service motivation and job performance of public sector employees in the Netherlands. *International Review of Administrative Sciences,75*(1), 35-52.

Lipsky, M. (2010). *Street-level bureaucracy*. Russel Sage Foundation.

Loon, N. van, Kjeldsen, A.M., Andersen, L.B., Vandenabeele, W. & Leisink, P. (2018). Only when the societal impact potential is high? A panel study of the relationship between public service motivation and perceived performance. *Review of Public Personnel Administration, 38*(2), 139-166.

Maier, M.J. (2015). *R: Einführung durch angewandte statistik* [R Package]. http://CRAN.R-project.org/package=REdaS.

Marume, S.B.M., Mutongi, C. & Madizyire, N.C. (2016). An analysis of public policy implementation. *Journal of Business and Management, 18*(4), 86-93.

Marvel, J.D. (2015). Public opinion and public sector performance: are individuals' beliefs about performance evidence-based or the product of anti-public sector bias? *International Public Management Journal, 18*(2), 209-227.

Marvel, J.D. (2016). Unconscious bias in citizens' evaluations of public sector performance. *Journal of Public Administration Research and Theory, 26*(1), 143-158.

Mendes, W.B. & Jamieson, J. (2011). Embodied stereotype threat. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: theory, process, and application*. Oxford University Press.

Miller, R.B. & Hollist, C. (2007). Attrition bias. In N. Salkind (Ed.), *Encyclopaedia of Measurement and Statistics* (pp. 57-60). Sage Publications.

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2020). *Trends & Cijfers 2020*. BZK.

Motowidlo, S.J. & Kell, H.J. (2003). Job performance. In W.C. Borman, D.R. Ilgen & R.J. Klimoski (Eds.) *Handbook of psychology: industrial and organizational psychology* (pp. 39-53). John Wiley & Sons Inc.

Moynihan, D. (2020). Populism and the Deep State: the attack on public service under Trump. In M.W. Bauer, B. Guy Peters, J. Pierre, K. Yesilkagit, S. Becker (Eds.), *Liberal-Democratic Backsliding and Public Administration* (pp. 1-26). Cambridge University Press.

Moynihan, D. & Roberts, A. (2021). Dysfunction by design: Trumpism as administrative doctrine. *Public Administration Review, 81*(1), 152-156.

Napier, J.L., Suppes, A. & Bettinsoli, M.L. (2020). Denial of gender discrimination is associated with better subjective well-being among women: a system justification account. *European Journal of Social Psychology, 50*(6), 1191-1209.

Needham, C. & Mangan, C. (2016). The 21st-century public servant: working at three boundaries of public and private. *Public Money & Management, 36*(4), 265-272.

Nguyen, H.H.D. & Ryan, A.M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314-1334.

Nielsen, V.L. (2006). Are street-level bureaucrats compelled or enticed to cope? *Public Administration, 84*(4), 861-889.

NOS. (2021, 14 March). Haagse plannen: bureaucratiestop, geld voor verpleging en gratis relatietheratpie. *NOS*. Retrieved at March 2 2022, from https://nos.nl/artikel/2372553-haagse-plannen-bureaucratiestop-geld-voor-verpleging-en-gratis-relatietherapie.

Nosek, B.A., Beck, E.D., Campbell, L., Flake, J.K., Hardwicke, T.E., Mellor, D.T., Veer, A.E van ‘t & Vazire, S. (2019). Preregistration is hard and worthwhile. *Trends in Cognitive Sciences, 23*(1), 815-818.

Nunan, D., Aronson, J. & Bankhead, C. (2018). Catalogue of bias: attrition bias. *BMJ Evidence Based Medicine, 23*(1), 21-22.

Olsen, J.P. (2005). Maybe it is time to rediscover bureaucracy. *Journal of Public Administration Research and Theory, 16*(1), 1-24.

Osborne, S.P. (2021). *Public service logic: creating value for public service users, citizens, and society through public service delivery.* Routledge.

Oxford Languages. (n.d.). *Bureaucrat.* Lexico. Retrieved at 27 March 2022, from https://www.lexico.com/definition/bureaucrat.

Pautz, M.C. & Warnement, M.K. (2013). Government on the silver screen: contemporary American cinema’s depiction of bureaucrats, police officers, and soldiers. *Political Science & Politics, 46*(3), 569-579.

Perry, J.L. (1997). Antecedents of Public Service Motivation. *Journal of Public Administration Research and Theory*, *1997*(2), 181-197.

Perry, J.L. & Wise, L.R. (1990). The motivational bases of public service. *Public Administration Review, 1990*, 367-373.

Peters, B.G. (2009). Still the century of bureaucracy?: the roles of public servants. In J. Fenwick & J. McMillan (Eds.), *Public Management in the Postmodern Era* (pp. 7-21). Edward Elgar Publishing, Inc.

Peterson, R.A. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters, 11*(3), 261-275.

Pino, E. del, Calzada, I. & Díaz-Pulido, J.M. (2016). Conceptualizing and explaining bureauphobia: contours, scope, and determinants. *Public Administration Review, 76*(5), 725-736.

R Core Team (2021). *R: A language and environment for statistical computing* [R Package]. R Foundation for Statistical Computing. https://www.R-project.org/.

Revelle, W. (2022). *Psych: procedures for personality and psychological research* [R Package]. https://CRAN.R-project.org/package=psych.

Richeson, J.A. & Shelton, J.N. (2011). Stereotype threat in interracial interactions. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: theory, process, and application*. Oxford University Press.

Rosen, B., Preisman, M., Read, H., Chaukos, D., Greenberg, R.A., Jeffs, L., Maunder, R. & Wiesenfeld, L. (2022). Resilience coaching for healthcare workers: experiences of receiving collegial support during the COVID-19 pandemic. *General Hospital Psychiatry, 75*(2022), 83-87.

Sager, F. & Rosser, C. (2009). Weber, Wilson and Hegel's theory of modern bureaucracy. *Public Administration Review, 69*(6), 1136-1147.

Sarnacchiaro, P., Camminatiello, I., D'Ambra, L. & Palma, R. (2019). How does public service motivation affect teachers self-reported performance in an education system? Evidence from an empirical analysis in Italy. *Quality & Quantity, 53*(5), 2521-2533.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350-353.

Schwarzer, R. & Knoll, N. (2008). Functional roles of social support within the stress and coping process: a theoretical and empirical overview. *International Journal of Psychology, 42*(4), 243-252.

Shannon, B.N., McGee, Z.A. & Jones, B.D. (2019). Bounded rationality and cognitive limits in political decision making. In *Oxford Research Encyclopedia of Politics* (pp. 1-25). Oxford University Press.

Shapiro, J.R. (2011). Different groups, different threats: a multi-threat approach to the experience of stereotype threats. *Personality and Social Psychology Bulletin, 37*(4), 464-480.

Shaxson, L. (2005). Is your evidence robust enough? Questions for policy makers and practioners. *Evidence & Policy, 1*(1), 101-111.

Shumaker, S.A. & Brownell, A. (1984). Toward a theory of social support: closing conceptual gaps. *Journal of Social Issues, 40*(4), 11-36.

Simon, H.A. (2000). Bounded rationality in social science: today and tomorrow. *Mind & Society, 1*(1), 25-39.

Simone, S. De, Cicotto, G., Pinna, R. & Giustiniano, L. (2016). Engaging public servants. *Management Decisions, 54*(7), 1596-1594.

Skok, J.E. (1995). Policy issue networks and the public policy cycle: a structural-functional framework for public administration. *Public Administration Review, 55*(4), 325-332.

Sommers, R. & Miller, F.G. (2013). Forgoing debriefing in deceptive research: is it ever ethical? *Ethics & Behavior, 23*(2), 98-116.

Sonnentag, S., Volmer, J. & Spychala, A. (2008). Job performance. In: J. Barling (Ed.), *Micro approaches* (pp. 427-447). Sage.

Spencer, S.J., Logel, C. & Davies, P.G. (2016). Stereotype threat. *Annual Review Psychology, 67*(2016), 415-437.

Steenhuisen, B. & Eeten, M. van. (2012). Patterns of coping with inconsistent demands in public service delivery. *Administration & Society, 45*(9), 1130-1157.

Stolwijk, S. (2019). The representativeness heuristic in political decision making. In *Oxford Research Encyclopedia of Politics* (pp. 1-25). Oxford University Press.

Suzuki, K. (2020). Why are some civil servants more committed to professional norms than others? *Universiteit Leiden.* Retrieved at 14 March 2022, on https://www.universiteitleiden.nl/en/research/research-projects/governance-and-global-affairs/why-are-some-civil-servants-more-committed-to-professional-norms-than-others.

Szydlowski, G., Boer, N. de & Tummers, L. (2022). Compassion, bureaucrat bashing and public administration. *Public Administration Review* [available, but not yet published], 1-59.

Tummers, L.G. (2016). The relationship between coping and job performance. *Journal of Public Administration Research and Theory, 27*(1), 150-162.

Tummers, L.G., Bekkers, V., Vink, E. & Musheno, M. (2013). *Handling stress during policy implementation: developing a classification of 'coping' by frontline workers based on a systematic review* [Conference]. IRSPM Conference Prague.

Tummers, L.G., Bekkers, V., Vink, E. & Musheno, M. (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, L.G. & Knies, E. (2016). Measuring public leadership: developing scales for four key public leadership roles. *Public Administration, 94*(2), 433-451.

Tummers, L.G. & Musheno, M. (2015). *Coping during public service delivery: developing a measurement instrument for survey research* [Conference]. Public Management Research Conference, Minneapolis (June 11-13).

Vandenabeele, W. (2008). Development of a Public Service Motivation measurement scale: corroborating and extending Perry's measurement instrument. *International Public Management Journal, 11*(1), 143-167.

Vandenabeele, W. (2009). The mediating effect of job satisfaction and organizational commitment on self-reported performance: more robust evidence of the PSM-performance relationship. *International Review of Administrative Sciences, 75*(1), 11-34.

Veelen, R. van & Veldman, J. (2020). Distancing from a stigmatized social identity: state of the art and future research agenda on self-group distancing. *European Journal of Social Psychology, 50*(6), 1089-1107.

Veldman, J., Laar, C. van, Meeussen, L. & Bue, S.L. (2021). Daily coping with social identity threat in outgroup-dominated contexts: self-group distancing among female soldiers. *Personality and Social Psychology Bulletin, 47*(1), 118-130.

Viswesvaran, C. & Ones, D.S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment, 8*(4), 216-226.

Walle, S. van de. (2004). Context-specific images of the archetypical bureaucrat: persistence and diffusion of the bureaucracy stereotype. *Public Voices, 7*(1), 3-12.

Walle, S. van de, Kampen, J.K., Bouckaert, G. & Maddens, B. (2003*). Zesde rapport – burgergericht besturen: kwaliteit en vertrouwen in de overheid.* Instituut voor de Overheid.

Wickham, H. (2016). *Ggplot2: elegant graphics for data analysis* [R Package]. https://ggplot2.tidyverse.org.

Wickham, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686.

Wickham, H. & Bryan, J. (2002). *Readxl: read Excel files* [R Package]. https://CRAN.R-project.org/package=readxl.

Wickham, H., François, R., Henry, L. & Müller, K. (2022). *Dplyr: A grammar of data manipulation* [R Package]. https://CRAN.R-project.org/package=dplyr.

Willems, J. (2020). Public servant stereotypes: it is not (at) all about being lazy, greedy and corrupt. *Public Administration, 2020*(98), 807-823.

Wright, B.E. & Davis, B.S. (2002). Job satisfaction in the public sector. *American Review of Public Administration, 33*(1), 70-90.

Xia, Y. & Yang, Y. (2018). RMSEA, CFI, and TLI in structural equation modelling with ordered categorical data: the story they tell depends on the estimation methods. *Behavior Research Methods, 51*(1), 409-248.

Zacka, B. (2017). *When the state meets the street.* The Belknap Press of Harvard University Press.

Appendix 1

English versions of vignettes

ENGLISH - CONTROL

Please read this *fictitious* newspaper article carefully and answer the subsequent questions as if you were a public servant working at this governmental organization.

An active labour market, but what about the activity of NSAA?

By the editorial board - 7 April 2022

The NSAA (National Social Affairs Agency) is subjected to critical scrutiny because it is taking the NSAA too long to update an outdated policy.

The NSAA's core task is developing active labour market policy, which helps to increase job opportunities for people. The NSAA's former active labour market policy was already outdated in 2015. Yet, it took until January 2022 for the NSAA to update its policy. Critics ask why it took NSAA *so* long (7 years) to adapt the policy. In other words, how active is NSAA?

There are three aspects within the policymaking process that contributed to this delay. First, to develop its policy, the NSAA has adopted several standard operating procedures to create a policy that is in line with the guidelines. In practice, these procedures were indeed generally followed.

Second, when a draft policy was developed, it was closely examined for its (un)intended effects. Third and finally, knowledge from experts had to be collected to create a high-quality policy. These three aspects explain why NSAA needed seven years to update its policy.

Thus, how active is the NSAA? The NSAA is certainly active, but the policymaking process can sometimes cause delay and is part of the deal to make high-quality policy.

ENGLISH - INTERVENTION

Please read this *fictitious* newspaper article carefully and answer the subsequent questions as if you were a public servant working at this governmental organization.

An active labour market, but what about the activity of NSAA?

By the editorial board – 7 April 2022

The NSAA (National Social Affairs Agency) is subjected to critical scrutiny because it is taking the NSAA too long to update an outdated policy.

The NSAA's core task is developing active labour market policy, which helps to increase job opportunities for people. The NSAA's former active labour market policy was already outdated in 2015. Yet, it took until January 2022 for the NSAA to update its policy. Critics ask why it took NSAA *so* long (7 years) to adapt to the changes and update the policy. In other words, how active is NSAA?

There are three aspects, mainly concerning the questionable functioning of public servants, that contributed to this delay. First, to develop its policy, the NSAA has adopted several standard operating procedures to create a policy that is in line with the guidelines. In practice, the procedures were too strictly followed, leading to inefficiency by public servants.

Second, when a draft new policy was developed, it was closely examined for its (un)intended effects. Third and finally, knowledge from experts had to be collected, because of the extent of incompetency and lack of appropriate knowledge of public servants. These three aspects explain why NSAA needed seven years to update its policy.

Thus, how active is the NSAA? The NSAA is certainly active, but the public servants within the NSAA have clearly slowed down the process, which is unacceptable.

Dutch versions of vignettes

NEDERLANDS – CONTROLE GROEP

Lees onderstaand fictief krantenartikel aandachtig door en beantwoord vervolgens de vragen alsof u een ambtenaar bent die bij deze overheidsorganisatie werkt.

Een actieve arbeidsmarkt, maar hoe zit het met de activiteit van NASZ?

Door de redactie – 7 April 2022

Het NASZ (Nationaal Agentschap voor Sociale Zaken) wordt kritisch onder de loep genomen, omdat het te lang nodig had om een om verouderd beleid te vernieuwen.

De belangrijkste taak van het NASZ is het ontwikkelen van actief arbeidsmarktbeleid, dat gefocust is op het vergroten van de arbeidskansen. Het vorige actieve arbeidsmarktbeleid was al verouderd in 2015. Toch had het NASZ tot januari 2022 nodig om dit beleid te vernieuwen. Critici vragen zich af waarom het NASZ zó lang (7 jaar) heeft geduurd om het beleid aan te passen. Met andere woorden, hoe actief is NASZ zelf?

Er zijn drie aspecten binnen het beleidsvormingsproces die bij hebben gedragen aan deze vertraging. Ten eerste, gebruikt het NASZ verschillende *standard operating procedures* om een beleid te creëren dat in lijn is met de richtlijnen. In de praktijk worden deze procedures inderdaad over het algemeen gevolgd.

Ten tweede werd een nieuw concept van het beleid telkens nauwkeurig onderzocht op (on)bedoelde effecten. Ten slotte moest kennis van experts worden verzameld om zo een kwalitatief hoogstaand beleid te creëren. Deze drie aspecten verklaren waarom NASZ 7 jaar tijd nodig had om het beleid te vernieuwen.

Dus, hoe actief is NASZ? Het NASZ is zeker actief, maar het beleidsvormingsproces kan soms leiden tot vertraging en moet op de koop toe worden genomen om kwalitatief hoogwaardig beleid te maken.

NEDERLANDS – INTERVENTIE GROEP

Lees onderstaand fictief krantenartikel aandachtig door en beantwoord vervolgens de vragen alsof u een ambtenaar bent die bij deze overheidsorganisatie werkt.

Een actieve arbeidsmarkt, maar hoe zit het met de activiteit van NASZ?

Door de redactie – 7 April 2022

Het NASZ (Nationaal Agentschap voor Sociale Zaken) wordt kritisch onder de loep genomen, omdat het te lang nodig had om een om verouderd beleid te vernieuwen.

De belangrijkste taak van het NASZ is het ontwikkelen van actief arbeidsmarktbeleid, dat gefocust is op het vergroten van de arbeidskansen. Het vorige actieve arbeidsmarktbeleid was al verouderd in 2015. Toch had het NASZ tot januari 2022 nodig om dit beleid te vernieuwen. Critici vragen zich af waarom het NASZ zó lang (7 jaar) heeft geduurd om het beleid aan te passen. Met andere woorden, hoe actief is NASZ zelf?

Er zijn drie aspecten, specifiek rondom het discutabele functioneren van de ambtenaren, die bij hebben gedragen aan deze vertraging. Ten eerste, gebruikt het NASZ verschillende *standard operating procedures* om een beleid te creëren dat in lijn is met de richtlijnen. In de praktijk worden de procedures te strikt gevolgd, wat leidt tot inefficiëntie bij ambtenaren.

Ten tweede werd een nieuw concept van het beleid telkens nauwkeurig onderzocht op (on)bedoelde effecten. Ten slotte moest kennis van experts worden verzameld, vanwege de mate van incompetentie en gebrek aan passende kennis bij ambtenaren. Deze drie aspecten verklaren waarom NASZ 7 jaar tijd nodig had om het beleid te vernieuwen.

Dus, hoe actief is NASZ? Het NASZ is zeker actief, maar de ambtenaren werkend binnen het NASZ laten duidelijk het proces vertragen; dit is onacceptabel.

Appendix 2

English version of the survey

Welcome to this survey! First of all, I would like to thank you for your time to participate in this research. My name is Julia Wesdorp and I am a student at the master Governance and Policy at Utrecht University. This survey is part of my master thesis, where I study the behavior of public servants in different situations.

All the information gathered in this survey will be treated confidentially and anonymously. The data will be saved in a secured location (Yoda) hosted by the University of Utrecht. Only my supervisor (Barbara Vis, Universiteit Utrecht) and myself have insight in the surveys. Information will only be presented in an aggregated form, which means that none of the data can be traced back to any individuals or organizations. Besides, I will process all the information according to the GDPR. You have the right to terminate your participation in this study at any time without any reason.

It will take approximately 8 minutes to fill out the survey.
If you have any questions or suggestions about this research, please do not hesitate to contact me at j.v.wesdorp@uu.nl or my supervisor Barbara Vis (b.vis@uu.nl).

Q1. Do you agree with your participation in this research?

0 Yes, I agree with the participation in this research.

-> Continue survey

0 No, I do not agree with the participation in this research.

-> Survey will be ended.

First, some general questions:

Q2. What is your gender? *not compulsory*

0 Male

0 Female

0 Other

0 Prefer not to say

Q3. At which kind of governmental organization do you work? *not compulsory*

0 Ministry

0 Province

0 Municipality

0 Waterboards

0 Umbrella organizations (e.g. IPO, Unie van Waterschappen, VDP, VNG)

0 Governmental agencies (e.g. NVWA, RWS, RIVM, DUO, IND)

0 Independent administrative bodies (e.g. AFM, KvK, CBR)

0 Other: ….

Q4. How many years of working experience do you have within the public sector? *not compulsory*

0 Between 1 and 5 years

0 Between 6 and 10 years

0 Between 11 and 15 years

0 16 or more years

HERE RESPONDENTS RANDOMLY RECEIVE EITHER CONTROL OR INTERVENTION VIGNETTE (SEE APPENDIX 1)

Q5. Within which policy domain is the governmental organization in the news article working?

0 Environmental domain

0 Education domain

0 Social domain

0 Infrastructure domain

Q6. Could you please rate the extent to which you feel each statement applies to the text you have read?

The extent to which public servants are criticized.

0 1 2 3 4

The extent to which public servants are attacked.

0 1 2 3 4

The extent to which public servants are blamed.

0 1 2 3 4

The extent to which public servants are generalized.

0 1 2 3 4

The extent to which public servants are criticized in a meaningless way.

0 1 2 3 4

[measured on five-point Likert scale: 0 = not at all; 1 = a little bit; 2 = fairly; 3 = quite a bit; 4 = very much]

The following questions ask you something about your performance as a public servant in your daily life.

Q7. Please rate the extent to which you feel the statement apply to you.

In my opinion, I contribute to the success of the organization.

1 2 3 4 5

I think I am performing well within this organization.

1 2 3 4 5

I think I am a good employee.

1 2 3 4 5

On average, I work harder than my colleagues.

1 2 3 4 5

[measured on five-point Likert scale -> 1 = Strongly disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly agree].

The next questions ask you something about how you would cope with a situation that is given in the news article (only for intervention group).

Q8. How would you deal with the news article that you have read? Indicate to what extent the statements would apply to you, as if you were working at this organization.

I would made light of the situation and refuse to get too serious about it.

0 1 2 3

I would go on as if nothing had happened.

0 1 2 3

I would not let it get to me and refuse to think much about it.

0 1 2 3

I would try to forget the whole news article.

0 1 2 3

I would try to look on the bright sight of things

0 1 2 3

[measured on four-point Likert scale; 0 = not used; 1 = used somewhat; 2 = used quite a bit; 3 = used a great deal].

Q9. Here are some other ways to deal with the news article that you have read. Indicate to what extent the statements would apply to you, as if you were working at this organization.

I would talk to a colleague to find out more about the situation.

0 1 2 3

I would talk to a colleague who could do something concrete for me about the negative effects I experience (e.g. supervisor or confidant).

0 1 2 3

I would ask a near colleague, who I respect, for advice.

0 1 2 3

I would talk to a colleague about how I am feeling.

0 1 2 3

I would seek for understanding from someone.

0 1 2 3

I would get professional help.

0 1 2 3

[measured on four-point Likert scale; 0 = not used; 1 = used somewhat; 2 = used quite a bit; 3 = used a great deal].

End of survey

You have finished the survey, thank you for your time! If you would like to receive the final thesis, you can insert your mail address here: [insert email address, not obligatory].

Before you close this page…

…I would like to tell you something more about this research. The research investigates how public servants behave in different situations, but more specifically I research the effect of negative stereotypes on public servants on their performance. This is done by doing an experiment, where there were two random groups. One group has read a news article with some general criticisms and no negative stereotypes regarding public servants, the other group has read a new article with specific individual criticisms focused on the functioning of the public servants. In this way, I can research what the effect is of negative stereotypes on public servants.

If you feel that you are disadvantaged by this experiment, please do not hesitate to contact me (j.v.wesdorp@uu.nl).

Dutch version of the survey

Welkom bij deze vragenlijst! Allereerst wil ik u bedanken voor de tijd die u wilt nemen om de vragenlijst in te vullen. Mijn naam is Julia Wesdorp en ik ben student van de master Bestuur en Beleid aan de Universiteit Utrecht. Deze vragenlijst is onderdeel van mijn masterscriptie, waarin ik het gedrag van ambtenaren in verschillende situaties bestudeer.

Alle informatie die wordt verzameld in deze vragenlijst worden anoniem en vertrouwelijk behandeld. De data wordt opgeslagen op een beveiligde locatie, welke wordt beheerd door de Universiteit Utrecht. Alleen mijn supervisor (Barbara Vis, Universiteit Utrecht) en ik hebben inzicht in de vragenlijsten. De informatie wordt enkel geaggregeerd gepresenteerd, wat betekent dat geen enkele data terug kan worden herleid naar een individu of een organisatie. Daarnaast behandel ik alle informatie in overeenstemming met de AVG. U heeft altijd het recht om deelname te beëindigen zonder opgaaf van reden.

Het invullen van de survey zal ongeveer 8 minuten in beslag nemen.

Als u vragen of suggesties heeft over dit onderzoek, aarzel niet om mij (j.v.wesdorp@uu.nl) of mijn supervisor (b.vis@uu.nl) te benaderen.

Q1. Stemt u in met deelname in dit onderzoek?

0 Ja, ik stem in met deelname in dit onderzoek.

-> Doorgaan met survey

0 Nee, ik stem niet in met deelname in dit onderzoek.

-> Survey wordt direct beëindigd

Allereerst een aantal algemene vragen:

Q2. Wat is uw geslacht? *niet verplicht*

0 Man

0 Vrouw

0 Anders

0 Wil ik liever niet zeggen

Q3. Bij welke overheidsorganisatie werkt u op dit moment? *niet verplicht*

0 Ministerie

0 Provincie

0 Gemeente

0 Waterschap

0 Koepelorganisatie (bijvoorbeeld IPO, Unie van Waterschappen, VDP, VNG)

0 Agentschappen (bijvoorbeeld NVWA, RWS, RIVM, DUO, IND)

0 Zelfstandig bestuursorgaan (bijvoorbeeld AFM, KvK, CBR)

0 Anders: …

Q4. Hoeveel jaar werkervaring heeft u binnen de publieke sector? *niet verplicht*

0 Tussen de 1 en 5 jaar

0 Tussen de 6 en 10 jaar

0 Tussen de 11 en 15 jaar

0 Meer dan 16 jaar

HIER KRIJGEN DE RESPONDENTEN WILLEKEURIG EEN CONTROLE OF EEN INTERVENTIE VIGNETTE (ZIE APPENDIX 1).

Q5. Binnen welk beleidsdomein is de overheidsorganisatie vanuit het nieuwsartikel actief?

0 Klimaat en milieu

0 Onderwijs

0 Sociale zaken

0 Infrastructuur

Q6. Kunt u aangeven in welke mate u het gevoel heeft dat onderstaande stellingen toepasselijk zijn op het gelezen nieuwsartikel?

De mate waarin ambtenaren worden bekritiseerd.

0 1 2 3 4

De mate waarin ambtenaren worden aangevallen.

0 1 2 3 4

De mate waarin ambtenaren van iets wordt beschuldigd.

0 1 2 3 4

De mate waarin ambtenaren op dezelfde wijze worden behandeld.

0 1 2 3 4

De mate waarin ambtenaren worden bekritiseerd op een nietszeggende manier.

0 1 2 3 4

[gemeten op een vijf-punts Likert schaal: 0 = helemaal niet; 1 = een klein beetje; 2 = behoorlijk; 3 = aardig veel; 4 = erg veel]

De volgende vragen gaan over uw prestatie als ambtenaar in het dagelijks leven.

Q7. Geef aan in hoeverre u het gevoel heeft dat de stelling op u van toepassing is.

Naar mijn mening draag ik bij aan het succes van de organisatie.

1 2 3 4 5

Ik denk dat ik goed presteer binnen mijn organisatie.

1 2 3 4 5

Ik denk dat ik een goede werknemer ben.

1 2 3 4 5

Gemiddeld genomen werk ik harder dan mijn collega's.

1 2 3 4 5

[gemeten op een vijf-punts Likert schaal: 1= helemaal oneens; 2 = oneens; 3 = neutraal; 4 = eens; 5 = helemaal eens]

De volgende vragen gaan over hoe u zou omgaan met een situatie, zoals gegeven in het nieuwsartikel (alleen voor de interventie groep).

Q8. Hoe zou u omgaan met het nieuwsartikel dat u heeft gelezen? Geef aan wat op u van toepassing is, als u zou werken in deze organisatie.

Ik zou luchtig over de situatie doen en het niet al te serieus nemen.

0 1 2 3

` Ik zou doorgaan alsof er niets is gebeurd.

0 1 2 3

Ik zou het me niet laten beïnvloeden en er niet te veel over nadenken.

0 1 2 3

Ik zou proberen het hele nieuwsartikel te vergeten.

0 1 2 3

Ik zou proberen de positieve kant ervan te bekijken.

0 1 2 3

[gemeten op een vijf-punts Likert schaal: 0 = niet gebruiken; 1 = soms gebruiken; 2 = regelmatig gebruiken; 3 = erg vaak gebruiken].

Q9. Ten slotte nog een aantal andere manieren die een beeld kunnen geven hoe u zou kunnen omgaan met het nieuwsartikel. Geef aan wat op u van toepassing is, als u zou werken in deze organisatie.

Ik zou met een collega praten om meer te weten te komen over de situatie.

0 1 2 3

Ik zou met een collega praten, die iets concreets voor me kan doen op het gebied van de negatieve effecten die ik ervaar (bijvoorbeeld een leidinggevende of een vertrouwenspersoon).

0 1 2 3

Ik zou een naaste collega, die ik respecteer, om advies vragen.

0 1 2 3

Ik zou met een collega praten over hoe ik me voel.

0 1 2 3

Ik zou naar begrip zoeken van iemand.

0 1 2 3

Ik zou professionele hulp zoeken.

0 1 2 3

[gemeten op een vijf-punts Likert schaal: 0 = niet gebruiken; 1 = soms gebruiken; 2 = regelmatig gebruiken; 3 = erg vaak gebruiken].

Einde van de survey

U heeft de survey afgerond, dank u voor uw tijd! Als u mijn scriptie uiteindelijk wilt ontvangen, kunt u hier uw email adres opgeven: [opgeven emailadres, niet verplicht].

Voordat u deze pagina sluit…

…wil ik u nog wat vertellen over dit onderzoek. Het onderzoek kijkt naar hoe ambtenaren zich gedragen in verschillende situaties, maar specifiek ben ik geïnteresseerd in het effect van negatieve stereotyperingen richting ambtenaren op hun prestatie. Dit is onderzocht aan de hand van een experiment met twee groepen. Een groep heeft een nieuwsartikel gelezen met algemene kritiek, maar geen negatieve stereotyperingen richting ambtenaren; de andere groep heeft een nieuwsartikel gelezen met specifieke individuele kritiek gericht op het functioneren van de ambtenaar. Op deze manier kan ik onderzoeken wat het effect is van negatieve stereotyperingen op ambtenaren.

Als u het gevoel heeft dat u bent benadeeld door dit experiment, aarzel niet om mij te benaderen (j.v.wesdorp@uu.nl).

Appendix 3 – Factor analysis

A3.1 What is a factor analysis?

The method chapter already described the results of the factor analyses. To create a better understanding of the results, it is relevant to create a small summary of the factor analysis and which settings are used.

One uses the factor analysis to look at a possible underlying structure within certain items. This means that we can expect that some of the items (the questions that are asked to measure a specific concept) of a construct are connected in a specific way, namely with a common underlying factor. This factor can explain the covariation among certain items. Consequently, the factor analysis also indicates whether certain items measure something different than other items, which shows that there are two (or more) different underlying factors (Julnes, 2008, pp. 528-530). Because this research uses constructs that have several items and because they are translated to Dutch, it is important to test the underlying factors.

To conduct a factor analysis, the researcher must choose a few settings. First, there is a lot of discussion within the literature about the minimal factor loading. Based on the study of Peterson (2000, p. 264), who has looked into this issue, this research sets the minimum on .32. Next, the researcher selects the type of rotation, where one can choose between an orthogonal or oblique rotation. With orthogonal rotation it is assumed that the identified factors are independent of each other and therefore cannot correlate; with oblique rotation factors can correlate. This research uses the *direct oblimin rotation* and checks the correlations. According to Tabachnick & Fiddell (2007, in: Brown, 2009), a correlation of .32 (or higher) shows that the factors correlate, which makes an oblimin rotation most suitable. If correlations are very low or do not exist, one switches to the orthogonal rotation (which was not necessary for this research). Next, the level of the eigenvalue is determined. The eigenvalue is the amount of variation a factor can explain; often a minimal value of 1 is used, also known as the Kaiser-rule, which this research uses (Cliff, 1988).

A3.2 Results of factor analysis

First, the KMO and Bartlett test are performed to test whether it is useful and valuable to perform a factor analysis. Because the control group does not fill out the questions regarding the coping mechanisms, the NA's are changed to 0. The results of both of the tests are in table A3.1. Both results turn out very good, showing that a factor analysis is valuable for this dataset (KMO should have a minimum of .50, the higher, the better; Bartlett's test should have a $p < .05$ (Field, 2012, p. 777)).

Table A3.1
Overview results KMO and Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.94
Bartlett's Test of Sphericity	Chi-squared	3777.74
	df	190
	sig.	.001

Note. Df = degrees of freedom. Sig. = significance level.

The results of the factor analysis are in table A3.2. The table shows that for meaningless bureaucrat bashing the five items have a common underlying factor: meaningless bureaucrat bashing. Also, for performance and social support the items have a high enough factor loading, meaning that these specific items have a common underlying factor. Only with the coping mechanism distancing does one item have a factor loading lower than the minimum of .32. This means that these items have probably a different underlying factor than all the other items. Looking at these items, it can indeed be the case that the low-loading item (*I would try to look on the bright sight of things*) has a different underlying factor than the other items: if you try to forget things or would go on as nothing has happened, one can dissociate from a certain stereotype. However, by looking at the bright sight of things, one can argue that someone accepts the stereotyping-situation and just looks at the positive side of the situation. This matches probably more with a factor such as positivity than distancing. Therefore, this item is deleted from further analysis. After deleting the item, the Weighted Omega for each of the scales is calculated: all the scales have a reliability greater than .70, which means that the reliability of the scales is good (see table A3.2 for the Weighted Omega).

Table A3.2

Factor loadings for Exploratory Factor Analysis with oblimin rotation of used scales

	Loading			
Construct and item	Bashing	Performance	Distancing	Social support
Meaningless bureaucrat bashing, Ω = .90				
The extent to which public servants are criticized.	**.806**	.001	.001	.001
The extent to which public servants are attacked.	**.907**	.001	.001	.001
The extent to which public servants are blamed.	**.832**	.001	.001	.001
The extent to which public servants are generalized.	**.411**	.120	.001	.001
The extent to which public servants are criticized in a meaningless way.	**.552**	.001	.001	.001
Performance, Ω = .79				
In my opinion, I contribute to the success of the organization.	.001	**.495**	.001	.001
I think I am performing well within this organization.	.001	**.768**	.001	.001
I think I am a good employee.	.001	**.798**	.001	.001
On average, I work harder than my colleagues.	.001	**.448**	.001	.001
Coping - distancing, Ω = .83				
I would made light of the situation and refuse to get too serious about it.	.001	.001	**.681**	.001
I would go on as if nothing had happened.	.001	.001	**.722**	.001
I would not let it get to me and refuse to think much about it.	.001	.001	**.849**	.001
I would try to forget the whole news article.	.001	.001	**.558**	.001
~~I would try to look on the bright sight of things~~	.001	.001	~~.235~~	~~.435~~

Coping - social support, Ω = .91				
I would talk to a colleague to find out more about the situation.	.001	.001	.117	**.676**
I would talk to a colleague who could do something concrete for me about the negative effects I experience.	.001	.001	.001	**.670**
I would ask a near colleague, who I respect, for advice.	.001	.001	.001	**.856**
I would tal to a colleague about how I am feeling.	.001	.001	.001	**.875**
I would seek for understanding from someone.	.001	.001	.001	**.780**
I would get professional help.	.001	.001	-.146	**.438**

Note. Factor loadings > .32 are in bold. Items with loading < .32 are striked through and deleted from further analysis.

Appendix 4 – Independent sample *t*-test

A4.1 What is an independent sample t-test?

One uses the independent sample t-test when there are two conditions (control and intervention) and each participant get one of the two conditions (and not both). In this research this is also the case: randomly some participants got the control vignette and some participants got the intervention vignette. The test uses the means of both the conditions and compares these: if the participants of the control and intervention group significantly differ, it shows us that the intervention possibly works (Field, 2009, pp. 334-342).

Before starting an independent sample t-test it is crucial to look at the extent the assumptions of this test are fulfilled. The t-test assumptions are (Field, 2009, p. 326):

- normal distribution of the data, meaning that the data is similarly distributed on the positive and negative side with most of the data in the centre. The normality assumption can either be tested with a histogram or a Q-Q-plot (Field, 2012, pp. 19, 170-171);
- homogeneity of variance or homoskedacity, meaning that the variances between the control and intervention groups are equal. To test this assumption, the Levene's test or the NCV test can be used (Non-Constant Error Variance) (Field, 2009, p. 186; Breusch & Pagan, 1979).

A4.2 Results of the independent sample t-test

The data fulfils the assumption of a normal distribution: with the help of a Q-Q plot, one could see almost a perfect diagonal line, indicating a fairly normal distribution. Regarding the homoscedasticity assumption, the data does not show any signs of heteroskedasticity, which is good (F (1, 180) = 0.056, p = 0.814). As already stated in the results chapter, there is not a significant difference between the mean of the control (M = 3.905) and intervention (M = 3.94) regarding performance (t (177,21)= -.497, p = .62).

Recall that the control and intervention group are balanced regarding the variables gender, working organization and years of work experience. In table A4.1 the differences between the means are visible, which indeed shows a small difference between the two groups. Something else that stands out from the performed analyses is that there is a significant difference on the variables bureaucrat bashing (p < .001), meaning that there is a significant difference between the experiences of bureaucrat bashing in the control and the intervention group. This can mean that the meaningless bureaucrat bashing is experienced within the intervention group. See for the other results the main text, specifically 4.2.

Table A4.1

Balance check for gender, working years and bashing

	Control		Intervention		
Variable	M	SD	M	SD	F-test value
Gender	1.60	.659	1.483	.503	1.797
Working years	2.511	1.285	2.609	1.333	0.257
Working organization	2.874	1.925	2.954	1.892	0.08
Bureaucrat bashing	2.156	0.624	3.356	0.811	126.451*

Note. M = mean. SD = standard deviation.

* $p < .001$.

Appendix 5 – Multiple regression

A5.1 What is multiple regression?

This research uses multiple regression analysis to describe the expected causal relationship between the independent and dependent variable (Allison, 1999, pp. 1-2). *Multiple* regression is suitable in this case, because not only the dependent and the independent variables are included within this analysis, but also some control variables and eventually some moderators.

Also, it is important to check whether the data fulfil the assumptions of multiple regression. The following assumptions are relevant according to Field (2012, pp. 271-295):

- normal distribution of data, see paragraph A5.1;
- independence of the error terms, which means that no error term of an observation should correlate with another error term. This can be tested with the help of a Durbin-Watson test; the results of this test can range between 0-4, where a result of 2 points towards no correlation or independence of the error terms;
- no multicollinearity, which means that the correlations between the independent variables or control variables should not be very strong, thus no perfect correlation. This can be tested by performing a correlation analysis and searching for correlations higher than .80 or .90; if this is the case, this points to multicollinearity;
- linearity, which means that the expected relationship between the dependent and independent variable is linear. This can be tested with by using a graph: on the X-axis the *fitted values* and on the Y-axis the error term, there should be a linear relationship visible between the given points.

The results chapter discusses all the results.

Appendix 6 - Moderation

A6.1 What is a moderation analysis?

A moderator changes the strength or the direction of the relationship between the independent and dependent variables (Baron & Kenny, 1986, p. 1174). This contrasts with a mediator, which explains the relationship between the dependent and independent variables. One can use the Baron and Kenny test to test the moderation, where the researcher creates an interaction variable of the independent variable combined with the moderating variable. This newly created variable is added to the regression model of hypothesis 1 next to the individual moderator. Note that this test is only possible if there exists a main effect, which means that there should be a relationship between meaningless bureaucrat bashing and performance (and thus hypothesis 1 is true) (Baron & Kenny, 1986, p. 1174). The test is interpreted with the help of Dawson's (2014, p. 9) freely available Excel sheet, which shows the results of the Baron & Kenny test in a graph.[18]

However, as no main effect exists between the dependent and independent variables in the intervention group, the Baron and Kenny-test is not possible. As an alternative, the research uses PROCESS of Andrew Hayes to get more information about the variables (Hayes & Rockwood, 2016).

[18] See: www.jeremydawson.com/slopes.htm (choice: two-way interactions, either linear or non-linear, which is dependent on results of analysis.

Appendix 7 – Excluding 2SD

Below you can find the results of the robustness analyses done without the participants whose duration to fill out the survey took more or less than two standard deviations from the mean.

A7.1 Multiple regression

One standard deviation is 5209.47 seconds (5209.47 * 2 = 10418,934 seconds). This is a big standard deviation, but when participants left the survey open in their browser and did something else, the time keeps increasing, which can explain the high standard deviation. The results in table A7.1 show no enormous deviations from the presented model in the results chapter. Also when bashing is added, there are no huge deviations in comparison to the model presented earlier for the control group as well as the intervention group (respectively $F(4, 87) = 2.22$, $p > .05$; $F(4, 81) = 2.56$, $p = .04$).

Table A7.1

Predictors of self-reported performance in control and intervention group without participants +-2SD

Variable	Model 1 (C)[a] B (SE)		Model 2 (I)[b] B (SE)		Model 3 (C)[a] B (SE)		Model 4 (I)[b] B (SE)	
Constant	3.93	***	4.09	***	3.59	***	3.82	***
	(.11)		(.10)		(.16)		(.23)	
Control variables								
Gender	-.11		-.24	*	-.13		-.22	*
	(.09)		(.09)		(.09)		(.09)	
Working organization	.02		.09		-.04		.08	
	(.09)		(.10)		(.10)		(.10)	
Years of work experience	.03		-.14		-.01		-.13	
	(.09)		(.10)		(.09)		(.10)	
Bureaucrat bashing					.19	**	.08	
					(.07)		(.06)	
R^2	.02		.09		.09		.11	
F	.59		2.80	*	2.22		2.56	*

Note. C = control group ; I = intervention group.

[a] n = 93. [b] n = 86.

* $p < .05$. **$p < .01$. ***$p < .001$

A7.2 Moderation

Also a moderation is done for the variables distancing and social support. For both variables, the moderation stays non-significant and there are no big differences in this model in comparison to the presented model earlier (social support: $\beta = -.13$, $t(6, 81) = -1.36$, $p > .05$; distancing: $\beta = .06$, $t(6, 81) = .57$, $p > .05$).

Appendix 8 – Excluding FMC

In this appendix, the results of the robustness analyses without those who answered the factual manipulation check wrong are visible.

A8.1 Multiple regression

The right answer on the factual manipulation check is *Social Affairs*, however some people (3 in control, 3.2%; 7 in intervention, 8.0%; 5.5% of total sample size) did not choose the right option. Therefore they are eliminated in this specific analysis. In table A8.1, the results are shown. It demonstrates that in the control group there are only some small differences: gender is not significant anymore. In the intervention group we see also some differences: the model is not significant anymore and has an lower F-statistic with a smaller R-squared ($F(4, 79) = 1.68$, $p > .05$). However, the variable meaningless bureaucrat bashing remains almost the same in comparison to the other analyses. Gender, working organization and the constant are lower than in the original model, which can explain the lower *F*-statistic of the model. In short, there are no deviations that are problematic.

Table A8.1

Predictors of self-reported performance in control and intervention group without wrong FMC

	Model 1 (C)[a]		Model 2 (I)[b]		Model 3 (C)[a]		Model 4 (I)[b]	
Variable	B (SE)		B (SE)		B (SE)		B (SE)	
Constant	3.91	***	4.09	***	3.59	***	3.88	***
	(.11)		(.11)		(.17)		(.24)	
Control variables								
Gender	-.10		-.20		-.11		-.19	
	(.09)		(.10)		(.08)		(.06)	
Working organization	.02		.04		-.03		.05	
	(.10)		(.11)		(.10)		(.11)	
Years of work experience	.05		-.14		-.01		-.13	
	(.10)		(.10)		(.09)		(.10)	
Bureaucrat bashing					.18	*	.06	
					(.07)		(.06)	
R^2	.02		.07		.09		.08	
F	.55		1.92		2.01		1.68	

Note. C = control group ; I = intervention group.

[a] n = 92. [b] n = 80.

* $p < .05$. **$p < .01$. ***$p < .001$

A8.2 Moderation

The moderation with the help of PROCESS of Andrew Hayes is done with the respondents who did not answer the factual manipulation check correct. Also here, both moderations for distancing and social support remain non-significant and no considerable differences are found (social support: $\beta = -.11$, $t(6, 76) = -1.12$, $p > .05$; distancing: $\beta = .16$, $t(6, 81) = 1.43$, $p > .05$). Only with the variable distancing, the beta and the F-statistic are slightly higher, but does not changes the *p* – value.