Understanding street-level bureaucrats' attitude towards clients: Towards a measurement instrument

Erasmus University Rotterdam

ABSTRACT

The attitude of street-level bureaucrats towards their clients has an impact on the decisions they take. Still, such attitudes have not received much scholarly attention, nor are they generally studied in much detail. This article uses Breckler's psychological multicomponent model of attitude to develop a scale to measure street-level bureau-crats' general attitude towards their clients. By means of a test study ($n = 218$) and a replication study ($n = 879$), the article shows that street-level bureaucrats' attitude towards clients consists of four different components: a cognitive attitude component, a positive affective attitude component, a negative affective attitude component and a behavioral attitude component. It also establishes a conceptual and empirical distinction from related attitudes, such as prosocial motivation, work engagement, bureau-crats' rule-following identities and self-efficacy, and suggests avenues for application and further validation among different groups of street-level bureaucrats. This instrument opens up opportunities for theory testing and causality testing that surpasses case-specific considerations.

2.1 INTRODUCTION

The Weberian model portrays bureaucratic encounters as characterized by impersonality and affective neutrality to prevent bureaucrats' personal attitudes from permeating these encounters (Katz & Danet, 1973a; Merton, 1940). In reality, however, it is impossible to separate street-level bureaucrats' attitudinal dispositions and their interactions with clients (Baviskar & Winter, 2017; Lipsky, 1980; Oberfield, 2014a). Bureaucrats rely on their attitudinal dispositions to process clients to cope with complex cases, ambiguous rules and resource strain (Lipsky, 1980). By means of these attitudinal dispositions, psychological simplifications of clients emerge that "redefine [. . .] the nature of the clientele to be served" (Lipsky, 1980, p. 141). As a result, street-level bureaucrats' attitude towards clients introduces bias to the bureaucratic encounter and its outcomes (e.g., Baviskar & Winter, 2017; Kroeger, 1975; Oberfield, 2014a; Prottas, 1979; Stone, 1981), making this attitude key to understanding bureaucratic decision-making, street-level performance and official-client relations.

Attitudes are among the most studied phenomena in survey research in the field of public administration [PA] (Lee et al., 2012) and current developments in this field, such as the emergence of the behavioral PA movement (e.g., Grimmelikhuijsen et al., 2017), triggered a revived interest in this phenomenon. Despite renewed attention for the study of attitudes, the concept of street-level bureaucrats' attitude towards clients is still plagued by conceptual unclarity and unsound measurement efforts (e.g., see Lee et al., 2012). Poor measurement bears a risk of generating faulty data (DeVellis, 2003) and, subsequently, reaching erroneous conclusions (Lee et al., 2012). Despite multiple calls to improve measurement quality in our field (e.g., Perry, 2012), scale development is still scarce in PA (Grimmelikhuijsen et al., 2017; for examples, see Perry, 1996; Tummers, 2012; Tummers & Musheno, 2015).

The aim of this article is to develop a measurement instrument for street-level bureaucrats' attitude towards clients, based on material collected among one specific type of bureaucrat. Grounded in psychological attitude theory, such a scale will allow for a multidimensional measurement of attitudes towards clients, at a sufficiently abstract level as to exclude case-specific considerations. The instrument also allows to conceptually distinguish bureaucrats' attitude towards clients from other related attitudes. More in particular, we consider street-level bureaucrats' prosocial motivation, work engagement, rule-following identities and self-efficacy—all attitudes that have been studied in a public work context but that are theoretically different. The use of the psychological multicomponent model of attitudes (see Breckler, 1984) furthermore allows to move beyond current practice to rely on one-dimensional scales when measuring bureaucrats' attitudes towards clients.

Developing a measurement instrument for street-level bureaucrats' general attitude towards clients has merit for multiple reasons. First, it opens up opportunities for the quantification of bureaucrats' general assessment of clients, aiding scholars to draw in-

ferences that surpass study and context specificities (see Grimmelikhuijsen et al., 2017), which is especially challenging for street-level attitudes (Winter, 2002). Second, a validated measurement instrument of street-level bureaucrats' client-attitude with sound theoretical foundations enables theory testing and the study of causality in the bureaucratic encounter (see Pandey & Scott, 2002; Perry, 2012).

The results of a series of statistical tests among a specific sample of street-level bureaucrats show that bureaucrats' attitude towards clients consist of four distinct components: a cognitive attitude component, a positive affective attitude component, a negative affective attitude component and a behavioral attitude component. Furthermore, it establishes that this attitude is conceptually and empirically distinct from related attitudes such as prosocial motivation, work engagement or rule-following identities.

In the next section, we position street-level bureaucrats' general attitude towards clients in the bureaucratic encounter and within the broader literature. We then introduce the concept of attitude and its components as used in social psychological scholarship. This is followed by a description of methods and the sample used to develop and test the new scale. First, a scale is developed following an elaborate item generation phase. Second, the preliminary measurement instrument is tested in a survey. The new scale's psychometric properties are assessed using exploratory and confirmatory factor analysis, followed by an assessment of its construct validity. Then, the resulting measurement instrument is tested and cross-validated in a replication study. We conclude with a discussion of the research findings and the applicability of the new scale for the study of bureaucratic encounters.

2.2 WHY BUREAUCRATS' ATTITUDE TOWARDS CLIENTS MATTERS

During their interactions with clients, street-level bureaucrats exercise a vast amount of discretion (Hupe & Hill, 2016). Decisions made in this discretionary space can greatly impact citizens' lives (e.g., Lipsky, 1980; Maynard-Moody & Musheno, 2003). Consequently, how bureaucrats exercise their discretion has always been carefully scrutinized (e.g., Goodsell, 1976; Prottas, 1979). Bureaucrats' use of discretionary space is based on their categorization of clients (Lipsky, 1980; Jilke & Tummers, 2018), such as based on clients' 'worthiness' or 'deservingness' (Maynard-Moody & Musheno, 2003), whether they are deemed 'desirable' or 'undesirable' (Hasenfeld & Steinmetz, 1981) or their perceived need (Kroeger, 1975). These categorizations subsequently inform the courses of action bureaucrats take (e.g., Maynard-Moody & Musheno, 2003). For instance, Lineberry's (1977) 'underclass hypothesis' posits that "clients who are perceived by bureaucrats as less worthy or desirable (e.g., the minorities and poor) are discriminated against in delivery manner and output" (Goodsell, 1981b, p. 763). Bureaucrats can, however, differ in their classifications and reactions thereto. For

example, clients' perceived need can trigger bureaucrats to put in extra effort to benefit a client's case (Kroeger, 1975; Maynard-Moody & Musheno, 2003), but can also trigger them to 'cream' clients, meaning that they are inclined to favor clients whose cases are most likely to be brought to a successful outcome anyway instead of those who are in most need (Jilke & Tummers, 2018; Lipsky, 1980).

How bureaucrats react to cues like perceived need depends on how they assess and interpret information on clients (e.g., Oberfield, 2014a; Raaphorst & Van de Walle, 2017). These interpretations are informed by their general attitude towards clients (e.g., Eagly & Chaiken, 1993). This general attitude is an abstract-level prototype that guides bureaucrats' assessments and categorizations of clients, as well as their subsequent courses of action (Baviskar & Winter, 2017; Oberfield, 2014a). Bureaucrats do not always have very positive views of their clients (Van de Walle &Lahat, 2016), and early studies on bureaucratic encounters have already found that bureaucrats with different general orientations towards clients react differently to the cues they receive from them. For instance, Kroeger (1975, p. 191) showed that client-oriented bureaucrats use greater discretion to benefit clients and were more prone to manipulation by clients. Stone (1981, p. 45) argues that bureau-crats with "a condemnatory moralistic view of clients", that is, a negative attitude towards clients, are more inclined to respond punitively to clients. And Billingsley (1964, p. 406) found that client-oriented bureaucrats are more willing to disregard policy in order to help clients. Despite this early evidence, scholars have mainly studied how bureaucrats categorize clients during decisions, but have looked less into the general attitudes bureaucrats hold towards their clients. Furthermore, current work does not generally distinguish between attitudes towards clients and other related attitudes of bureaucrats. This article attempts to unravel the general attitude of bureaucrats towards clients into its different dimensions and give substantive body to this concept in a street-level context.

2.3 WHAT ARE ATTITUDES? THE MULTICOMPONENT ATTITUDE MODEL

To elaborate a scale for measuring street-level bureaucrats' attitude towards clients, we rely on insights from social psychology which treats attitudes as a summary evaluation. Social psychology has a long tradition of studying attitudes. In this field, the attitude concept is commonly understood as an object's evaluation on a continuum (Ajzen, 2001; Eagly & Chaiken, 1993; Fabrigar, MacDonald, & Wegener, 2005; Katz, 1960). This evaluation is often believed to incorporate the cognitive, affective and behavioral pieces of information that are called into one's memory upon confrontation with an attitude object (Haddock & Huskinson, 2004; Maio & Haddock, 2009; Olson & Kendrick, 2008). Incorporating these

pieces of information is crucial to come to a full understanding of attitudes (Fabrigar et al., 2005; Fazio & Olson, 2003).

A definition of street-level bureaucrats' attitude towards clients should thus contain three elements: bureaucrats' general evaluation of clients, an index on which to place this evaluation (e.g., a bipolar continuum) and the pieces of information that constitute this evaluation. To this end, we fuse the attitude definitions by Maio & Haddock (2009) and Petty et al. (1997) and define street-level bureaucrats' attitude towards clients as: *their summary evaluation of clients along a dimension ranging from positive to negative that is based on the bureaucrats' cognitive, affective and behavioral information on clients.*

Such a detailed definition of street-level bureaucrats' attitudes is rare in studies of the bureaucratic encounter, where the attitude concept is frequently characterized by conceptual ambiguity: scholars rarely make explicit what they understand by 'attitude' (e.g., Baviskar & Winter, 2017; Liou & Cruise, 1994; Scherer & Scherer, 1980; Soydan, 1995; Stone, 1981; Stone & Feldbaum, 1976) or they conceptualize it by referring to related concepts, such as 'values' (e.g., Winter, 2002). They also tend to use different concepts to measure similar phenomena (i.e., views of the other party in the bureaucratic encounter), suggesting conceptual overlap and the interchangeable use of concepts. For instance, authors refer to 'orientations' (Blau, 1960), 'assessments' (Nelson, 1981), 'perceptions' (Berman, 1997; Borgatta et al., 1960), 'views' (Oberfield, 2012) and 'attitudes' (Baviskar & Winter, 2017; Oberfield, 2012; Stone, 1981; Wilson, 1989). Some exceptions among bureaucracy scholars are Wilson (1989) and Oberfield (2012), who both refer to 'evaluation' in their attitude conceptualizations, which closely links to attitude conceptualizations in the field of psychology. Furthermore, this attitude has generally been considered to be a one-dimensional construct, mostly tapping into the cognitive dimension of bureaucrats' assessment of clients (e.g., see Borgatta et al., 1960; Scherer & Scherer, 1980; Winter, 2002).

Pieces of information belonging to the same type are commonly referred to as *attitude components* (Fazio & Olson, 2003; Maio & Haddock, 2009). Smith (1947) was the first to formally distinguish among three attitude aspects in his use of the attitude concept (Breckler, 1984, p. 1192). The multicomponent model of attitude took flight in the 1960s and has since been recurrent in attitude theory (Breckler, 1984, p. 1192; e.g., see Eagly & Chaiken, 1993; Haddock & Huskinson, 2004; Maio &Haddock, 2009; Olson & Kendrick, 2008; Rosenberg & Hovland, 1960). In this study, we use the multicomponent model as depicted by Breckler (1984, p. 1196). In this model, there are three attitude components: the cognitive, the affective and the behavioral one. The *cognitive attitude component* refers to the beliefs and attributes that bureaucrats associate with clients (see Fabrigar et al., 2005; Fazio & Olson, 2003). For instance, juvenile probation officers might perceive their clients as disrespecting order and authority. That belief links clients to the attribute 'disrespect'. Subsequently, this piece of information contributes to a negative assessment of clients.

The *affective attitude component* refers to bureaucrats' emotional responses evoked in the bureaucrat by confrontations with the attitude object 'clients' (see Breckler, 1984; Haddock & Huskinson, 2004). This confrontation thus acts as a stimulus that brings the feelings and emotions bureaucrats associate with clients to the forefront (see Maio & Haddock, 2009). This affect subsequently informs their general evaluation of clients. An example is provided by Blau (1960) who found that new employees of a public welfare agency experience a reality shock when they realize that clients are dishonest. This 'shock' can make the bureaucrat feel threatened. These feelings subsequently pressure the bureaucrat to change his attitude towards clients.

The *behavioral attitude component* refers to past behavioral responses to clients (see Haddock & Huskinson, 2004; Maio & Haddock, 2009) and is best explained by Bem's (1972) self-perception theory. Self-perception theory posits that individuals infer their attitudes from their past behaviors, as long as they were not impelled to undertake those behaviors (Bem, 1972). Following this theory, bureaucrats are argued to (subconsciously) reflect on their chosen past behaviors towards clients—which exclude behaviors prescribed by bureaucratic rules and regulations—and infer their attitude towards clients from these observations (Bem, 1972). If these behaviors have been mostly negative, they can inform a bureaucrat of his negative client-attitude (see Eagly & Chaiken, 1993). Thus, a bureaucrat who observes that he is always rude to clients might acquire that he must have a negative attitude towards them from this behavioral piece of information.

For conceptual clarity, it is necessary to distinguish between behavior as an attitude component and as a consequence of attitude. In PA research, attitudes have mostly been explored as antecedents of behaviors relevant to bureaucratic decision-making (Oberfield, 2012), rather than as an attitude component. Behavior can both function as an attitudinal indicator and as an attitudinal consequence (Eagly & Chaiken, 1993; Fazio & Olson, 2003). Attitudes itself are unobservable mental constructs (Fazio & Olson, 2003). They become observable by means of the cognitive, affective and behavioral pieces of information linked to the attitude object in the mind of the individual (Himmelfarb, 1993). All information types can subsequently work to elicit behavioral responses (Eagly & Chaiken, 1993). Even though behavior can both serve as an attitudinal indicator and consequence thereof, behaviors can only be an attitudinal indicator if they "relate to the dimension of favorability-unfavorability towards the attitude object" (Himmelfarb, 1993, p. 63) and thus harbor evaluative properties.

2.4 METHODS

Measurement of scale development

The efforts to develop a measurement instrument for bureaucrats' general attitude towards clients were based on the scale development guidelines of DeVellis (2003). We first generated a pool of items for each of the three components. The item-generation procedure consisted of multiple steps: a literature study; a series of qualitative interviews; multiple cognitive interviews with street-level bureaucrats and a brainstorm session with academic experts. Subsequently, this item pool was subjected to a test among a sample of bureaucrats to establish whether the three attitude components existed using EFA, followed by CFA. This resulted in the inclusion or exclusion of individual items. Finally, the construct validity of the therewith obtained scale was further assessed: after a test of convergent validity, discriminant validity was tested by exploring the relations within the new attitude construct (bureaucrats' general attitude towards clients) and between this new construct and three related constructs (prosocial motivation, work engagement and bureaucrats' rule-following identities). To further validate the resulting measurement instrument, we conducted a replication study in which discriminant validity was tested using bureaucrats' rule-following identities and self-efficacy.

Case selection and test sample

For our scale development purposes, we focused on one single type of bureaucrat. Tax bureaucrats are a classic example of street-level bureaucrats (see Finer, 1931). They work in a classic bureaucratic framework yet have ample discretionary space that enables attitudes to affect their work (Aberbach & Christensen, 2007). The main advantage of using the tax administration as a case is that all citizens are its potential clients. This differentiates it from many other street-level bureaucracies whose bureaucrats tend to deal with very specific types of clients, with very specific characteristics, such as poor people or criminals (e.g., Dubois, 2010). The latter could bias the development of a measurement instrument. Another advantage is the rather similar design and tasks of tax administrations across (capitalist) countries (see Campbell, 2005), which facilitates future cross-case comparison.

We develop and test the multicomponent model in the Dutch tax administration, among bureaucrats who have face-to-face interactions with clients. Their task is to audit tax filings of entrepreneurs in the small and medium-sized enterprises segment (1–50 employees). They typically meet with small entrepreneurs to discuss tax declarations, ask for clarifications and make decisions about deductions or fines, often in confined settings. They experience a range of client actions and reactions ranging from joy, over fear and suspicion, to threats and violence.

The final item pool was tested in a cross-sectional survey of Dutch tax bureaucrats. We selected one of the five Dutch tax regions. All bureaucrats in this tax region who fit our

research population were invited to participate ($n = 433$). Data were collected using an email survey with two reminders, sent to the bureaucrats' work address in the summer of 2015. Seven-point Likert scales ranging from 'never' to 'always' were used. The response rate was 67.4% ($n = 292$). We included only those respondents who confirmed that they actually had contact with citizens and had valid replies on the attitude items, resulting in a sample of 218 valid observations. The respondents' mean age was 52.4, which is representative of the Dutch tax administration's workforce (Belastingdienst, 2012). Of the respondents, 19.7% were female and 75.2% were male.

2.5 ITEM GENERATION

Item generation for the scale was based on the literature on bureaucratic encounters and in-depth interviews with tax bureaucrats. This literature offers rich descriptions of official-client interactions (e.g., Dubois, 2010; Maynard-Moody & Musheno, 2003; Prottas, 1979). In these studies, scholars, for example, describe critical incidents between bureaucrats and clients (Savaya, Gardner, & Stange, 2011) or show how bureaucrats cope with clients (Lipsky, 1980). These accounts hold crucial pieces of cognitive, affective and behavioral information relevant to the bureaucratic encounter. For our literature search, we used Google Scholar so both relevant articles and books could be selected. To collect literature, we searched for (combinations of) the following terms: 'bureaucratic encounter', 'civil servant', 'official', 'public official', 'public servant', 'front-line official', 'street-level bureaucrat', 'administrator', 'attitude', 'orientation', 'perception', 'evaluation', 'interaction', 'citizen', 'citizen-client', 'client' and 'the public'. A study was deemed relevant if it described cognitive, affective and/or behavioral elements in the encounter between officials and the public. Studies mostly stemmed from the field of PA, but also studies from the field of psychology were included since the multicomponent model stems from this discipline. Whenever a publication was selected, we checked the publications citing that piece for relevance. Search results were checked until saturation occurred.

To increase the validity of the item generation, we validated and complemented the pieces of information from the literature study with those from 15 interviews with tax bureaucrats who work in the SME-segment. In these interviews, bureaucrats were first administered an open-ended measure of the three components—as illustrated by Maio and Haddock (2009): they were asked to name all characteristics, feelings and chosen behaviors that they associate with clients in general. Second, respondents were invited to tell the story of one memorable negative and positive experience with clients. Follow-up questions targeted the clients' characteristics in those interactions, emotions that they evoked in the bureaucrat, and the courses of action that the respondent took. The interviews were transcribed afterwards.

Pieces of information were coded on whether they stemmed from literature from the PA field or psychology (literature study only); which party in the bureaucratic encounter was evaluated (i.e., the bureaucrat or the client); and the component a piece of information represented. Items were selected based on their recurrence in the literature and the interviews. Each resulting preliminary item pool was refined by means of feedback from three cognitive interviews with street-level tax bureaucrats and a brainstorm session with a group of academics who specialize in survey research and/or street-level bureaucracy. The cognitive interviews inquired about the bureaucrats' understanding of the items and their applicability to their work context.

These efforts resulted in three initial item pools that contained 15 items for each attitude component (see Appendix 1). Since the aim of this study is to develop a measurement instrument with potential applicability across different types of street-level bureaucrats, all items were constructed as templates. This means that the attitude object the items refer to can be adapted from 'clients' to a term befitting the type of street-level bureaucrat being surveyed (see Tummers, 2012). In the current case, using templates meant that the term 'clients' was replaced by 'taxpayers'. The pieces of information central to each component are described below.

The cognitive attitude component

Central pieces of information to the cognitive component are clients' honesty and cooperativeness, how clients approach the street-level bureaucrat, and clients' knowledgeability. Clients' honesty and cooperation are critical for successful bureaucratic encounters. To decide on the allocation of government benefits and sanctions, bureaucrats often depend on input from clients (Lipsky, 1980); for instance, determining whether clients are eligible for welfare requires that clients honestly report their income. Such (monetary) incentives can cause clients to seek to influence the bureaucrat to alter the outcomes of the bureaucratic encounter to their advantage, for example, by offering bribes or emphasizing their destitution (Hasenfeld & Steinmetz, 1981; Katz & Danet, 1973b). Such tactics can force bureaucrats to put more effort into these cases and are hence more burdensome for them (see Wilson, 1989).

The second recurring element is the way clients approach bureaucrats. Many bureaucratic encounters are non-voluntary (Lipsky, 1980). Citizens do not want to get arrested by the police. They interact with social welfare bureaus because they depend on their services for their livelihood, and they pay taxes because the government obliges them to do so (see Lipsky, 1980). This non-voluntary nature can cause clients to approach bureaucrats with resistance and hostility (e.g., Dubois, 2010; Savaya et al., 2011; Worden, 1989). On the other hand, accounts of bureaucratic encounters also link clients to, for instance, gratitude for the benefits that they receive or the efforts that bureaucrats invest in their cases (Blau, 1960; Lipsky, 1980), and being sympathetic (Finer, 1931), or mild-mannered (Borgatta et

al., 1960). The latter types of traits are frequently recurring in descriptions of how clients should behave (e.g., Finer, 1931).

The last piece of information we identified encompasses clients' knowledgeability. Knowledge is often a requirement for successful bureaucratic encounters; clients applying for welfare must know where to apply, which documents to bring, and how to fill in certain forms (see Dubois, 2010; Gordon, 1975; Lipsky, 1980). Insufficient knowledge can cause clients to give improbable explanations for their situation or make inappropriate appeals to bureaucrats (Katz & Danet, 1973b). Clients can also use their knowledgeability to manipulate the bureaucratic encounter to their advantage (Hasenfeld & Steinmetz, 1981; Prottas, 1979, p. 40).

The affective attitude component

Because affect has been the focus of much prior research, we used the well-established Positive and Negative Affect Schedule (PANAS) scales by Watson, Clark, and Tellegen (1988) as the foundation for this attitude component. These scales consist of "two 10-item mood scales that comprise the PANAS" (Watson et al., 1988, p. 1063). These scales, however, do not encompass the object inducing those moods, whereas we focus on moods specifically evoked by clients. To ensure the content validity of the affective attitude component, we revised the PANAS scales in three steps, based on the literature study, substantive interviews and cognitive interviews.

First, we identified PANAS items that were not suitable for the street-level work context. To identify these poorly fitting items, we administered the Dutch translation of the PANAS scales (Peeters, Ponds, & Vermeeren, 1996) to the respondents of the cognitive interviews. Items had a poor fit because their meaning was ambiguous to them (e.g., 'proud'), they did not fit their work context (e.g., 'guilty'), they showed content overlap (e.g., 'scared' and 'afraid') or they were not interpreted as affect (e.g., 'strong').

Second, to decide which items to discard, we combined these insights from the cognitive interviews with insights from previous research on the PANAS scales: Crawford and Henry (2004) found covariation among the PANAS items, suggesting item redundancy in the PANAS scales (Thompson, 2007). To not jeopardize the validity of the PANAS foundation of the affective attitude component, we ensured that each covarying item group was represented in the final PANAS item selection. This resulted in the selection of the PA items of alert, active, determined and inspired and the NA items of upset, afraid, irritable, ashamed and nervous.

Third, having refined the PANAS scales for our research purposes, we then complemented the remaining PANAS items with observations from the literature study and the substantive interviews. One piece of affective information complementing the PANAS items is a feeling of threat (see Blau, 1960; Savaya et al., 2011). This threat can be either physical or psychological. A feeling of physical threat may result from client aggression; hostility; anger and actual,

attempted, or threatened physical violence by clients (Savaya et al., 2011, p. 65). Psychological threats result from stimuli that affect bureaucrats' mental well-being, such as feeling inexperienced or unable (Blau, 1960; Lipsky 1980); guilt or embarrassment (Prottas, 1979, p. 107); or frustrated, unappreciated or uncertain (Savaya et al., 2011). As a form of protection against these threats, bureaucrats might emotionally distance themselves from their clientele and hence become emotionally detached from them (Blau, 1960; Dubois, 2010).

Another piece of affective information is a desire to help that characterizes many street-level bureaucrats upon entering the public sector (Blau, 1960). Bureaucrats have a tendency to show concern for their clients (Stone & Feldbaum, 1976). Consequently, client-encounters can "provide workers with their greatest sense of accomplishment" (Maynard-Moody & Musheno, 2003, p. 21).

The behavioral attitude component

At the core of the behavioral attitude component are the past behaviors towards clients that street-level bureaucrats engaged in by choice (see Bem, 1972). Pieces of information central to this component are the 'fair' treatment of clients, the way bureaucrats approach clients and bureaucrats' helpfulness.

Despite its recurrence (e.g., Almond & Verba, 1973; Lipsky, 1980), an unequivocal understanding of what is meant by 'fair treatment' is yet to emerge. Bureaucracies are traditionally designed to treat clients equally (Gastelaars, 2006); however, many emphasize the need for a personal approach and responsive services (e.g., Almond & Verba, 1973; Brenninkmeijer, 2013). Lipsky (1980, p. 22) even refers to "the apparent unfairness of treating people alike." Perceived unfairness can motivate bureaucrats to manipulate rules for clients (e.g., Maynard-Moody & Musheno, 2003).

The second piece of behavioral information concerns bureaucrats' approach to clients. Here, accounts of how bureaucrats should behave are prominent (e.g., Brenninkmeijer, 2013; Gastelaars, 2006). Scholars argue that bureaucrats should be courteous (Nelson, 1981), polite (Finer, 1931), civil (Katz & Danet, 1973a) and respectful (Brenninkmeijer, 2013). Scholars also regularly criticize bureaucrats' behavior, accusing them of the inhumane, abusive, unreflective and apathetic treatment of clients (Bartels, 2013, p. 473; Goodsell, 1981b, p. 763).

The third piece of behavioral information relates to bureaucrats' helpfulness (e.g., Blau, 1960; Finer, 1931; Wilson, 1989). Many behaviors can be categorized as such, like being accessible and responsive; disseminating information; showing concern and empathy (King and Stivers, 1998, p. 66); being deliberative (Bartels, 2013); listening to clients attentively and considerately (Almond & Verba, 1973, p. 43) and "making application easy, volunteering unsolicited information, providing prompt responses" (Hasenfeld & Steinmetz, 1981, p. 88). On the opposite end of this spectrum, one can find behaviors that obstruct clients' cases, such as "making application difficult, withholding information, evading questions, and applying stigmatizing labels" (Hasenfeld & Steinmetz, 1981, p. 88).

2.6 TESTING AND VALIDATING THE MULTICOMPONENT MODEL

We now test the multicomponent model by means of EFA and CFA and assess the construct validity of this measure.

Analyzing dimensionality: EFA

Principal axis factoring with oblique rotation was used as the extraction method (direct oblimin). Conditions for arriving at a stable, reliable factor solution were present. Because all items were designed to measure one specific component, groups of items were expected to cluster together (see DeVellis, 2003). Upon inspecting the correlation matrix, we omitted six items (see Appendix 1) that had no or almost no correlations above the .3 threshold (Field, 2013).

We used the scree test and the theoretical interpretability of factors to determine the number of factors to retain. This resulted in a four-factor solution that diverged from the hypothesized three-factor solution by its distinction between the positive and negative affective attitude items. This finding is consistent with prior research: the PANAS scales in which the affective component was grounded were found to represent two orthogonal (i.e., distinctive) dimensions of affect (Watson et al., 1988), suggesting that affective items are unipolar in nature. This means that, if affective items are placed on a measurement continuum, its endpoints do not pertain to each other's opposite—as they would for a bipolar item—but to "different degrees of the presence of the same attribute" (Schwarz, 2008, p. 43). Thus, if we placed 'clients make me feel active' on a continuum, the endpoints would represent 'very active' and 'not active at all' (unipolar) instead of 'very active' and 'passive' (bipolar).

To determine the most optimal, parsimonious four-factor solution, we deleted 22 additional items (see Appendix 1) on multiple grounds: they had factor loadings under .4 (Field, 2013), had cross-loadings greater than .3, had low communalities (close to .3 or under) and/or were judged to lack clarity from a theoretical point of view. As a result, we obtained a four-factor model that explained 59.5% of the variance in the remaining 17 items. The Cronbach's alphas were respectable for all dimensions (see DeVellis, 2003), ranging from .70 to .83. No items could be removed to increase the reliability of the subscales, thus showing no sign of item redundancy. Table 2.1 shows the EFA results.

The correlations between the four attitude subscales are listed in Table 2.2. They show that the components are sufficiently distinctive (e.g., no multicollinearity). The positive and negative affective components are uncorrelated to each other, affirming that the affective items are indeed unipolar. A similar distinction is found between the cognitive and positive affective items.

Table 2.1. Summary of EFA results for the final item pool

	Rotated Factor Loadings							
	Test study (*n* = 218)				Replication study (*n* = 879)			
Item	F1	F2	F3	F4	F1	F2	F3	F4
Cognitive attitude component								
Taxpayers are manipulative	-.720				.628			
Taxpayers are hostile	-.547				.644			
Taxpayers are unpredictable	-.575				.570			
Taxpayers are stubborn	-.508				.487			
Taxpayers are dishonest	-.703				.593			
Positive affective attitude component								
Taxpayers make me feel alert		.681				.645		
Taxpayers make me feel inspired		.674				.682		
Taxpayers make me feel determined		.559				.687		
Taxpayers make me feel active		.809				.808		
Negative affective attitude component								
Taxpayers make me feel upset			.585				.530	
Taxpayers make me feel afraid			.514				.570	
Taxpayers make me feel nervous			.765				.679	
Taxpayers make me feel insecure			.705				.646	
Taxpayers make me feel uncomfortable			.814				.830	
Behavioral attitude component								
I explain things to taxpayers				.662				.758
I make taxpayers feel at ease				.710				.739
I help taxpayers				.610				.739
Eigenvalues	4.35	2.66	1.67	1.43	3.65	2.90	1.78	1.58
% of variance	25.59	15.64	9.83	8.44	21.46	17.06	10.49	9.27
α	.77	.77	.83	.70	.73	.80	.79	.79
M	2.86	3.98	1.78	5.35	2.79	4.17	1.81	5.06
SD	.61	.92	.48	.89	.59	.93	.48	1.01

EFA, exploratory factor analysis; SD, standard deviation.

Confirmatory factor analysis

To test the hypothesized structure that resulted from the EFA, we conducted a CFA using AMOS version 24. In the CFA, all the items from one hypothesized component are constrained to freely load on that component but to have zero loadings on the other attitude components (Byrne, 2010). The fit indices of the multicomponent model are: χ^2 = 217.18 (df = 113, *p* =.000), the ratio of the minimum discrepancy to degrees of freedom (CMIN/DF) = 1.92, the goodness-of-fit index (GFI) = .90, the comparative fit index (CFI) = .91, the root mean square error of approximation (RMSEA) = .065. The χ^2 is significant, which

means there is no perfect fit between the hypothesized model and the data. However, the χ^2 assumption of perfect fit has limitations that are offset by other goodness-of-fit indices, such as the CMIN/DF (Byrne, 2010). CMIN/DF values between 1 and 3 represent a good model fit (Vermeeren, Kuipers, & Steijn, 2011). The GFI and CFI indicate a good fit if their value is close to 1 (Byrne, 2010). A generally used cut-off point in the social sciences is .9 (Vermeeren et al., 2011). These fit parameters thus support the hypothesized fit of the multicomponent model. The RMSEA value is indicative of reasonable model fit (Byrne, 2010).

Table 2.2. Intercorrelations between the attitude subscales

		Cog.	**PosAff.**	**NegAff.**	**Behav.**
Test study (*n* = 218)					
Cognitive att. comp.	Pearson Correlation	1			
	Sig. (2-tailed)				
Positive affective att. comp.	Pearson Correlation	.085	1		
	Sig. (2-tailed)	.212			
Negative affective att. comp.	Pearson Correlation	.408	-.011	1	
	Sig. (2-tailed)	.000	.875		
Behavioral att. comp.	Pearson Correlation	-.216	.194	-.306	1
	Sig. (2-tailed)	.001	.004	.000	
Replication study (*n* = 879)					
Cognitive att. comp.	Pearson Correlation	1			
	Sig. (2-tailed)				
Positive affective att. comp.	Pearson Correlation	.107	1		
	Sig. (2-tailed)	.002			
Negative affective att. comp.	Pearson Correlation	.324	-.068	1	
	Sig. (2-tailed)	.000	.043		
Behavioral att. comp.	Pearson Correlation	-.112	.229	-.165	1
	Sig. (2-tailed)	.001	.000	.000	

Construct validity tests

To further validate the new scale, we assessed the multicomponent model's construct validity by means of multiple tests that evaluate its convergent validity, discriminant validity and reliability.

Convergent validity

To assess the convergent validity of the multicomponent model, we first examined how well the items measure the component that they are hypothesized to represent. Significant construct loadings are indicative of convergent validity (Vermeeren, Kuipers, & Steijn, 2014). The construct loadings range from .565 to .795 (Table 2.3) and are all significant at the *p* = 0.001 level, supporting the claim of convergent validity.

As a second assessment of convergent validity, we calculated the average variance extracted (AVE) for each component. The AVE is calculated by averaging the squared correlation estimates of the individual items that represent the same component (Farrell and Rudd, 2009). AVEs over .5 indicate that the amount of variance captured by a component is greater than the amount of variance due to measurement error (Fornell & Larcker, 1981, p. 46). The AVEs are .39 (cognitive component), .47 (positive affective component), .50 (negative affective component) and .45 (behavioral component), somewhat under the .5 threshold. The AVEs are listed in Appendix 2.

Table 2.3. Means, standard deviations (SD), standardized estimates, R^2

	Test study ($n = 218$)				Replication study ($n = 879$)			
	M	SD	Stand. Regression Weights (S.E.)	R^2	M	SD	Stand. Regression Weights (S.E.)	R^2
Cognitive component								
…are manipulative	2.93	.88	.645 (.060)	.416	2.88	.88	.616 (.031)	.379
…are hostile	2.23	.79	.594 (.055)	.353	2.14	.71	.644 (.025)	.412
…are unpredictable	2.92	.90	.620 (.062)	.384	2.81	.87	.504 (.032)	.254
…are stubborn	3.15	.84	.584 (.059)	.341	2.96	.88	.628 (.031)	.394
…are dishonest	3.08	.81	.705 (.054)	.497	3.16	.89	.589 (.032)	.347
Positive affective comp.								
…make me feel alert	4.19	1.16	.690 (.078)	.476	4.38	1.15	.684 (.038)	.468
…make me feel inspired	3.62	1.18	.668 (.080)	.447	3.78	1.24	.504 (.041)	.254
…make me feel determined	3.97	1.16	.565 (.081)	.319	4.21	1.13	.647 (.037)	.419
…make me feel active	4.16	1.26	.795 (.083)	.632	4.31	1.19	.798 (.038)	.637
Negative affective comp.								
…make me feel upset	1.59	.55	.681 (.035)	.464	1.62	.64	.565 (.022)	.319
…make me feel afraid	1.55	.58	.655 (.037)	.430	1.50	.57	.613 (.347)	.376
…make me feel nervous	1.90	.67	.761 (.042)	.579	1.89	.65	.750 (.021)	.563
…make me feel insecure	1.86	.62	.690 (.040)	.477	1.97	.68	.659 (.023)	.434
…make me feel uncomfortable	2.00	.66	.741 (.041)	.548	2.09	.68	.713 (.022)	.508
Behavioral component								
I explain things to taxpayers	5.97	.99	.743 (.073)	.551	5.62	1.07	.766 (.036)	.587
I make taxpayers feel at ease	5.24	1.14	.690 (.084)	.476	5.05	1.23	.749 (.041)	.561
I help taxpayers	4.86	1.22	.571 (.090)	.326	4.50	1.30	.723 (.043)	.523

Divergent validity

Discriminant validity is achieved when items representing the same component are more strongly correlated with one another than with items measuring either the other components or different constructs (Farrell & Rudd, 2009). This assumption was first tested by comparing the χ^2 of the unconstrained measurement model with the χ^2 of the constrained

model, in which the covariance between the four components is constrained to equal 1. The constrained model equals a one-factor representation of the multicomponent model. If the unconstrained model has a significantly lower χ^2 than the constrained model, discriminant validity is established (Bagozzi & Philips, 1982). The χ^2 of the constrained model is 670.41 (df = 119).[4] The resulting $\Delta\chi^2(6)$ is 453.23, which is significant at $p < .001$ and supportive of the discriminant validity of the multicomponent model.

For the second test of discriminant validity, we again used the AVEs. If the AVEs for any two constructs are both larger than the variance that these two constructs share discriminant validity is established (Farrell & Rudd, 2009, p. 5). For this test, we assessed the relations *within* the attitude construct (i.e., between the attitude components only) and *between* the four attitude components and three similar constructs from the social sciences: prosocial motivation, work engagement and the bureaucrats' rule-following identities, all measured in the same survey.

Prosocial motivation is defined as "the desire to benefit other people" (Grant, Parker, & Collins, 2009, p. 48). High prosocial motivation has previously been linked to positive perceptions of others in the work environment (Grant et al., 2009). We therefore classify high prosocial motivation as a positive attitude. To measure prosocial motivation, we used the four-item scale developed by Grant et al. (2009). In our study, this instrument exhibited high reliability ($\alpha = .87$).

Schaufeli, Bakker, and Salanova (2006, p. 701) define work engagement as "a positive, fulfilling work-related state of mind that is characterized by vigor, dedication, and absorption." We perceive high work engagement as a positive attitude and measured this construct using the one-factor, nine-item short version of the Utrecht Work Engagement Scale developed by Schaufeli et al. (2006). In our study, the α for this instrument was .95.

Reliability

Finally, we turn to the reliability. The R^2 in Table 2.3 indicates "how consistently the observed variable measures the latent dimension" (Perry, 1996, p. 14). R^2 values range between 32.6% and 63.2%. This finding shows that although there is some variability in the reliabilities, the components generally explain an adequate portion of the variance in their items (see Perry, 1996).

---

4   The other fit parameters of the constrained model are: CMIN/DF = 5.63, GFI = .69, CFI = .50, RMSEA = .146.

2.7 REPLICATION STUDY

A common practice in scale development is to split samples for scale validation purposes (DeVellis, 2003). This practice is problematic for multiple reasons: first, "two subsamples are likely to be much more similar than two totally different samples" (DeVellis, 2003, p. 99); second, performing EFA and CFA on two subsamples does not allow the cancelling out of effects the omitted items had on respondents' replies to the final scale items (DeVellis, 2003). For these reasons, we conducted a replication study that followed the same steps as the test study. The replication sample and the results of the cross-validation efforts are discussed below.

Replication sample

The test survey was administered in one of the five Dutch tax regions. The replication survey was conducted in the other four tax regions, in the summer of 2016. The selection- and sample procedures were the same for both studies. The sample for the replication study contained 2257 street-level tax bureaucrats. The response rate was 55.2% (n = 1245). Again, we only included tax bureaucrats who confirmed to be tax auditors with client-contact and valid replies on the attitude items. This resulted in a final sample of 879 street-level tax bureaucrats. The respondents' mean age was 53.7 and the gender division was 78.8% male versus 21.2% female.

Exploratory- and confirmatory factor analysis

An EFA identified four factors with an Eigen Value over 1, together explaining 58.3% of the variance. Table 2.1 shows that no cross-loadings above .3 were found and the factor loadings indicate that all items contribute substantially to their respective component (Field, 2013). The Cronbach's alphas varied between .73 and .80, all well within the respectable range, and showed no signs of item redundancy. The correlations support the distinctness of the components (see Table 2.2). However, they do imply that a distinction based on the orthogonality of the components is less prominent in this sample.

    The CFA supported the hypothesized fit of the model. The fit indices are: χ^2 = 394.32 (df = 113, p = .000), CMIN/DF = 3.49, GFI = .95, CFI = .93, RMSEA = .053. Only the CMIN/DF is somewhat above the 3.0 threshold. However, since this parameter is affected by sample size authors frequently suggest that values below 5.0 can be deemed reasonable (Marsh & Hocevar, 1985).

Construct validity tests

Significant construct loadings and AVEs over .5 are indicative of convergent validity. Table 2.3 shows that all construct loadings are significant at the p = 0.001 level, indicative of the multicomponent model's convergent validity. The AVEs are .36 (cognitive component),

.51 (positive affective component), .44 (negative affective component) and .56 (behavioral component), similar to those of the test study.

Discriminant validity is assessed through a comparison of the unconstrained measurement model with the constrained model, an AVE-analysis of the within-construct and between-construct shared variance, and the reliability of the measured items. The χ^2 of the unconstrained model was 394.32 (df=113). The χ^2 of the constrained model is 2749.69 (df=119). The resulting $\Delta\chi^2(6)$ of 2355.37 is significant at $p < .001$ and affirms the model's discriminant validity.[5]

For the AVE-analysis of discriminant validity, we use one construct from the test survey, that is, workers' rule-following identities[6] ($\alpha = .78$), and introduce one new construct from the social sciences: bureaucrats' self-efficacy. Self-efficacy refers to the "belief in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). We view self-efficacy as a positive attitude towards the self and assess this construct using Chen, Gully, and Eden's (2001) eight-item new general self-efficacy scale ($\alpha = .88$).

Appendix 2 shows that the AVE-analysis of both the within-construct and between-construct shared variances endorse the discriminant validity of the multicomponent model: the AVEs for the four components were substantially higher than the variance shared by either any combination of the attitude components or by the components and bureaucrats' rule-following identities and self-efficacy.

Lastly, the reliabilities of the measured items vary between 25.4% and 63.7% (see Table 2.3). Although some variability in the reliabilities can be observed between the test study and the replication study, the general tendencies of the replication study's reliabilities are similar to those of the test study.

2.8 DISCUSSION AND CONCLUSION

Based on the psychological multicomponent model of attitude, we developed a measurement instrument for street-level bureaucrats' attitude towards clients that was tested and validated in two surveys among Dutch street-level tax bureaucrats. This study revealed four different attitude components—rather than the expected three—measured by seventeen items: the cognitive attitude component, the positive affective attitude component, the negative affective attitude component and the behavioral attitude component. This measurement instrument demonstrated factorial and construct validity.

---

5 The other fit parameters of this constrained model are: CMIN/DF = 23.10, GFI = .67, CFI = .38, RMSEA = .159.

6 EFA and reliability analysis again showed that the item without any reference to rules needed to be removed.

In addition to a substantive distinction according to the attitude components, we found that a distinction following the positive and negative connotation of items is inherent to the multicomponent model: the positively and negatively worded affective items were shown to form two orthogonal dimensions of affect, rather than loading on a single affective dimension. The cognitive and behavioral attitude components are also characterized by this positive–negative distinction, although in a different manner: like the affective attitude components, the cognitive component and behavioral attitude component only hold *either* positively worded *or* negatively worded items. However, in contrast to the affective items, the positively and negatively worded cognitive and behavioral items did not form separate dimensions of cognition and behavior in the test study. An explanation therefore can be found in the unipolar versus bipolar nature of the items: where the affective items are *unipolar*, the cognitive and behavioral items are *bipolar* in nature. Their bipolarity allows these items to grasp the relevant spectrum of information, regardless of their unidirectional wording. For example, whereas the endpoints of a measurement continuum for the positive affective item 'clients make me feel active' are 'very active' and 'not active at all' (unipolar), the endpoints are 'very honest' and 'very dishonest' (bipolar) for the cognitive item 'clients are dishonest'. Thus, one cannot characterize a bureaucrat who does not feel active as feeling passive, whereas a client who is not honest can be classified as being dishonest.

At first sight, the orthogonal distinction between the positive and negative affective items, and the positive affective items and the cognitive items seemingly moves to the background in the replication study: in this study, we do find significant correlations between these components. An explanation therefore can be found in the contextual factors of both studies: the test study had a lower data to respondent ratio than the replication study. A small data to respondent ratio combined with a small sample size can cause distortions to correlation values (DeVellis, 2003). Since EFA and CFA tests of factorial validity in the replication study support the orthogonal nature of the affective items, these results mainly indicate that the four attitude components are indeed related, but distinct components.

This study has some limitations. First, the aim of this study was to develop and validate a measurement instrument for street-level bureaucrats' attitude towards clients. In general, the fit parameters of the multicomponent model indicate a good model fit. Nevertheless, the model parameters are somewhat ambivalent with regard to the AVE assessment of convergent validity, mainly for the cognitive attitude component. A possible explanation for this observation might be that this AVE application is a relatively conservative measure for assessing convergent validity (Fornell & Larcker, 1981, p. 46). Nevertheless, the AVEs allude to the added value of further model testing.

Second, for the test study, these parameters were obtained using the same data for both the EFA and CFA. Performing EFA and CFA on the same sample of respondents can inflate model fit indices. CFA parameters obtained under these circumstances do not suffice as a stand-alone test of factorial validity. However, they can be used as an assessment of model

validity when their parameters are used for a comparison with CFA parameters derived from an entirely different sample of street-level bureaucrats. Moreover, that the replication study was totally unaffected by distortions from the items omitted in the test study strongly offsets this initial limitation.

Third, there is a risk of survey distortions resulting from social desirability bias and the use of cognitive measures to capture affective and behavioral attitudinal information. Social desirability bias incentives are deemed small in anonymous, self-administered survey research in PA, the more since most topics under inquiry in this academic field are not personally sensitive (Lee et al., 2012). Moreover, given the scale-development purposes of this study, social desirability is less of an issue due to our interest in how the scale items relate to each other, rather than their absolute values—assuming that social desirability bias would affect all four attitude items equally.

Also, cognitive measures may not have full access to an individual's affective and behavioral information (Breckler, 1984, p. 1193). Despite this reservation, the measurement instrument developed in this study was specifically designed for use in survey research. Survey research largely confines scholars to using cognitive measures only. Due to this restriction, it is commonly accepted to assess affective and behavioral information with cognitive measures (e.g., Eagly & Chaiken, 1993).

Fourth, the multicomponent model discerns different types of attitude content, but attitudes can also differ in their formation process: attitudes can form through implicit, subconscious or explicit, conscious processes (Olson & Kendrick, 2008, p. 118). In the literature on street-level bureaucracies, cues and information on clients are often assumed to be processed implicitly (e.g., Jilke & Tummers, 2018). And previous research has shown that attitudes that form through implicit processes tend to be more strongly grounded in an affective attitudinal base (Olson & Kendrick, 2008). Even though surveys are believed to make implicit attitudes explicit through their recall function (see DeVellis, 2003), the next steps to advance the understanding of street-level bureaucrats' attitude towards clients lie in exploring how this attitude is formed by unravelling the content and process drivers of this attitude.

Lastly, a precondition for any such instrument is its applicability across different types of street-level bureaucrats. The measurement instrument discussed here was mainly developed based on data gathered among street-level tax bureaucrats. Although bureaucrats' tasks differ depending on the type of bureaucracy they work for, they find common ground in their assessment of clients underlying their client-related activities. Consequently, seemingly different street-level work contexts can be compared due to structural similarities across all types of street-level bureaucracies (Lipsky, 1980). Nevertheless, before claims about applicability across cases can be made, this instrument requires testing beyond the tax administration to verify its external validity.

Despite some limitations, this study has merit for multiple reasons. Knowledge of core PA concepts "can be achieved only through more systematic attention to concept development and measurement specification" (Pandey & Scott, 2002, p. 578). The current study answers a loudly voiced call to PA scholars to use insights from other fields to advance PA theory and measurement (Grimmelikhuijsen et al., 2017; Perry, 2012).

This article showed that studies exploring street-level bureaucrats' attitude towards clients often merely touch upon some elements thereof, and that thorough conceptualizations of this construct are scarce. Building on psychological attitude theory, this study brought the cognitive, affective and behavioral attitudinal elements interwoven with the rich descriptions and narratives of the bureaucratic encounter together in one cohesive conceptual framework of street-level bureaucrats' attitude towards clients. This conceptual framework served as the foundation of a measurement instrument with sound psychometric qualities that showed that the multicomponent model of bureaucrats' attitude towards clients measures a new, distinctive construct in PA. It also showed that, to fully capture this attitude, four components should be used in future measurements thereof. By means of these merits, the multicomponent assessment of bureaucrats' attitude towards clients allows scholars to move beyond current practices that mainly tap into the cognitive dimension of this attitude. As a result, this study is exemplary of how introducing psychological insights to PA attitude theory and measurement can advance our understanding of core PA concepts.

In addition to avenues for future testing, a future research agenda should include exploring the applicability of the multicomponent measurement instrument for bureaucrats' attitude towards clients. A main use of PA research lies in identifying causal relationships (Lee et al., 2012). Examples of causal relations to be explored include those between bureaucrats' attitude towards clients and a variety of bureaucratic behaviors, such as the coping strategies they employ to deal with pressures inherent to working in a street-level bureaucracy (see Baviskar & Winter, 2017; Tummers & Musheno, 2015). A second avenue lies in the link between bureaucrats' general client attitude and their categorizations of clients, such as their perceived deservingness (see Jilke & Tummers, 2018).