


Understanding Felt Accountability

The institutional antecedents of the felt accountability of agency-CEO's to central government

Thomas Schillemans¹ | Sjors Overman¹ | Paul Fawcett² |
Matthew Flinders³ | Magnus Fredriksson⁴ | Per Laegreid⁵ |
Martino Maggetti⁶ | Yannis Papadopoulos⁶ | Kristin Rubecksen⁵ |
Lise Hellebø Rykkja⁵  | Heidi Houlberg Salomonsen⁷ |
Amanda Smullen⁸ | Matt Wood³

¹Utrecht University, The Netherlands

²University of Melbourne, Australia

³University of Sheffield, UK

⁴University of Gothenburg, Sweden

⁵University of Bergen, Norway

⁶University of Lausanne, Switzerland

⁷Aarhus University, Denmark

⁸Australian National University, Australia

Funding information

Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Number: NWO vidi 452-14-008

Abstract

The literature on autonomous public agencies often adopts a top-down approach, focusing on the means with which those agencies can be steered and controlled. This article opens up the black box of the agencies and zooms in on their CEO's and their perceptions of hierarchical accountability. The article focuses on felt accountability, denoting the manager's (a) expectation to have to explain substantive decisions to a parent department perceived to be (b) legitimate and (c) to have the expertise to evaluate those decisions. We explore felt accountability of agency-CEO's and its institutional antecedents with a survey in seven countries combining insights from public administration and psychology. Our bottom-up perspective reveals close connections between de facto control practices rather than formal institutional characteristics and felt accountability of CEO's of agencies. We contend that felt accountability is a crucial cog aligning accountability holders' expectations and behaviors by CEO's.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Governance* published by Wiley Periodicals LLC.

The disaggregation of the state into a plethora of quasi-autonomous arm's-length agencies represents a core element of the transition from "government to governance" (Dommett & MacCarthaigh, 2016; Pierre & Peters, 2020). This is reflected in the literature on the "unbundling" (Pollitt & Talbot, 2004) and "unraveling" (Hooghe & Marks, 2003) of the central state. The extending "chain of delegation" between central governments and the quasi-autonomous agencies that deliver public services, regulate certain sectors or make critical decisions about the distribution of public resources, has generated an extensive subfield of scholarship on accountability and control in public administration and political science (Bovens, Schillemans, & Goodin, 2014; Busuioc & Lodge, 2016; Dubnick & Romzek, 1993; Verhoest, Van Thiel, Bouckaert, Lægheid, & Van Thiel, 2012).

A major gap in the existing research base can be identified in understanding how institutional design and control are *actually perceived* on an individual level by key decision-makers in agencies in public administration (Olsen, 2013; Romzek, LeRoux, & Blackmar, 2012). Very few studies have opened the lid on the black box of unbundled governance in ways that unite "governance as theory" and "governance as practice" (Chhotray & Stoker, 2008). Most studies in public administration and political science start off from a top-down perspective and focus on the control-problems faced by politicians and government departments (Brehm & Gates, 1999; Verhoest et al., 2012). Very few studies have examined how the leadership of agencies actually perceive and experience those attempts. Yet, such a bottom-up perspective is crucial if we are to understand the impact of the institutional environment on their decisions and behaviors in practice.

This article seeks to address this gap in the existing research base through a focus on the perception of hierarchical accountability by the chief executive officers (CEOs) of quasi-autonomous agencies. We combine theories and insights from public administration and psychological research on control and accountability. We focus on felt accountability which is "a natural bridging construct between the individual and institutional levels of analysis" (Lerner & Tetlock, 1999:256). Felt accountability denotes the expectation that one's decisions and behaviors will be evaluated (Hall, Frink, & Buckley, 2017; Hochwarter et al., 2007) by an "able" (Mulgan, 2003) accountability forum with authority (Simon, 1997). Felt accountability has been translated to public administration research from psychology, where its effects on individual decisions and behaviors have been studied in hundreds of experimental studies.¹ Individually felt accountability is conceived as a *response* to the accountability environment in which the individual operates (Hall, Bowen, Ferris, Royle, & Fitzgibbons, 2007).

The research question is: "what are the institutional antecedents of CEOs' felt accountability? This question is relevant from a public administration and political science perspective, as much has been written about the design and control of agencies. Some scholars have focused on the appropriate institutional design of agencies while others believe de facto control practices are more important. Divergent factors such as formal legal status, organizational autonomy, the use of sanctions, monitoring or (in)formal contacts have all been noted as important (Bertelli, 2006; Koop, 2014; Van Slyke, 2006; Van Thiel & Yesilkagit, 2011; Verhoest, Peters, Bouckaert, & Verschuere, 2004) and little is known about which of these factors has the strongest impact. We study these factors as potential institutional antecedents of felt accountability of agency-CEO's. This is also relevant in light of psychological research, as the (institutional) antecedents of felt accountability are under-explored (Hall et al., 2017).

Our analysis is based on survey data (N = 498; response rate 45%) from the CEO's of (semi-) autonomous agencies in seven countries: Australia, Denmark, the Netherlands, Norway,

Sweden, Switzerland, and UK. The article is not comparative but focuses on similarities found across those countries.

As great expectations regarding public sector agencies often remain elusive (Overman, 2016), our findings on felt accountability matter for a number of reasons. First of all, the findings provide an insight into CEO's perceptions and feelings in relation to their direct principals, offering important theoretical clues on what dimensions of their institutional context really matter to them. Secondly, our results also have immediate policy implications for the design and governance of agencies (Verhoest et al., 2012). Finally, by combining insights from psychology with those from public administration and political science, we contribute to the expanding new literatures on individual behaviors in public administration (Battaglio Jr et al., 2019), focusing particularly on behavioral responses to accountability (Schillemans, 2016).

The article will first develop the concept of felt accountability (dependent variable), followed by the institutional antecedents (independent variables) used in this study. After the methods section, the article will discuss the results, including both full and partial confirmations of expectations as well as some surprises. These results will be discussed in the discussion section.

1 | FELT ACCOUNTABILITY

In public administration studies of accountability, the complexity and multiplicity of the concept (Bovens et al., 2014; Romzek & Dubnick, 1987), in combination with its missing, contradictory or even perverse effects (Halachmi, 2014; Koppell, 2005), often take central stage. Notwithstanding the notorious complexity, authors signal the rise of a “minimal conceptual consensus” (Bovens et al., 2014:6). Key to this is the relational nature of the concept—some actor is accountable to some forum, principal or audience—and the expectation of answerability. Accountability is furthermore also consequential; the actors' conduct or behavior is evaluated and sanctions and rewards may be used. These elements of the minimal conceptual consensus are consistent with most important typologies and definitions in different disciplines (Ferris, Rosen, & Barnum, 1995; Hall et al., 2007; Lerner & Tetlock, 1999; Mulgan, 2003; Olsen, 2013; Romzek & Dubnick, 1987).

In psychology, accountability is seen as a crucial “bridging mechanism” that links individual behaviors to social and institutional contexts (Hall et al., 2017). Much of this literature is based upon or inspired by Tetlock's social contingency model which “encourages communications” (1992:331) between individual-level behavioral research and scholars in for instance organization- or political science. The social contingency model assumes that people are concerned with their image, social status and power position. As a result, their responses to accountability are driven by their attempts to enhance or protect their social standing vis-à-vis important “others”. This aligns with recent studies in public administration on the reputational bases of accountability (Busuioac & Lodge, 2016).

Tetlock (1992) identifies three individual coping strategies in response to accountability, depending on how accountability is enacted. Decision-makers as “intuitive politicians” will first of all try to please their accountability forums when this is possible. They will then shift their position and subscribe to the forum's view on the issue (*acceptability heuristic*). When they are held accountable for past events, however, they are likely to resort to *defensive bolstering* in which they rationalize past behaviors and self-justify. The third coping strategy is *pre-emptive self-criticism*. The latter option is theoretically superior in lieu of the quality of decision-making

(Tetlock, 1983), yet requires high levels of cognitive effort which individuals will only choose to invest when they feel they have to.

Crucially, the social contingency model assumes that accountability is a state of mind rather than a state of affairs (Tetlock, 1992). That is, what matters is how accountability is perceived by the individual, rather than its formal, objective characteristics (Ferris et al., 1995). We therefore study the individual's felt accountability (Frink & Klimoski, 1998; Hall et al., 2017), which is defined as:

“The 1) implicit or explicit expectation that one's decisions or actions will be subject to evaluation by 2) some salient audience(s) with the belief that there exists the potential for one to receive either rewards or sanctions based on this expected evaluation.” (Hochwarter et al., 2007: 227, emphasis added).

This definition combines the *expectation* of accountability with the *salience*² of, in our words, the accountability forum, which should be capable of *evaluation* (Hall et al., 2007; Hall, Frink, & Buckley, 2017b). Felt accountability thus consists of several dimensions. We will focus on three: the *expectation* to be held accountable by a *legitimate* accountability forum with sufficient *expertise* to evaluate.

Felt accountability first of all denotes “expected” accountability. The individual expects that her decisions are likely to be scrutinized in the future which increases the efforts put into decision-making and the extent to which the agent anticipates the viewpoints of her accountability forum. Lerner & Tetlock 1999: 257 discuss the different effects of pre-decisional (expected) versus postdecisional (unexpected) accountability. Decision-makers who expect accountability yet do not know the exact preferences of their forum are found to process more information, reflect more broadly and thoroughly on their choices and take better-informed decisions (Hall et al., 2017; Lerner & Tetlock, 1999). They are likely to engage in pre-emptive self-criticism. Expected accountability may also reduce decision biases that threaten to punctuate the quality of decision making (Tetlock, 1983), although accountability is also sometimes found to increase biases (Klimoski & Inks, 1990).³ Expected accountability denotes the *anticipation of future accountability* which aligns with recurring findings in public administration research on the anticipation of accountability (Meijer, 2001; Papadopoulos, 2003:482).

When accountability comes as a surprise (postdecisional accountability) it typically prompts defensive bolstering in which individuals rationalize their behavior and try to save their social status (Lerner & Tetlock, 1999). They self-justify rather than self-reflect and this may lead to normatively more questionable responses, such as lying, but also perceptions of red tape (Bozeman & Scott, 1996) or work overloads (Power, 2003).

The second dimension of felt accountability is the perceived salience of the accountability forum which is implied in Hochwarter et al.'s (2007) definition. The ministry's salience is dependent on its “sanctioned acceptance” (Simon, 1997). Without the CEO's (implicit) acceptance of authority, she is not likely to be influenced in the, from a top down perspective, desirable way. To be seen as an able accountability forum, the ministry should nurture the impression of both its “authority of sanctions” and its “authority of ideas” (Simon, 1997, 187).

The authority of sanctions refers to the perceived *legitimacy* of accountability (Lerner & Tetlock, 1999, 258). In line with Suchman (1995, 574) we understand legitimacy to be a “generalized perception or assumption”, held, here, by the actor, that the accountability forum's scrutinizing role is “desirable, proper or appropriate” and in line with “norms” and “values” of the democratic system. Legitimate authority requires input based on respect and bears the actor's

“willingness to obey” (Simon, 1997, 180). This dimension is important, as “people’s acceptance of specific accountability measures and of accountability systems will also be influenced by the historical credibility of accountability measures” (Bergsteiner, 2012, 454). The legitimacy of accountability is specifically relevant in public administration settings where actors are accountable to multiple accountability forums (Busuioc & Lodge, 2016; Olsen, 2013).

Thirdly, the “authority of ideas” refers to the perceived *expertise* of the accountability forum. Do CEO’s acknowledge the parent department’s expertise to properly evaluate their decisions? The perception of expertise is important as many agencies perform highly specialized tasks (Maggetti & Ewert, 2018) and the delegation to agencies aggravates information asymmetries. Expertise is also crucial for the effective use of authority so that one’s “decisions will be accepted as decisional premises by the other” (Simon, 1997:189). The forum’s capacity to oversee the agent and to prevent her from “shirking or sabotage” (Brehm & Gates, 1999) is a necessary precondition for any accountability relationship to work (Strøm, 2000) and there is ample evidence of failing accountability forums in public administration (Schillemans & Busuioc, 2015). The capacity to monitor and evaluate requires knowledge, skills and abilities (Bergsteiner, 2012, 455) and are crucial to accountability (Ferris et al., 1995).

Felt accountability, thus, denotes three related subconcepts of *expected* accountability to a *legitimate* accountability forum with the *expertise* to appropriately evaluate the agent’s behavior. The felt relational accountability scale captures and further elaborates on these three dimensions (Overman, Schillemans, & Grimmelikhuijsen, 2020; see also Appendix D). These subdimensions do not necessarily add up or develop in line. One may expect accountability to a forum with or without expertise and with or without legitimacy. And, obviously, a legitimate forum with high expertise may nevertheless not be expected to demand any accountability at all. The three dimensions are nevertheless related as they are constitutive parts of a successful accountability alignment of agent to forum. When the CEO of the agency expects, accepts and acknowledges her upward accountability to central government, her behavior can also be expected to be aligned to the expectations of central government. Conversely, however, one could expect that a setting of high expected accountability with low perceived legitimacy and expertise could lead to the negative responses to accountability documented in the literature, such as gaming, lying and obstruction (Bevan & Hood, 2006; Halachmi, 2014).

2 | INSTITUTIONAL ANTECEDENTS OF FELT ACCOUNTABILITY

In psychological studies, the individual’s felt accountability is understood as a response to her accountability environment (Hall et al., 2007), although there will be differences between individuals in the same conditions (Frink & Klimoski, 1998). According to Hall et al. (2007, 2017), accountability environments vary along four dimensions. Accountability *source* refers to what we call the accountability forum: the person or institution to whom accountability is rendered. In this study, the source of accountability studied is central government as the principal of the CEO. Accountability *focus* is the extent to which individuals are held accountable for decision processes or outcomes (not the object of this study). Accountability *salience* refers to whether one is held accountable for important decisions or outcomes. We study the accountability of CEO’s for decisions regarding their organizations’ most important task. This implies that we study cases of high accountability salience. Finally, accountability *intensity* denotes the multiplicity of accountability sources. The four features of the accountability environment are seen

as potential antecedents of felt accountability, yet they have not been modeled as such in existing studies (Hall et al., 2017:209).

We focus on the first element: the accountability source (or forum) as antecedent of felt accountability. More specifically, we focus on crucial institutional features of the governance of agencies in public administration which are expected to affect CEO's felt accountability. In the literature on governmental control of agencies, a large number of factors have been identified as relevant. Authors have for instance discussed the relevance of legal types (Bertelli, 2006), dimensions of autonomy (Verhoest et al., 2004) patterns of interaction (Van Thiel, 2016), exchange of information (Koop, 2014), and monitoring and sanctioning (Braithwaite, 1997). All of these factors suggest an underlying *logic of proximity*, comparable to the idea of accountability salience: when the agency is "closer" to its parent department, higher levels of felt accountability for CEOs can be expected. In those cases, parent departments are more present in the professional life world of CEO's, triggering a higher sense of accountability. This proximity however can materialize in different ways and revolves around the tension between formal institutional design versus de facto inter-organizational interactions.

2.1 | Formal Institutional design versus de facto control practices

At a general level, central governments can use two generic strategies in their relations with agencies. The first is *formal institutional design* of agencies. Central governments shape the formal-structural conditions in which agencies operate. Laws, rules and regulations are the first causes of individual perceptions of accountability (Ferris et al., 1995:187) and the degree of power and control of the forum is one of the most critical external influences on agents (Bergsteiner, 2012:451). Central governments construct specific legal-organizational types, determine the levels of formal autonomy and assign and remove tasks. According to the *structural perspective* on governance (Christensen & Lægreid, 2006), such baseline formal dimensions are crucial in shaping how agencies relate to central government. Existing studies show how their formal legal status (Van Thiel, 2012) or autonomy (Carpenter, 2001) are crucial for agencies. Such formal factors may affect the felt accountability of agency CEO's. The underlying logic is that some agencies work closer to central government and have less autonomy in taking policy or financial decisions than others. Formal constraints necessitate them to take their principal into account when taking decisions and they, thus, feel more accountable.

A second general strategy refers to actual practices of control, such as contacts and the exchange of information. Although formal, structural perspectives have been widely used and developed in the literature, they have also been criticized because they are not always descriptive of real relations in the public sector (Skelcher, 2010) or because central theoretical assumptions are not found to hold empirically (Schillemans & Busuioc, 2015). Alternative approaches to the governance of agencies therefore focus on informal relations (Romzek et al., 2012), relational contracting (Bjurstrøm, 2020) or stewardship theory (Van Slyke, 2006). What those alternatives have in common, is that they assume that actual practices of interaction between individuals from different organizations are more important than formal institutional settings. This aligns with behavioral research suggesting that mechanisms of social control are more important than formal institutional design in generating perceptions of accountability (Ferris et al., 1995:187). De facto control and autonomy, as enacted in specific interactions between both parties, are then more relevant than institutional designs. In this perspective, the CEO's felt accountability is most likely to be shaped in face-to-face, enacted, formal but also informal,

interactions (Romzek et al., 2012). More frequent interactions then lead to a higher salience of the forum and additionally create social ties which lead to higher felt accountability. Interactions may either refer to the exchange of information, formal and informal meetings as well as the use of sanctions and rewards. The latter may come with relational costs, as will be subsequently explained.

2.2 | Six potential antecedents

Formal design and informal interactions are, although theoretically divergent, not mutually exclusive strategies. De facto interactions are predicated on de jure rights, rules and regulations. The potential institutional antecedents of felt accountability are, to varying degrees, marked by both strategies. We identify six potential antecedents, ordered on a continuum starting with the formal legal type of agencies—which is the exclusive product of formal-institutional design—and ending with inter-organizational contacts, which have an (almost) exclusively de facto character.

i. Organizational type is the most commonly identified structural characteristic of agencies (Verhoest et al., 2012). In most countries, there are different legal types of agencies, characterized by their specific legal status and proximity to central government. Organizational type has been found to be related to accountability Bertelli (2006). It can be expected that CEO's in more “distant” types of agencies subsequently experience less felt accountability toward their parent departments.

ii. Financial autonomy relates to the extent to which agencies can take specific financial decisions themselves (Verhoest et al., 2004). For agencies, the financial dimension of autonomy is crucial and, in combination with resources, important for their influence and power (Meier (1980). Public Management reforms have often sought to increase the financial autonomy of agencies while simultaneously constraining their policy autonomy with quantitative targets (Pollitt, 2005). It could be assumed that financial autonomy also affects felt accountability, in the sense that low financial autonomy leads to a stronger sense of felt accountability toward parent departments on whom the agency is then more dependent for crucial financial decisions.

iii. Policy autonomy is a second important institutional feature. Bureaucratic autonomy refers to agencies being separated from the actors who control them and capable of “sustained patterns of actions consistent with their own wishes” (Carpenter, 2001:14). This is related to organizational type yet it is more fine-grained and also carries a clear informal dimension on the basis of formal rights and regulations (Christensen & Lægreid, 2006). Policy autonomy is about the extent to which an agency is autonomous in making crucial policy decisions. One could expect, everything else being equal, an inverse relationship between policy autonomy and felt accountability, where increases in the one are mirrored by decreases in the other.

iv. Sanctions and rewards are crucial parts of the hierarchical relationship between departments and agencies. As principals, departments may sanction or reward (un)satisfactory performance or decisions by agencies with various means, ranging from budgetary measures, the (re) allocation of tasks, binding directives, dismissal or promotion. Sanctions are usually seen to form a ladder, escalating from the “ability to shame” via “lighter weapons such as certificates or formal (dis)approvals” to the “nuclear weapon of liquidation” (Hood, 2005:47). The possibility and severity of sanctions are crucial in making individuals feel accountable to specific

accountability “sources” (forums) in contexts of multiple accountability (Hall et al., 2007:408). Sanctions are most important in situations characterized by ambiguity and operational autonomy (Ferris et al., 1995:181), as in our case of semi-autonomous agencies. The “sanctions lobby” as Bergsteiner (2,102:294) nicely puts it, would claim that this is the most important element in an accountability relationship. More specifically, we would expect that sanctions and rewards are mainly positively related to the first subdimension of felt accountability: expected accountability.

v. *Reporting practices* are further important parts of the glue connecting (quasi)autonomous agencies to their parent departments. Agencies report in a multitude of ways, ranging from evaluations to annual reports, letters, external audits and informal talks. Much of this reporting is mandatory although scholars have also found high levels of voluntary reporting by agencies (Koop, 2014). Intense performance monitoring is an important management tool and conveys important signals as to what goals or actions are seen as important. Intense reporting practices can thus be expected to lead to higher levels of felt accountability by CEO's (Ferris et al., 1995, 183–4).

vi. *Inter-organizational contacts* between parent departments and agencies may be important to grease the wheels of collaboration. Agencies and parent departments are mutually dependent and collaborate on a long term basis in a setting with formally shared goals but practically diverging interests. A commonsensical adage to follow in this setting would be “Don't be a stranger!”. Various scholars stake the claim that frequent (in)formal contacts lead to trust and are imperative to proficient working relations between agencies and departments (Van Slyke, 2006; Van Thiel & Yesilkagit, 2011). High contact frequencies may be necessary “bonding costs” (Van Thiel, 2016:48; Ferris et al., 1995) and higher contact frequencies can be expected to be positively related to felt accountability.

2.3 | Negative feedback

The active use of monitoring and sanctioning may be conducive of organizational compliance but the use of monitoring and sanctions may come with relational costs (Braithwaite, 1997). Monitoring and sanctioning may have a negative impact on the subdimension of legitimacy, as parent departments may be experienced to be over-demanding (Flinders & Tonkiss, 2016). CEOs may find high levels of information-demanding by departments to be a nuisance and a form of accountability overload (Halachmi, 2014) or audit explosion (Power, 2003). Behavioral research suggests that “abusive supervision” (Hall et al., 2017:205) or “illegitimate accountability” (Lerner & Tetlock, 1999:258) generates negative feedback. Higher levels of monitoring and sanctioning could result in managers feeling that the expected accountability is not at all legitimate, because the department could be seen to be micro-managing, even meddling with, intra-organizational affairs (Greiling & Spraul, 2010). Constraining people can lead to reactance (Ferris et al., 1995:184). Reactance is that individuals who feel that their freedoms are threatened will be emotionally aroused to regain their former freedoms. On this basis, the “motivation lobby” Bergsteiner (2012:294) champions empowering forms of accountability rather than sanctions. We thus expect that active use of sanctions and rewards, while likely to be positively related to expected accountability, is negatively related to the subdimension of forum legitimacy. These are the relational costs of controlling behaviors (Braithwaite, 1997).

3 | DATA AND METHODS

3.1 | Case selection and sample

We studied felt accountability among CEO's of agencies in seven advanced democracies: Australia, Denmark, the Netherlands, Norway, Sweden, Switzerland, and the UK. These countries have all gone through numerous reforms over the last decades, affecting the formal position of agencies, systems of accountability and public sector management in general, although with important national differences (Christensen & Lægheid, 2006; Verhoest et al., 2012). We selected 1,096 agencies at the national level in those countries. These were all agencies that carry out public tasks, but are not structurally a part of the parent department. We selected all of the agencies that can be classified as an internally autonomous agency without legal personality or as an autonomous agency with legal personality (cf. type 1 and 2 in the classification by Van Thiel, 2012).

We circulated an invitation to take part in the electronic survey among the agency heads between May and December 2017. We received a total of 661 responses, of which 496 were fully completed. Additional data on the institutional relationship between agencies and parent departments was collected from various external sources. This part of the data-collection was finished in March 2018.⁴

Table 1 specifies the number of complete responses per country. In most countries, a response rate of about 50% was achieved and there was an overall response rate of 45%.

In addressing the agencies, we used the strategy adopted by the earlier cost-cobra surveys (Verhoest et al., 2012) and aimed for the CEO of an organization, indicating that some of the descriptive parts of the survey could be delegated. Direct emails were used where possible, with several follow up emails. In the UK (but also in Switzerland) we were unable to collect as many direct email addresses as in the other countries, which probably explains some of the lower response rates. The survey was translated to the various national languages, except for Switzerland where we used an English version. Approximately half of the items were derived from existing scales, which are available in all five languages (see Table 3 for sources). Other items, including the felt accountability scale, were available in two languages and had to be translated three times. This was coordinated by the bilingual PI and RA and was performed and tested by the country teams.

TABLE 1 Overview responses

Country	Complete responses	Total invitations	Response rate
Australia	91	170	54%
Denmark	32	53	60%
Netherlands	56	105	53%
Norway	85	168	51%
Sweden	115	241	48%
Switzerland	49	116	42%
United Kingdom	68	243	28%
Total	496	1,096	45%

3.2 | Variables

“Felt accountability” (DV) was measured with the Felt Accountability Scale (see Appendix B). The scale has been validated and has now been applied in 19 different populations of civil servants, including the seven for this survey. Items and the reliability of the scales are listed in Table 2.

Organizational type was coded on the basis of the typology developed by Van Thiel (2012) by experts from the seven countries on the basis of public records and available databases of agencies. We focused on (a) semi-autonomous agencies without legal entity, and (b) statutory bodies or agencies with a public legal entity.

Financial autonomy was measured on a three-item scale, which measured whether the agency is mandated to take financial loans, shift between its budgets for personnel and running costs, and shift between budgets over multiple years. Together, these items formed an adequate scale ($\alpha = .65$).

Policy autonomy was measured on a five-item scale, which measured whether agencies had autonomy to set overall goals, identify target groups, choose between policy instruments, publicly communicate about the policy, and prioritize tasks. These items formed a reliable scale ($\alpha = .76$).

The *credibility of sanctions and rewards* was measured with a seven-item scale, identifying the availability of seven different sanctions and rewards and the perceived likelihood of use. These items formed a reliable scale ($\alpha = .77$).

Reporting practices were measured with a five-item scale that tapped into the frequency with which the parent department receives particular types of information relating to the agency's self-identified most important task. The items correlate and form an adequate scale ($\alpha = .68$).

Contact frequency was measured with a four-item scale measuring frequency of formal and informal contacts on different hierarchical levels. Answers formed a reliable scale ($\alpha = .86$).

We added control variables, both on the organizational level as well as on the individual level, on the basis of available public records in the seven countries, websites, annual reports,

TABLE 2 Felt accountability items

Expected accountability (Spearman Brown = .70)	I am held very accountable for our most important task.
	The parent department holds me accountable for all of my decisions.
Forum legitimacy ($\alpha = .65$)	When the parent department changes its views we just have to comply with this new reality.
	It is a good thing, that we are ultimately accountable to the parent department.
	I am willing to work in the interest of the parent department.
Forum expertise ($\alpha = .76$)	The parent department applies clear / understandable standards to evaluate our most important policy.
	The parent department provides constructive feedback on our work.
	Opinions from the parent department are generally unambiguous.
	The parent department has sufficient substantive or technical expertise about our work to oversee / evaluate our duties.

and existing databases. Data collection was coordinated by the PI yet performed in collaboration with academic experts on agencies and accountability in all seven countries.

- *Organizational size.* We used numbers of FTE's as a measure for size and coded this from various sources.
- *Policy field.* This was coded on the basis of websites and policy documents on the 10 international COFOG-criteria used to classify governmental functions.
- *Location of main office,* in either the (political) capital or elsewhere (peripheral).
- *Length of tenure.* Respondents were asked to indicate the number of years they had worked for the agency (or a predecessor).
- *Country.* We used the seven countries in our analysis.

Appendix A presents the descriptive statistics of all variables. Table B1 provides further analyses.

3.3 | Analysis

We conducted a regression analysis on all three dimensions of felt accountability. To deal with potential heteroskedasticity or influential cases, robust standard errors were estimated (Cribari-Neto & da Silva, 2011). We used a fixed effects model to control for the country differences. All analyses were conducted in R 3.4.3 using the sandwich 2.3–4 package for the estimation of robust standard errors (Zeileis, 2006), and the qvcalc 0.9–1 package for the calculation of quasi-variances to estimate country contrasts (Firth, 2003).

We controlled for common method bias which was necessary as several of our independent variables have been collected in the same survey. A statistical control was applied to reduce the risk of overestimated effect sizes. We conducted Harman's single factor test (Harman, 1976). This test is an exploratory factor analysis of all relevant variables from the same source. If, based on the eigenvalues, a single factor emerges, or if one factor captures more than 50% of the variance, this is an indication that common method bias is present. All of our models had three eigenvalues greater than one, and the proportion explained variance of the largest factor varied between 22.2% and 26.4%, which is well under the cut-off value of 50%. Therefore, common method bias is not likely to be influencing our results. In addition, as we are measuring attitudes and perceptions, the risk of common method bias might also be smaller than for research on other organizational aspects such as performance (Favero & Bullock, 2015).

4 | RESULTS: UNPACKING FELT ACCOUNTABILITY

4.1 | Felt Accountability: expectation, legitimacy, expertise

CEO's of autonomous agencies face a double bind. They are expected to focus more strongly on their own organization's performance and also on what societal stakeholders expect, yet they are simultaneously bound to their parent departments for political directives, legal status, and resources (Pollitt, 2005). Critics have been concerned that CEO's may turn their backs on their parent departments and wrestle free from central control. Our results now suggest that this concern is not generally warranted: the CEO's in our sample can be said to feel fairly accountable

in relation to their parent departments on all three dimensions of felt accountability. Responses vary on the three subdimensions, across the countries and also between individual respondents (see the Appendix for details).

The *expectation* of accountability is most pronounced among our respondents ($M = 5,73$, scale 1–7), signifying that CEO's of autonomous agencies are still more like “close neighbors” than “distant friends” to their parent departments (Van Thiel & Yesilkagit, 2011). This holds for the agencies in all seven countries, yet to varying degrees. This expected accountability is deemed to be quite *legitimate* by the respondents ($M = 5,13$), although the average answer is now more moderated. Accountability is thus a little more expected than that it is considered to be legitimate. When we move from the normative dimension of legitimacy to the cognitive dimension of perceived *expertise*, the mean response again drops a bit. The expertise of the parent department as accountability forum is the relatively weakest element in the perceptions of our respondents, and the mean is not much above the neutral response ($M = 4,61$).

We thus see relatively high levels of expected accountability to parent departments who are not always seen as legitimate, and even less often deemed to have proficient expertise. For policy-makers in parent-departments, the first takeaway, then, is that they should focus on how their expertise is perceived if they would want to increase felt accountability among CEO's. This is particularly relevant, as some of the policies are explicitly delegated to those agencies in order to cluster expertise (Van Thiel & Yesilkagit, 2011) and the delegation process widens the knowledge gap.

4.2 | Institutional antecedents of felt accountability

When CEOs feel more accountable toward their parent departments as legitimate principals with sufficient expertise, it is more likely that they will be more responsive toward the governments' policies and that the actions of their organizations are aligned with central governmental expectations. Felt accountability and its subdimensions are undoubtedly in part dependent on individual characteristics, but this is not part of this study. We also see some significant differences between the countries in our data, suggesting the relevance of cultural and systemic differences (see Appendix A). Beyond that, however, our analysis suggests there are robust relations between some of the institutional variables identified and (subdimensions of) felt accountability. Table 3 below displays the results on the institutional antecedents of felt accountability at a glance.

4.3 | De facto control: some strong findings

As Table 3 shows, our analysis suggests that higher reporting frequencies and a higher credibility of sanctions and rewards are both strongly related to all three dimensions of felt accountability. CEO's who report more often to their principal and consider it more likely that sanctions and rewards will be used, do expect to be accountable to their parent departments and also find this more salient in terms of legitimacy and expertise. Also, as will be discussed later, we did not find negative effects of sanctions and rewards on forum legitimacy, as was expected.

The most direct and straightforward interpretation of this finding would be that, contrary to some expectations in the literature, *hierarchy* may actually work and be relevant, even in times of more fragmented governance (Bergsteiner 2012; Pierre & Peters, 2020). In the control

TABLE 3 Regression analyses antecedents of felt accountability

	Expected Accountability		Forum Legitimacy		Forum Expertise	
	Estimate	SE	Estimate	SE	Estimate	SE
Organizational Type (ref. Type 1)						
Type 2	−0.882	0.589	−0.267	0.352	−0.241	0.458
Financial autonomy	0.250	0.147	0.221*	0.101	−0.085	0.126
Policy autonomy	−0.158	0.080	−0.387*	0.069	−0.202*	0.082
Credible sanctions and rewards	0.340*	0.116	0.648*	0.105	0.250*	0.117
Reporting practice	0.336*	0.136	0.293*	0.121	0.390*	0.134
Contact frequency	0.153	0.078	0.048	0.077	−0.034	0.076
Location Main Office	−0.030	0.104	0.239	0.098	−0.019	0.107
Size	−0.001	0.031	0.051	0.027	−0.032	0.031
Tenure	0.028	0.014	0.014	0.014	0.012	0.012
Intercept	5.63*	0.788	4.249*	0.536	4.361*	0.688
ΔF country**	3.518*	Δdf = 6	5.225*	Δdf = 6	6.527*	Δdf = 6
ΔF policy field**	1.474	Δdf = 9	2.403*	Δdf = 9	0.942	Δdf = 9
F	5.055*	df = 25; 430	11.18	df = 25; 442	3.888	df = 25; 443
R ²	0.227		0.387		0.179	
N	456		468		469	

* $p < .01$. Estimations and robust standard errors.

**All analyses include country and COFOG-classification effects, which are omitted in the table and presented in appendix if statistically significant. F-values based on ANOVA model comparison.

perspective embodied in principal-agent theory, monitoring and sanctioning stand out as key tools with which to curb the asymmetries between principal and agent. Our results suggest that CEO's feel more accountable to stronger principals who are more actively informed and may more credibly retaliate or reward. This suggests that the role of hierarchical control, not only practically but also in normative terms, may be reconsidered in the age of fragmented governance (Hill & Lynn, 2004; Jarvis, 2014).

There is however more to this, as the findings signify more than hierarchical control only. Our measures include mandatory *and* voluntary reporting (cf Koop, 2014) and sanctions *and* rewards. This suggests that the results are not only about enforcing control by central government but also about the CEO expecting rewards for good performance and opening up voluntarily to provide even more information than is requested. Also, only a very limited number of respondents indicate that sanctions are actually used (2,2–6,1%, depending on type). This is in line with earlier studies suggesting sanctioning comes with relational costs as “big sticks rebound” (Braithwaite, 1997). These findings do all in all suggest that it is important for government departments to induce the suggestion that they will use their powers of coercion *and* reward when this is necessary and appropriate, without necessarily using them.

An additional interpretation is that accountability serves more purposes than control only (Busuioac & Lodge, 2016; Dubnick & Romzek, 1993). Accountability can also be sense-giving and suggestive of status, reputation and relevance (Schillemans & Busuioac, 2015). By

demanding information and credibly signaling the intention to act upon it, parent departments signal that they actually *care* about the agency and that its work is significant. Departments then “convey social-relational signals to [agency] decision-makers” (Patil, Vieider, & Tetlock, 2014:72) suggesting they really care about their decisions and will act when necessary.

4.4 | Organizational type and autonomy: mixed findings

As Table 3 suggests, more formal potential antecedents of felt accountability had only limited and mixed effects on felt accountability.

On the one hand, the autonomy inherent in the legal status of the agency was not related to felt accountability of CEOs in our sample. The legal type of an agency is the most fundamental formal distinction between different agencies (Van Thiel, 2012) and is crucial in determining the boundaries between central governments and agencies. The legal type is also important, as it is frequently the object of administrative reform (Dommett and Maccarthaigh 2016). Governments tend to design comparable governance regimes for agencies with the same legal status (Bertelli, 2006). However, as important as legal status may be in many ways, it was not related to the felt accountability of CEO's.

On the other hand, policy autonomy—the ability to take substantive decisions on your own as an agency—was negatively related to the perceived expertise and legitimacy of the parent department while financial autonomy—the ability to take some financial decisions—was found to have a positive relation with forum legitimacy. When CEO's manage agencies with more policy autonomy, they react more negatively to upward accountability and find their hierarchical accountability relationship to be less legitimate. They may possibly feel that they are being micro-managed (Flinders & Tonkiss, 2016). CEO's of agencies with more financial autonomy find their accountability relationship with their parent department more legitimate. CEO's with a lot of financial autonomy by and large accept the accountability relationship they experience with their direct principal. It seems that financial autonomy may smoothen interactions between agencies and departments while policy autonomy, to the contrary, makes their interactions more difficult.

Autonomy features prominently in the literature on autonomous agencies (Christensen & Lægreid, 2006; Verhoest et al., 2004) and the design of appropriate formal rules and institutions is important from a policy perspective. Its impact on felt accountability, though, is limited and mostly related to the subdimension of legitimacy. Assuming felt accountability is desirable, our findings suggest that agencies should have some financial autonomy yet less policy autonomy.

4.5 | No indications of relational benefits or negative feedback

Our expectations regarding relational benefits and negative feedback were not confirmed. To begin with, we expected that warm and frequent interorganizational contacts would be conducive of proficient working relations. High levels of contact are said to be necessary to grease the wheels of cooperation between the two separate entities (Van Thiel & Yesilkagit, 2011). Frequent contacts could thus be expected to have a positive effect on felt accountability. However, as is clear in Table 3, we find no relations between contact frequency and felt accountability. This would suggest that even though frequent contacts may be relevant as such, the other dimensions have a much stronger effect on all dimensions of felt accountability.

Further, more frequent use of controlling measures such as sanctions and rewards could be expected to incur negative feedback on these relationships and particularly affect the

subdimension of legitimacy. Against the background of recurring criticisms of the audit society (Power, 2003), and accountability overloads (Halachmi, 2014), high levels of reporting and expected sanctioning could be expected to go hand in hand with much lower levels of forum legitimacy. We find, however, that they are even positively related to the subdimension of forum legitimacy. It seems that departments requesting more information and credibly signaling that they will respond with sanctions or rewards, strengthen felt accountability and in this process also build up their own legitimacy in the eyes of the CEOs.

5 | CONCLUSION AND REFLECTIONS

This article has set out to study how CEO's of (semi) autonomous agencies experience their accountability relationship with central government. We adopted a bottom-up approach, focusing on governance as practice rather than governance as theory (Chhotray & Stoker, 2008), exploring how institutional design and de facto control are *actually perceived* on an individual level (Olsen, 2013; Romzek et al., 2012). This article has gauged whether and when CEO's *expect* to be held accountable for their decisions, experience this as *legitimate* and whether their principals are felt to have sufficient *expertise* to properly evaluate them. Among the institutional antecedents of felt accountability two stand out: reporting practices and the credible use of sanctions and rewards. An informed proximity by central government as accountability forum that actually cares about what the agency does as an "exteriority within" (Wolters, 2019), and the anticipation of positive or punitive consequences, helps to make CEO's of government agencies feel more accountable.

This finding is of immediate relevance to public policy. The appropriate design and management of autonomous agencies are the objects of recurring policy initiatives in many established democracies, including the seven we studied for this article (Verhoest et al., 2012). Among the antecedents of felt accountability we included formal and "hard" institutional factors, such as the legal status of agencies, with "softer" yet, apparently crucial, de facto control practices. The analysis suggests that the latter clearly have the strongest impact on the felt accountability of CEO's. The policy implication is that policy-makers in central government would be well advised to shift some of their attentions away from formal-legal aspects to the actual interaction patterns with "their" agencies. They also need to assure that they have the necessary expertise to evaluate with their agencies properly.

The major theoretical contribution this article makes is that it integrates insights on felt accountability developed in psychology (Hochwarter et al., 2007; Hall et al., 2017) in the study of governance as practice, in line with the recent development toward a behavioral public administration (Battaglio Jr et al., 2019). This is relevant from the perspective of the psychological literature on accountability, as there has been little research on the antecedents of felt accountability and the role of the accountability audience (Hall et al., 2017). Also, by analyzing agency CEO's in seven countries, we conducted a study with higher ecological validity which was sensitive to national and cultural variations (Bergsteiner, 2012). By integrating micro-level research from psychology in mesolevel studies of public administration, we shed light on how and when salient institutional factors in governance affect key individuals in agencies.

The study of felt accountability on an individual level is important, as it has been related to many salient aspects of individual decision-making in laboratory studies (Hall et al., 2017>; Lerner & Tetlock, 1999). In their review of that body of literature, Hall et al., 2017: 208) contend that felt accountability may be "... the most pervasive (and perhaps even the most powerful)

single influence on human social behavior". This can be related to other "constructs of the mind" in public administration, such as public service motivation (Moynihan & Pandey, 2007) or emotional labor (Newman, Guy, & Mastracci, 2009). This is relevant if we are to understand governance as practice and want to gauge how institutional settings affect important individual decision-makers in public administration, whose behaviors and decisions have significant consequences for society.

ACKNOWLEDGMENTS

We would like to thank Maj Grove Jeppesen (Utrecht University) first of all, and also Daniel Bailey (University of Sheffield), Martin Moos (Aarhus University), Annbjørg Ryssdal (University of Bergen), Manuel Quaden (Utrecht University), Amanda Waldenström (University of Gothenburg), Ella Weisbrot (Australian National University) for their excellent research assistance.

ORCID

Lise Hellebø Rykkja  <https://orcid.org/0000-0001-7486-2349>

ENDNOTES

- ¹ See for reviews: Lerner & Tetlock, 1999; Hall et al., 2017; Harari & Rudolph, 2017; Aleksovska, Schillemans, & Grimmelikhuijsen, 2019.
- ² In psychological studies, "salience" may refer to the audience to whom accountability is rendered (Hochwarter et al., 2007:227; Ferris et al., 1995:186). But salience is also a dimension of the accountability environment. Accountability salience then denotes that decisions taken by the individual are "related in important ways to the overall effectiveness (or mission) of the group or organization" (Hall et al., 2007:408). We focus on the salience of the accountability forum.
- ³ There are more studies reporting attenuation than increases of bias (Aleksovska et al., 2019 et al, fig. 7. Appendix 3).
- ⁴ The dataset is available via <https://easy.dans.knaw.nl/>, DOI: 10.17026/dans-2ay-9avv.

REFERENCES

- Aleksovska, M., Schillemans, T., & Grimmelikhuijsen, S. (2019). Lessons from five decades of experimental and behavioral research on accountability: A systematic literature review. *Journal of Behavioral Public Administration*, 2(2), 1–18.
- Battaglio Jr., R. P., Belardinelli, P., Bellé, N., & Cantarelli, P. (2019). Behavioral public administration ad fontes: A synthesis of research on bounded rationality, cognitive biases, and nudging in public organizations. *Public Administration Review*, 79(3), 304–320.
- Bergsteiner, H. (2012). *Accountability theory meets accountability practice*. Melbourne: Emerald.
- Bertelli, A. M. (2006). Delegating to the Quango: Ex ante and ex post ministerial constraints. *Governance*, 19(2), 229–249.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538.
- Bjurström, K. H. (2020). Principal–Agent or Principal–Steward: How Ministry–Agency Relations Condition the Impact of Performance Management in the Steering of Government Agencies. *Public Performance & Management Review*, 1–25.
- Bovens, M., Schillemans, T., & Goodin, R. E. (2014). Public Accountability. In M. Bovens, R. E. Goodin, & T. Schillemans (Eds.), *The Oxford Handbook of Public Accountability* (pp. 1–20). Oxford: Oxford University Press.
- Bozeman, B., & Scott, P. (1996). Bureaucratic red tape and formalization: Untangling conceptual knots. *The American Review of Public Administration*, 26(1), 1–17.

- Braithwaite, J. (1997). On speaking softly and carrying big sticks: Neglected dimensions of a republication separation of powers. *The University of Toronto Law Journal*, 47(3), 305–361.
- Brandsma, G. J. (2010). *Backstage Europe: comitology, accountability and democracy in the European Union*. Utrecht University.
- Brehm, J. O., & Gates, S. (1999). *Working, shirking, and sabotage: Bureaucratic response to a democratic public*. Ann Arbor: University of Michigan Press.
- Busuioac, E. M., & Lodge, M. (2016). The reputational basis of public accountability. *Governance*, 29(2), 247–263.
- Carpenter, D. (2001). *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies* (pp. 1862–1928). Princeton University Press.
- Chhotray, V., & Stoker, G. (2008). *Governance theory and practice: A cross-disciplinary approach*. Basingstoke: Palgrave.
- Christensen, T., & Lægreid, P. (2006). *Autonomy and regulation: Coping with agencies in the modern state*. Edward Elgar Publishing.
- Cribari-Neto, F., & da Silva, W. B. (2011). A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. *AStA Advances in Statistical Analysis*, 95(2), 129–146.
- DeVellis, R. F. (2009). *Scale Development: Theory and Applications Applied Social Research Methods Series* (Vol. 26, 2nd ed.). Thousand Oaks, CA: Sage Publication.
- Dommett, K., & MacCarthaigh, M. (2016). Quango reform: the next steps? *Public Money & Management*, 36(4), 249–256.
- Dubnick, M., & Romzek, B. (1993). Accountability and the Centrality of Expectations in American Public Administration. In J. Perry (Ed.), *Research in Public Administration* (pp. 37–78). Greenwich: JAI, Press.
- Favero, N., & Bullock, J. (2015). How (Not) to Solve the Problem: An Evaluation of Scholarly Responses to Common Source Bias. *Journal of Public Administration Research and Theory*, 25(1), 258–308.
- Ferris, G., Rosen, S. D., & Barnum, D. T. (Eds.). (1995). *Handbook of human resource management*. Hoboken: Blackwell.
- Firth, D. (2003). Overcoming the reference category problem in the presentation of statistical models. *Sociological Methodology*, 33(1), 1–18.
- Flinders, M., & Tonkiss, K. (2016). From ‘poor parenting’ to micro-management: coalition governance and the sponsorship of arm’s-length bodies in the United Kingdom, 2010–13. *International Review of Administrative Sciences*, 82(3), 490–515.
- Frink, D. D., & Klimoski, R. J. (1998). Toward a Theory of Accountability in Organizations and Human Resource Management. In *Research in Personnel and Human Resources Management* (Vol. 16, pp. 1–51). US: Elsevier Science/JAI Press.
- Greiling, D., & Spraul, K. (2010). Accountability and the challenges of information disclosure. *Public Administration Quarterly*, 338–377.
- Halachmi, A. (2014, 2014). Accountability overloads. In M. Bovens, R. E. Goodin, & T. Schillemans (Eds.), *The Oxford Handbook of Public Accountability*. Oxford: Oxford University Press.
- Hall, A. T., Bowen, M. G., Ferris, G. R., Royle, M. T., & Fitzgibbons, D. E. (2007). The accountability lens: A new way to view management issues. *Business Horizons*, 50(5), 405–413.
- Hall, A. T., Frink, D. D., & Buckley, M. R. (2017). An accountability account: A review and synthesis of the theoretical and empirical research on felt accountability. *Journal of Organizational Behavior*, 38, 204–224.
- Harari, M. B., & Rudolph, C. W. (2017). The effect of rater accountability on performance ratings: A meta-analytic review. *Human Resource Management Review*, 27(1), 121–133.
- Harman, H. (1976). *Modern factor analysis* (3rd ed.). Chicago: Univ. of Chicago Press.
- Hill, C. J., & Lynn, L. E. (2004). Is hierarchical governance in decline? Evidence from empirical research. *Journal of Public Administration Research and Theory*, 15(2), 173–195.
- Hinkin, T. R. (1998). A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods*, 1(1), 104–121. <https://doi.org/10.1177/109442819800100106>
- Hochwarter, W. A., Ferris, G. R., Gavin, M. B., Perrewé, P. L., Hall, A. T., & Frink, D. D. (2007). Political Skill as Neutralizer of Felt Accountability—job Tension Effects on Job Performance Ratings: A Longitudinal Investigation. *Organizational Behavior and Human Decision Processes*, 102(2), 226–239. <https://doi.org/10.1016/j.obhdp.2006.09.003>

- Hood, C. (Ed.). (2005). *Controlling modern government: Variety, commonality and change*. Edward Elgar Publishing.
- Hooghe, L., & Marks, G. (2003). Unraveling the central state, but how? Types of multi-level governance. *American Political Science Review*, 97(2), 233–243.
- Jarvis, M. D. (2014). Hierarchical Accountability. in: Bovens, M, R. E. Goodin, and T. Schillemans. 2014. *The Oxford Handbook of Public Accountability*. Oxford: Oxford University Press., 405–420.
- Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. *Organizational Behavior and Human Decision Processes*, 45(2), 194–208.
- Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* (4th ed.). Guilford Publications.
- Koop, C. (2014). Theorizing and explaining voluntary accountability. *Public Administration*, 92(3), 565–581.
- Koppell, J. G. (2005). Pathologies of accountability: ICANN and the challenge of “multiple accountabilities disorder”. *Public Administration Review*, 65(1), 94–108.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255–275.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model Modifications in Covariance Structure Analysis: The Problem of Capitalization on Chance. *Psychological Bulletin*, 111(3), 490–504.
- Maggetti M., Ewert C. (2018) Comparative Regulatory Regimes and Public Policy. In: Ongaro E., Van Thiel S. (eds) *The Palgrave Handbook of Public Administration and Management in Europe*. Palgrave Macmillan, London
- Meier, K. J. (1980). Measuring organizational power: Resources and autonomy of government agencies. *Administration & Society*, 12(3), 357–375.
- Meijer, A. (2001). Accountability in an information age: opportunities and risks for records management. *Archival Science*, 1(4), 361–372.
- Moynihan, D. P., & Pandey, S. K. (2007). The role of organizations in fostering public service motivation. *Public Administration Review*, 67(1), 40–53.
- Mulgan, R. (2003). *Holding power to account: accountability in modern democracies*. New York: Springer.
- Newman, M. A., Guy, M. E., & Mastracci, S. H. (2009). Beyond cognition: Affective leadership and emotional labor. *Public Administration Review*, 69(1), 6–20.
- Olsen, J. P. (2013). The institutional basis of democratic accountability. *West European Politics*, 36(3), 447–473.
- Overman, S. (2016). Great expectations of public service delegation: A systematic review. *Public Management Review*, 18(8), 1238–1262.
- Overman, S. (2020). Aligning accountability arrangements for ambiguous goals: the case of museums. *Public Management Review*, 1–21. <https://doi.org/10.1080/14719037.2020.1722210>
- Overman, S., Schillemans, T., & Grimmelikhuijsen, S. (2020). A validated measurement for felt relational accountability in the public sector: gauging the account holder's legitimacy and expertise. *Public Management Review*, 1–20. <https://doi.org/10.1080/14719037.2020.1751254>
- Papadopoulos, Y. (2003). Cooperative forms of governance: Problems of democratic accountability in complex environments. *European Journal of Political Research*, 42(4), 473–501.
- Patil, S. V., Vieider, F., & Tetlock, P. E. (2014). Process versus outcome accountability. In Bovens, M., Goodin, R. E., & Schillemans, T. (eds.), *The Oxford handbook of public accountability*, Oxford: Oxford University Press, pp. 69–89.
- Pierre, J., & Peters, B. G. (2020). *Governance, Politics and the State* (2nd rev ed.). London: Palgrave.
- Pollitt, C. (2005). Performance management in practice: A comparative study of executive agencies. *Journal of Public Administration Research and Theory*, 16(1), 25–44.
- Pollitt, C., & Talbot, C. (Eds.). (2004). *Unbundled government: A critical analysis of the global trend to agencies, quangos and contractualisation*. London: Routledge.
- Power, M. (2003). Evaluating the audit explosion. *Law & policy*, 25(3), 185–202.
- Pressman, J. L., & Wildavsky, A. B. (1973). *Implementation: How Great Expectations in Washington Are Dashed in Oakland; or, Why It's Amazing That Federal Programs Work at All; This Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers, Who Seek to Build Morals on a Foundation of Ruined Hopes Oakland Project Series* (Vol. 1, Paperback ed.). Berkeley, CA: University of California Press.

- Romzek, B. S., & Dubnick, M. J. (1987). Accountability in the public sector: Lessons from the Challenger tragedy. *Public Administration Review*, 47, 227–238.
- Romzek, B. S., LeRoux, K., & Blackmar, J. M. (2012). A preliminary theory of informal accountability among network organizational actors. *Public Administration Review*, 72(3), 442–453.
- Schillemans, T. (2016). Calibrating Public Sector Accountability: Translating experimental findings to public sector accountability. *Public Management Review*, 18(9), 1400–1420.
- Schillemans, T., Bokhorst, M., van Genugten, M., & Oude Vrielink, M. (2018). Voorlopige contouren van bestuursgericht toezicht. Empirische inzichten in jonge praktijken van bestuursgericht toezicht. *Bestuurskunde*, 27(4), 54–66.
- Schillemans, T., & Busuioac, M. (2015). Predicting public sector accountability: From agency drift to forum drift. *Journal of Public Administration Research and Theory*, 25(1), 191–215.
- Schillemans, T., & Kremers, G. (2020). The unknown world of transnational organizations in Europe—challenges for accountability. *Public Money & Management*, 40(4), 330–334.
- Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). MIT press.
- Skelcher, C. (2010). Fishing in muddy waters: Principals, agents, and democratic governance in Europe. *Journal of Public Administration Research and Theory*, 20(suppl_1), i161–i175.
- Strøm, K. (2000). Delegation and accountability in parliamentary democracies. *European Journal of Political Research*, 37(3), 261–290.
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610.
- Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, 46(4), 285–292.
- Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. *Advances in Experimental Social Psychology*, 25(3), 331–376.
- Van Slyke, D. M. (2006). Agents or stewards: Using theory to understand the government-nonprofit social service contracting relationship. *Journal of Public Administration Research and Theory*, 17(2), 157–187.
- Van Thiel, S. (2012). Comparing Agencies across Countries. In K. Verhoest, S. Van Thiel, G. Bouckaert, & P. Lægheid (Eds.), *Government Agencies: Practices and Lessons from 30 Countries*. New York: Palgrave Macmillan.
- Van Thiel, S. (2016). *A principal-agent perspective*. In *Theory and Practice of Public Sector Reform* (pp. 54–70). Routledge.
- Van Thiel, S., & Yesilkagit, K. (2011). Good neighbours or distant friends? Trust between Dutch ministries and their executive agencies. *Public Management Review*, 13(6), 783–802.
- Verhoest, K., Peters, B. G., Bouckaert, G., & Verschuere, B. (2004). The study of organisational autonomy: a conceptual review. *Public Administration and Development*, 24(2), 101–118.
- Verhoest, K., Van Thiel, S., Bouckaert, G., Lægheid, P., & Van Thiel, S. (Eds.). (2012). *Government agencies: practices and lessons from 30 countries*. London: Springer.
- Waldo, D. (1948). *The Administrative State*. The Ronald Press Company.
- Wolters, R. (2019). Accountability matters. In *Enacting accountability in a profession's web of relations*. Nijmegen: Radboud University.
- Zeileis, A. (2006). Object-Oriented Computation of Sandwich Estimators. *Journal of Statistical Software*, 16(9), 1–16.

How to cite this article: Schillemans T, Overman S, Fawcett P, et al. Understanding Felt Accountability. *Governance*. 2020;1–24. <https://doi.org/10.1111/gove.12547>

APPENDIX A.

TABLE A1 Descriptive statistics and correlations

	1	2	3	4	5	6	7	8	9	10	11	12
1. Expected accountability	1.00											
2. Forum legitimacy	0.51*	1.00										
3. Forum expertise	0.41*	0.45*	1.00									
4. Organizational type ^a	0.09	0.10	0.05	1.00								
5. Policy autonomy	-0.14*	-0.32*	-0.20*	0.03	1.00							
6. Financial autonomy	0.18*	0.17*	-0.04	0.20*	0.16*	1.00						
7. Contact frequency	0.18*	0.16*	0.11	-0.20*	-0.10	0.04	1.00					
8. Reporting practice	0.24*	0.24*	0.18*	-0.16*	-0.12*	0.13*	0.67*	1.00				
9. Credible sanctions rewards	0.31*	0.48*	0.22*	0.06	-0.32*	0.13*	0.18*	0.21*	1.00			
10. Location ^b	-0.05	0.06	0.04	0.00	0.02	0.07	0.08	0.07	-0.09	1.00		
11. Size (log)	0.11	0.22*	0.11	-0.05	0.03	0.23*	0.35*	0.37*	0.23*	0.15*	1.00	
12. Tenure	0.03	0.01	0.02	0.00	0.01	-0.01	-0.11	-0.10	-0.02	-0.02	-0.01	1.00
N	473	490	491	541	526	522	504	509	495	542	537	489
M	5.73	5.13	4.61		4.15	2.01	1.78	1.21	1.94		4.64	6.60
SD	1.16	1.22	1.14		0.74	0.58	0.94	0.55	0.53		2.04	3.52
Range	1-7	1-7	1-6.75	0-1	1-5	1-3	0-5.30	0-3.47	1-3.43	0-1	0-9.96	1-11
Source	FA Scale			Van Thiel, Verhoest			Brandsma,					
				2012			et al.,			2010		
							2012					

Note: most items based on existing (validated) scales; sources provided here.

^aOrganizational type: 0 = type 1 (28.9%), 1 = type 2 (70.8%).

^bLocation: 0 = in capital (54.2%), 1 = outside capital (45.8%). For the Netherlands, the political capital is used (the Hague) as the seat of the national government.

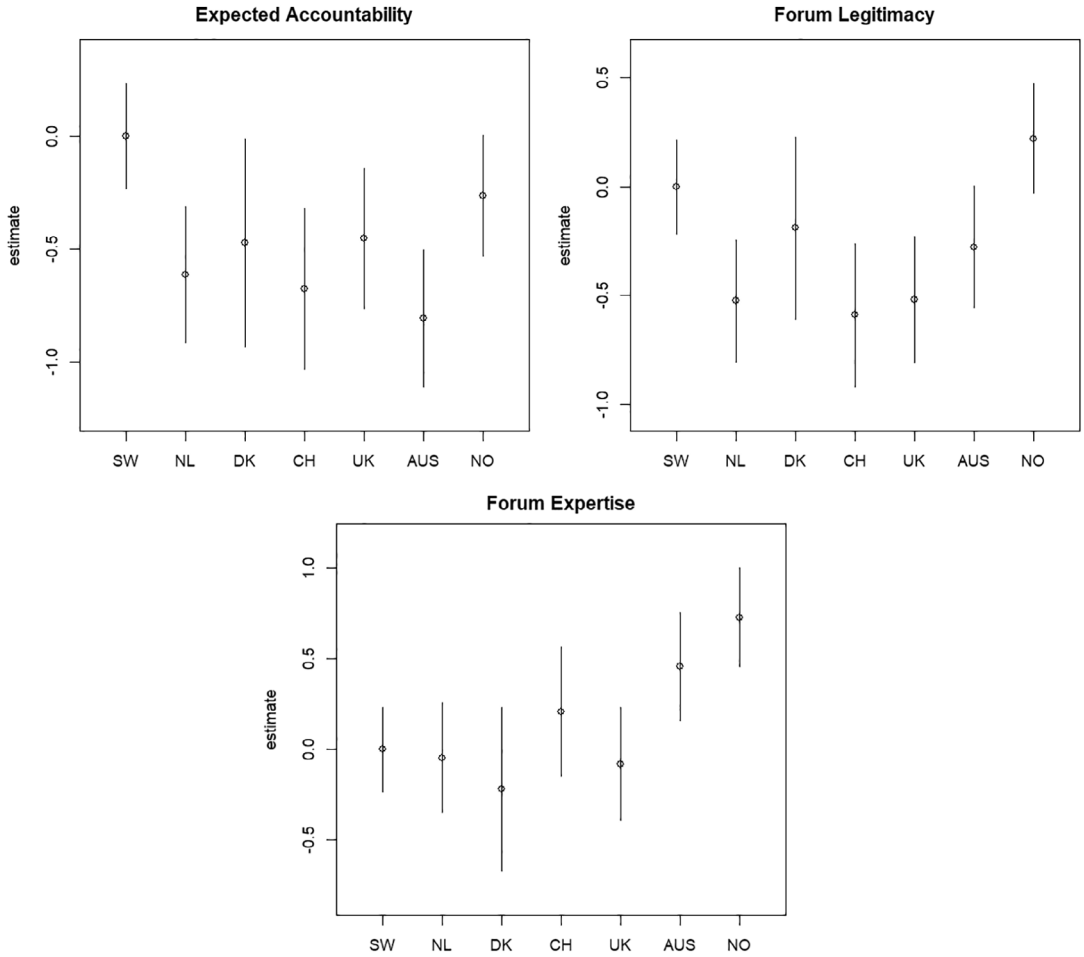
APPENDIX B.

TABLE B1 Multilevel models

	Expected Accountability		Forum Legitimacy		Forum Expertise	
	Estimate	SE	Estimate	SE	Estimate	SE
Organizational Type (ref. Type 1)						
Type 2	-0.712	0.096	-0.168	0.421	-0.305	0.484
Financial autonomy	0.125	0.327	0.196*	0.093	-0.151	0.107
Policy autonomy	-0.089	0.013	-0.346*	0.068	-0.172*	0.078
Credible sanctions and rewards	0.315*	0.252	0.648*	0.101	0.254*	0.115
Reporting practice	0.357*	0.187	0.278*	0.117	0.383*	0.134
Contact frequency	0.099	0.148	0.029	0.067	-0.033	0.077
Location Main Office	-0.038	0.450	0.193*	0.095	0.000	0.109
Size	-0.001	0.204	0.054*	0.026	0.040	0.030
Tenure	0.030	0.185	0.015	0.013	0.013	0.015
COFOG classification (ref. general public services)						
Defense	-0.275	0.171	-0.237	0.324	-0.046	0.372
Public order and safety	0.038	0.213	0.145	0.183	0.096	0.210
Economic affairs	-0.070	0.096	-0.172	0.144	-0.146	0.165
Environmental protection	-0.249	0.027	-0.018	0.250	0.249	0.286
Housing and community	0.084	0.013	0.396	0.410	0.556	0.471
Health	-0.097	0.327	0.117	0.202	0.144	0.231
Recreation Culture Religion	0.270	0.187	0.395*	0.183	0.287	0.209
Education	-0.350*	0.148	-0.008	0.168	-0.112	0.192
Social protection	-0.101	0.252	0.165	0.210	0.052	0.241
Intercept	5.151*	0.603	3.96*	0.50	4.492*	0.690
Country variance	0.064	0.254	0.110	0.3320	0.112	0.334
Residual variance	0.864	0.929	0.853	0.924	1.128	1.062
N (organizations)	456		468		469	
N (countries)	7		7		7	

Note: hierarchical linear model with random intercepts (country). Estimates and robust standard errors.

APPENDIX C: Country effects



Note: Estimations and 95% confidence intervals based on quasi standard errors (Firth, 2003).

APPENDIX D: The Felt Accountability Scale

The concept

The concept of felt accountability has been developed in psychological studies and has been used as dependent variable, independent variable or moderator in a large number of experimental studies (Aleksovska et al., 2019; Hall et al., 2017; Lerner & Tetlock 1999). Hochwarter et al. (2007) have developed a measurement instrument which has been used in many of those studies. The scale is useful and valid yet does not measure all dimensions of the overarching concept (Hall et al., 2017), most notably dimensions relating to the salience of the relationship to the “accountability forum”. Additionally, the instrument has not been used in public administration settings, where accountability relations are multidimensional (Bovens et al., 2014; Koppell, 2005; Romzek & Dubnick, 1987) and “peculiar” (Pressman & Wildavsky, 1973; Waldo, 1948). Against this background, a measurement instrument was developed, using items from the original Hochwarter scale in combination with new items for expertise and legitimacy of accountability, to be applied in public administration settings. The validated felt relational accountability measurement instrument was published in 2020 (Overman et al., 2020).

Procedure

The validated measurement was developed using commonly used recommendations in public administration research (DeVellis, 2009; Hinkin, 1998). The questionnaire contains a limited set of the original Hochwarter-items (which was established as a reliable scale with Cronbach's $\alpha = .73$) in combination with items denoting forum expertise and forum legitimacy.

A questionnaire containing the various items was administered to a total of two samples of Dutch public sector employees (from five different types of organizations). The first study was conducted among public sector employees effectively working below the national level in local and regional governments (municipalities and provinces) as well as in the police in November 2016. We fielded a second wave among employees in the national government (ministries and executive agencies) in February 2017. The first wave contains 1,983 responses (response rate: 36.3%), and the second wave contains 1,256 responses (response rate: 35.3%).

The measurement instrument was developed in a two-step strategy. First, the initial model was calibrated to the data from the first sample. Calibration involves the adaptation of the initial model based on the data in the first sample (Kline, 2015). The second sample was used to test and cross-validate the model to prevent overfitting or capitalizing on chance based on the available data in the first sample (MacCallum, Roznowski, & Necowitz, 1992).

Dimensions

The validation procedure led to a scale with three reliable subdimensions: expected accountability, forum legitimacy and forum expertise.

Expected accountability captures the original set of items on felt accountability and denotes whether decision-makers expect to be held accountable for their decisions in the future (Hochwarter et al., 2007).

Forum Legitimacy denotes whether the decision-maker finds it legitimate that (s)he will be held accountable in the future by a designated accountability forum in the future.

Forum Expertise finally denotes whether the decision-maker experiences that its accountability forum has the expertise to properly assess her decisions.

Applications

This scale has been applied in this study but also in three other empirical projects where the felt accountability of decision-makers in organizations with (some) autonomy from central government was measured: museums (Overman, 2020), transnational European organizations (Schillemans & Kremers, 2020), and independent regulators (Schillemans, Bokhorst, van Genugten, & Oude Vrielink, 2018). In all of these applications, the three items formed reliable scales. As the table below shows, there was also meaningful variance, both across the samples as well as between the three items of felt accountability. The means and N in red below are from this article, the other three from the other empirical applications of the scale in other papers.

Felt Accountability Scale: Means comparison (*n*) across empirical applications.

Felt Accountability (items)	Agencies in seven countries	Transnational organizations	Regulatory agencies	Museum
Expected accountability (Spearman Brown reliability: .71)	5.73 (473)	2.67 (43)	4.78 (75)	5.15 (96)
(1) I am held very accountable for the most important task. (2) The actors hold me accountable for all of my decisions.				
Forum legitimacy ($\alpha = .65$)	5.13 (490)	2.53 (43)	5.11 (75)	4.31 (97)
(3) When the actors change its views we just have to comply with this new reality. (4) It is a good thing, that we are ultimately accountable to the actors. (5) I am willing to work in the interest of the actors.				
Forum expertise ($\alpha = .76$)	4.61 (491)	3.42 (43)	4.24 (75)	3.79 (97)
(6) The actors apply clear/understandable standards to evaluate the most important tasks. (7) The parent department provides constructive feedback on our work. (8) Opinions from the actors are generally unambiguous. (9) The actors have sufficient substantive or technical expertise about our work to oversee /evaluate our duties.				