Comparing Governance, Agencies and Accountability in Seven Countries

CPA Survey Report

Sjors Overman, Thomas Schillemans, Per Laegreid, Paul Fawcett, Magnus Fredriksson, Martino Maggetti, Yannis Papadopoulos, Kristin Rubecksen, Lise Hellebø Rykkja, Heidi Salomonsen, Amanda Smullen, Matthew Wood

Preface

In 2017, a comparative survey was fielded by an international research consortium in seven countries: Australia, Denmark, Norway, the Netherlands, Sweden, Switzerland and the United Kingdom.

The survey is part of the broader *Calibrating Public Accountability* research project (https://accountablegovernance.sites.uu.nl/).

This report provides an overview of initial findings and some first analyses. It serves as a first peek at some of the most important findings. Further analyses and reflections will be published in time.

The research consortium responsible for the survey in the various countries represents the following universities: Utrecht University; Aarhus University; Australian National University; University of Bergen; Canberra University; University of Gothenburg; University of Lausanne; University of Sheffield.

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

We would also like to thank the many high level respondents from important public sector agencies in the seven countries for finding the time to respond to the questions in our survey. We hope we will be able to repay some of that investment by furthering our understanding of the governance of agencies, the impact of accountability on the top level of agencies and on how external accountability and organizational performance can be productively linked.

Table of Contents

1.   Aim of the Survey

This survey aimed to analyze external accountability and its impact on decision makers in public sector agencies. The focus is on accountability for policy and results; not on financial or legal accountability and decision-making.

The survey focused on the top managers in agencies. The survey mapped the formal accountability regimes of the organizations, real accountability practices and individually felt accountability. This last aspect is not only relevant on a personal level; how managers perceive accountability is obviously highly relevant for the organization as well. CEOs are (more than) likely to translate their individual perceptions of accountability in strategic decisions and specific instructions that will direct their organizations. The individual response by the CEO is in that sense highly relevant also for the organization as such.

The survey aims to innovate public administration research on public sector accountability in a number of ways.

The survey is, to begin with, the first available international research project in our field principally zooming in on 'accountability' and its effects on top level decision makers in public sector agencies. The survey maps both formal organizational accountability as well as individually felt accountability and this is, in a descriptive sense and as a mapping exercise, valuable research output as such. It will allow us to discuss similarities and differences between countries (and sectors, etc.) in accountability of agencies in a number of western democracies.

Secondly, the project combines institutional factors (such as 'type' of organization or 'task'), commonly used in public administration and political science research, with situational factors (summarized in the CPA-model; Schillemans 2016), commonly used in experimental and psychological research. This is important in a theoretical sense. It will contribute to our ability to explain (and to some degree predict) the effects of accountability mechanisms. The survey will obviously not evolve into a full-fledged and finalized theory. However, given the variety of items used, both institutional as well as situational and psychological, the relative strengths of these various factors in explaining felt accountability can be assessed.

Thirdly, the survey will trace how superordinate (mostly: departmental) accountability – on an institutional level – relates to felt accountability at the top of the organization; thus on an individual level.

Institutional accountability mechanisms should affect organizational behavior, but this relationship goes through individual perceptions and actions. If we really want to understand the impact of external accountability mechanisms on organizations, we should 'apply' Coleman's boat – suggesting that institutional factors affect individuals and that we should follow that line in our research (Coleman 1994) – at heart. This survey focuses on the first conversion from formal accountability on an institutional level to felt accountability on a personal level.

Figure 1: from institutions to individuals: [Coleman's] accountability boat [1]
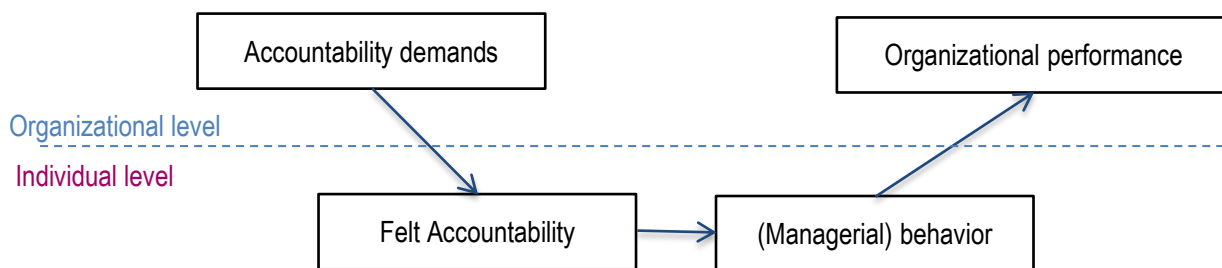
The results of the survey will (potentially) be relevant empirically as well as theoretically and practically.

As it is the first focused international accountability project in our field, the results have descriptive value on their own.

Furthermore, as we combine and test theories and insights from different academic disciplines (public administration and social psychology), and as this project is connected to a number of related (sub-) projects, it is theoretically highly innovative and ambitious.

Finally, as we start zooming in on the effects of accountability on behavior within public sector agencies, we will be able to detect (institutional or situational) factors that enhance or impede decision making in public sector agencies. This is likely to be practically relevant.

---

[1] Taken from Overman et al (2018).

2.   Country selection and responses

The survey was distributed to respondents from public sector agencies in advanced, western democracies. These countries have all gone through numerous reforms over the last decades, affecting the position of agencies, systems of accountability and public sector management in general (Christensen & Laegreid 2011; Verhoest et al 2012). Although there are important differences between the countries, of course, they have all been affected by the same reform movements. The relative cohesion of the set of countries, compared with the relatively broad reach, enables us to reach beyond obvious country-differences that would come to the fore in a broader sample, for instance in comparisons with southern or eastern European countries or in comparisons to non-western and developing countries.

The survey was thus distributed to the leadership of government agencies in 2017 in the following countries:

1. Australia
2. Denmark
3. Netherlands
4. Norway
5. Sweden
6. Switzerland
7. United Kingdom

All data in this study was collected between May and December 2017. We received a total of 661 responses, of which 499 were fully completed surveys. Table 1 specifies the number of responses per country. In most countries, a response rate of about 50% was achieved.

Table 1: Overview Responses

Country	Complete Responses	Total invitations	Response Rate
Australia	92	171	54%
Denmark	32	53	60%
The Netherlands	56	105	53%
Norway	86	169	51%
Sweden	115	241	48%
Switzerland	50	116	43%
United Kingdom	68	243	28%

We will now turn to some of the most important findings.

3. Accountability Regime: Working in a Web of Accountability

Government agencies work in networks where many other organizations may have an impact on their policies and decisions. The concept 'accountability regime' is generally understood as the sum total of all accountability mechanisms and/or relationships within which an agency operates (Scott 2000: 55; Bovens 2007; Warren 2014: 46).

The principal focus of this survey is on the dyadic accountability relationship of agencies for policy towards their parent-department. It is nevertheless important to tap into the broader accountability regime of agencies, because the dyadic accountability relationship between parent department and agency is most likely affected by the wider 'web of accountability' (Page 2006) or is even more properly conceived of as being part of that.

The broadness of the accountability regime can be measured by identifying the relevance of a number of potentially important accountability forums. We will use the well-known distinction between vertical, diagonal and horizontal accountability (Schillemans 2008) to organize the number of potentially relevant accountability forums.

Vertical accountability is directed towards hierarchical superiors, *horizontal accountability* towards non-hierarchical superiors, while *diagonal accountability* is towards persons or bodies working independently from hierarchical superiors yet appointed by them. The list of potential accountability forums is based on the section 'mechanisms' in Bovens et al's (2014) Oxford Handbook of Public Accountability. We have additionally also looked at 'diagonal' and 'horizontal' forms of accountability.

The results of the survey indicate that, in general, the parent department is the most relevant partner for the agency, which is in line with the assumptions for this study.[2] With relevant we mean that their opinions or evaluations have considerable impact on strategic decisions and how the agency performs its core task. On a 1 (not at all) – 5 (highly) scale for relevance, agency heads consider the parent department as relevant or as highly relevant (M=4.51, see Table 2). This score is significantly higher than for any other partners, including the government-as-a-whole. [3]

---

[2] The analyses in this section do not include Australia as answer categories differed.
[3] Paired sample t-tests yield values between 8.01 and 24.51, all p-values < .01.

The parent department is most important in all countries, but when countries are compared, the parent department is least important in Switzerland.[4] The government or cabinet are significantly more important in Denmark, Norway, Sweden, and the UK than in the Netherlands and Switzerland.[5] In Denmark, the department with a coordinating role regarding agencies is also significantly more relevant than in the Netherlands, Switzerland and Sweden.[6]

Table 2: Partner relevance (vertical accountability)

Partner	Mean relevance (1-5)
Parent department	4.51
Government or cabinet	4.04
Other governmental agencies	3.81
Ministries with agency coordination role	3.61
Supranational bodies, incl. EU	3.44
Other departments	3.38

Agencies also have relationships with regulatory or judicial bodies that are not directly responsible for the agency's behavior ('diagonal accountability'), as well as with partners at a non-hierarchical, horizontal level. Regarding bodies charged with oversight, agency heads see their boards as most relevant for their organizations (M=3.83). Regulators, the Supreme Court of Audit, and others are considered less relevant. We also tested the relevance of horizontal accountability forums, including advisory bodies or commissions, interest groups, unions, and media. Among this group, client and expert advisory bodies are considered most relevant (M=3.66; M=3.63), followed by interest groups and news media.

Table 3: Partner relevance (diagonal oversight bodies and horizontal relations)

	Partner	Mean relevance (1-5)
Diagonal accountability / Oversight bodies	Board	3.83
	Supreme Court of Audit	3.51
	Inspection or Regulatory Body	3.63
	Courts	3.19

---

[4] The country effect in an ANOVA yields $F_{(5,433)}=4.94; p<.01; \eta^2=.06$

[5] $F_{(5,437)}=7.741; p<.01; \eta^2=.09.$

[6] $F_{(5,404)}=3.39; p=.01; \eta^2=.05.$

	Periodic Evaluation Committee	3.08
	Ombudsman	2.79
Horizontal accounta-bility / societal partners	Advisory body representing clients	3.66
	Advisory body representing experts	3.63
	Interest Groups	3.62
	News Media	3.48
	Civil Society Organizations	3.15
	Public Sector Unions	3.19

In general, these external partners hold either similar or different views on the agency's most important task. Asking about the view of external partners, respondents tend to choose for either of the two. Only 7% of the respondents give a neutral answer to this question. The groups that remain are of equal size, with 43% of the agencies having stakeholders with slightly or strongly different views and 50% having stakeholders with fairly or strongly similar views.

The news media form a specific, non-institutional yet sometimes quite important, external forum that may have an impact on public sector agencies. To assess the importance of the media for our respondents, we asked the agency heads to rate four statements about the role of the media in their work on a 1-7 scale. These four items (based on Schillemans 2012) were:

- Stories in the news media are important and informative for our work
- I regularly adjust my daily schedule and priorities because of questions from or stories in the media
- The question "How will this be seen by the media?" is generally in the back of my mind
- Strengthening the media-reputation of our organization is important

Together, these four media-related items form a reliable scale (α=.80). This composite scale correlates with the relevance of the media as a partner, as asked elsewhere in the survey (r=,57, p<.01), which supports convergent validity of the scale. In Denmark and Norway, media are considered significantly more important than in Sweden, the Netherlands and Switzerland.[7]

---

[7] $F_{(5,435)}$=11,71; p<.01; η^2=.13.

4.   Organizational autonomy

Autonomy is one of the key issues in discussions on disaggregated governance (Christensen & Laegreid 2011; Verhoest et al 2012). Agencies are often even *defined* and *categorized* on the basis of the formal autonomy they are granted. Autonomy is also highly relevant in relation to accountability. Accountability becomes more salient when agencies are more autonomous in their decision-making which may drive calls for explanations and thus accountability. Simultaneously, however, the very autonomy they are granted may inhibit agencies from claims for accountability.

The autonomy of agencies consists of a number of sub dimensions. We have focused on two dimensions.

Financial autonomy. We assume that this probably strongly affects felt accountability, in the sense that low financial autonomy presumably leads to a stronger sense of felt accountability towards the parent department. Financial dependence in combination with policy autonomy is furthermore a recipe for conflictual relations. This issue consists of two elements: budgetary autonomy on the one hand and financial management autonomy on the other. Or, put differently, it is about how the agency acquires its budget and then about how it is allowed to spend it.

Policy autonomy is about the extent to which an agency is autonomous in making crucial decisions regarding implementation of its most important policy. One could expect, everything else being equal, an inverse relationship between policy autonomy and felt accountability, where increases in the one are accompanied by a decrease in the other.

4.1   Policy Autonomy

Agencies have a lot of formal autonomy and many can make strategic decisions without the parent department. We asked respondents about five topics (based on the cobra-survey; Verhoest et al 2012) if their agency could make decisions themselves, conjointly, or if the ministry makes decisions for them. These topics include

- goal setting
- target group identification
- choice of policy instruments
- public communication, and
- prioritization of tasks.

Parent departments have most influence on overall goals of the agency, where 24% of the respondents indicate that the parent department makes decisions, and 17% indicate that they make decisions conjointly with the parent department. However, the majority of agencies makes even these decisions about goals themselves (59%). Agencies can make most other decisions themselves: 82% of agencies make decisions regarding identification of target groups, 80% decide on policy instruments, 88% on public communication, and 88% on prioritization of tasks. All in all, it seems agencies have, or at least experience, a fair amount of autonomy regarding the implementation of policies.

Most country patterns regarding policy autonomy are largely similar. Norwegian agencies, however, have significantly more autonomy in making their own decisions than agencies in other countries.[8] One other exception is that Danish agencies and ministries make more decisions conjointly than other countries, in particular decisions about goals, target groups, and policy instruments. The Swedish make less decisions together, but instead give one of the organizations decision making power.

4.2   Financial Autonomy

We measured financial autonomy by considering the extent to which an agency was able to govern its own budget. We asked agency CEOs whether they could

- take loans for investments,
- shift between budgets for personnel and running costs within the same year,
- and shift budgets between years.

Most agencies could do all of these, with or without approval from the parent ministry, except in taking loans, which only 38% could do. In our sample, 52% could participate in legal entities, 61% can set tariffs, 73% can shift budgets between years, and 84% can shift between budgets for personnel and running costs.

When we compare the seven countries, we see that Swedish, Dutch, and Danish agencies have more freedom in taking loans than Australian, Norwegian, Swiss, and British agencies.[9]

---

[8] $F_{(6,516)}=20.72$; $p<.01$; $\eta^2=.24$.

[9] $X^2_{(12)}=165.61$; $p<.01$; Cramer's V = .40

Danish, Dutch, and Norwegian agencies have more freedom to participate in legal entities, whereas most Australian and Swedish agencies' freedoms in this respect are limited.[10]

The majority of agencies have freedom to shift between the budgets for personnel and running costs, although this freedom is less common in Swiss agencies than in the other countries.[11] Dutch, Danish, Norwegian and Swedish agencies also have more freedom to shift between budgets of different years than Swiss and British agencies.[12]

5. Formal Accountability

Accountability processes evolve in three phases (Bovens 2007). In the first phase, the information phase, the agency as the accountable actor informs its parent department about its conduct and performance. In the second phase, often coined debating phase, this account is assessed in a discussion, or meeting, or otherwise. In the third phase, the accountability forum may resort to sanctions to punish or reward the agency for its behavior. In our survey we asked a number of questions, initially based on Brandsma (2010), relating to those three phases of accountability.

5.1   Information phase

We asked questions about five ways in which agencies could provide information about their most important task to their parent department and asked for the frequency: periodic reports, formal evaluation, external audit, in writing after a formal query, informally.

We observe largely similar country patterns in the requests for information, with the exceptions of Denmark (more requests) and Switzerland (less). Some notable exceptions include the external audits, which occur relatively frequently in Sweden, Denmark and the UK, and significantly less in Dutch and Swiss agencies. A possible explanation for these findings is that in Sweden, external audits substitute other methods of providing performance information.

---

[10] $X^2(12)=181.01$; p<.01; Cramer's V = .42
[11] $X^2(12)=84.16$; p<.01; Cramer's V = .28
[12] $X^2(12)=70.48$; p<.01; Cramer's V = .26

5.2   Debating phase

The average contact frequency and the requests for information vary per country in our sample. We estimate that CEOs have on average 4.76 formal and 10.04 informal contacts with the parent department per year; in other words, CEOs and their parent department have a formal meeting about every two months, and an informal meeting about every month. Other agency employees have on average 11.56 formal and 24.51 contacts with the parent department per year. In other words, there is monthly formal interaction between agency and parent department, and bi-weekly informal interaction.

In general, Danish agencies have the most contact with their parent ministry. Swedish, Swiss, and Norwegian agencies have significantly less frequent contact with their parent ministry than agencies in most other countries.[13,14]

5.3   Consequences

The application of consequences by parent ministries is not widespread. We focused on 7 possible sanction or reward possibilities that ministries might apply to the agency. These possibilities range from public praise or reprimands to granting or reducing autonomy. Whenever there is a possibility to apply a consequence, most respondents indicate that the ministry is unlikely to do so. The full list of items is found in the appendix. The possibility and likelihood that the different sanctions are expected to be applied is correlated and can, therefore, be considered as a single scale measure of perceived possibility for sanctions.[15]

On average, Danish and Norwegian agency heads perceive their ministries to be most likely to apply consequences, and Australian, Swiss, and English agency heads perceive their ministries least likely to do so.[16]

---

[13] Contact frequency is measured with a 4-item scale with 8 answer options, ranging from never to daily. Items included a 2x2 of formal / informal contacts between agency head / other agency employees and the ministry. The answers were recoded into numbers of annual contacts (0-200) and log transformed. The items formed a reliable scale (α=.86). The variable used here is the mean log annual contact frequency.

[14] ANOVA: $F(6,500)$=15.30; $p<.01$; η^2=.18.

[15] α = .77.

[16] $F(6,491)$=21.05; $p<.01$; η^2= .26.

6.   Felt accountability

The causal link between accountability mechanisms on the one hand and behaviors and outcomes on the other entails an intermediary step. At the heart of this causal link is the individual's perception, within his institutional environment, of accountability. Individual behavior is predicated on perceptions of accountability. Agents take most decisions on the basis of what they find appropriate (Olsen 2013) and they may anticipate formal accountability in their actual decision making. In other words, felt accountability is key in our understanding of the causal link between institutional accountability arrangements and behavioral outcomes.

Felt accountability is based on the perception held by actors that they will be held accountable (Frink and Klimoski 1998; Hall, Frink, and Buckley 2015). Felt accountability denotes a relational and normative dimension: is the accountability forum authoritative or not? And it denotes a substantive core: what is it that actors believe they are accountable for?

Felt accountability consists of three dimensions (Overman et al 2018). First of all, expected accountability. Does the agency expect to be held accountable for decisions relating to its most important task? Secondly, the legitimacy of this expected accountability: does the agency accept this and find it appropriate and legitimate. And finally, the estimated expertise of the accountability forum: does, in this case, the parent department have sufficient understanding and expertise of the task at hand?

To measure expected accountability, we used a short version of the Utrecht Felt Accountability scale. The list of items is found in the appendix. The analysis shows that Swiss and Australian agencies experience, on average, a lower level of expected accountability, whereas the Scandinavian countries expect a higher level of accountability.

Expected Accountability
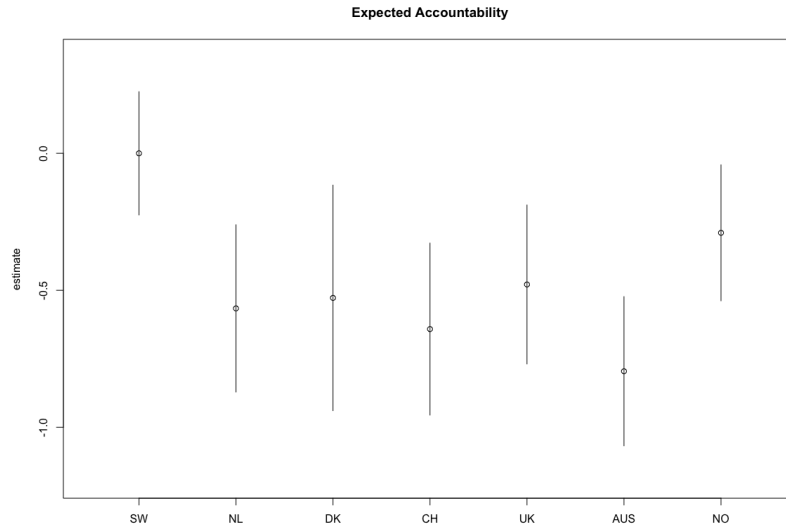
Note: country effect sizes estimates and quasi-standard errors, based on regression analyses in Table 4. In these three figures the country effects in each dimension of felt accountability are controlled for the other variables in the regression analysis in table 4.

Forum expertise is comparable among the countries, except for Norway, where expertise is considered notably higher.

Forum Expertise
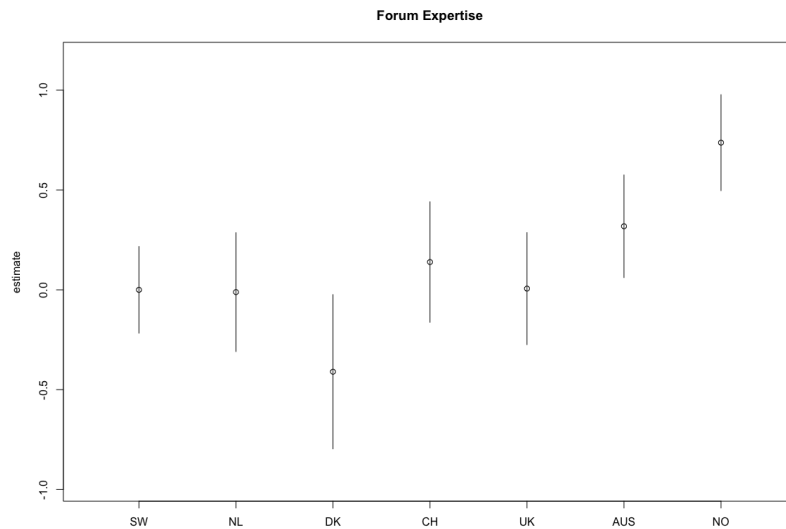
On legitimacy of the parent-department as an accountability forum, agencies in Scandinavia score high, and Switzerland and Australia score relatively low. This implies that, taken together, felt accountability

is highest in Norway and the other Scandinavian countries, and lowest in Switzerland, followed by Australia.

Forum Legitimacy
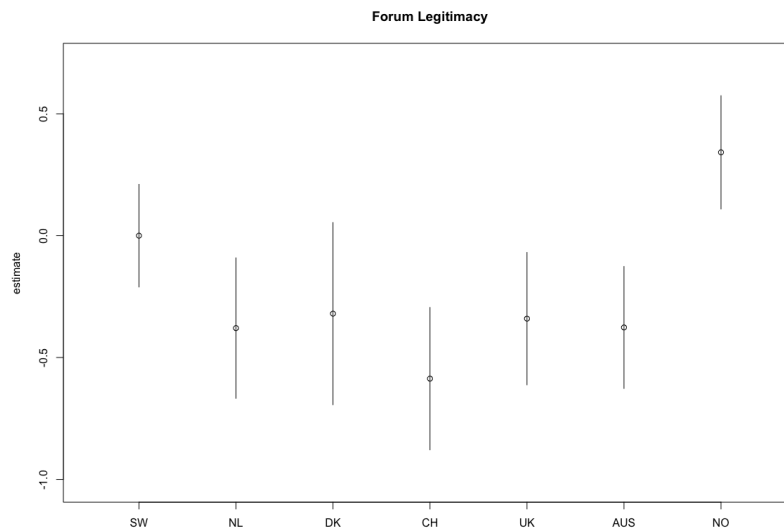
6.1 Relating external accountability to felt accountability

Now that we have seen that felt accountability differs between countries, but also between respondents obviously, we were able to analyze how the three phases of formal accountability (information phase, debating, and sanctioning) (cor)relate to felt accountability.

The frequency with which the ministry requests performance information correlates to felt accountability among agency heads in all countries, yet the effects are small (r^2 for performance information and the dimensions of felt accountability varies between .03 and .06).

The contact frequency between agencies and ministries is also correlated to felt accountability among agency heads. Contact frequency has a positive correlation with all dimensions of felt accountability. It has a stronger, more positive, effect in Sweden than in the other countries.

The likelihood of use of consequences is strongly correlated to felt accountability. In particular, the correlation between the application of consequences and forum legitimacy is high (r=.69; p<.01; r^2=.24). Also, the correlations between the application of consequences and expected accountability (r=.32; p<.01) and forum expertise (r=.22; p<.01) are relevant. These effects are comparable in all countries.

6.2 Combined effects

To further understand felt accountability we looked at the effects of a number of variables on felt accountability. We looked at the three phases of accountability in combination with the autonomy of organizations and a composite measure relating to the perceived clarity of use of targets.

A regression analysis of the separate effects of the independent variables on the dimensions of felt accountability explains a substantial fraction of the variance in felt accountability. It is unsurprising that forum expertise is least correlated with the predictors in this model. It is probably the dimension of felt accountability that is mostly affected by the capacity of the ministry in time, effort, and expertise. But there are stronger effects of the current predictors on the expectation of accountability, as well as on the forum legitimacy. Table 4 presents the results of the regression analyses. The regression analysis was conducted as one multivariate regression analysis, followed by the protected individual analyses as presented here.

Table 4: Regression analyses

	Expected Accountability		Forum Expertise		Forum Legitimacy	
Intercept	**3.66****	0.42	**2.44****	0.41	**1.97****	0.40
Information requests	**0.35****	0.13	**0.37****	0.12	**0.32****	0.12
Contact frequency	0.12	0.08	-0.03	0.07	0.06	0.07
Likelihood of use of consequences	**0.46****	0.11	**0.41****	0.10	**0.86****	0.10
Targets	0.09	0.06	**0.17****	0.05	**0.14****	0.05
Autonomy	0.17	0.10	-0.13	0.09	0.15	0.09
R^2	0.20		0.18		0.33	

Note: * p<.05; **p<.01. All analyses are controlled for country effects, which are omitted in the table, see figures.

We observe that contact frequency in the debating phase has no statistically significant effect on expected accountability when controlled for the other variables. Performance information requests have a small but statistically significant effect on all dimensions of felt accountability. Doubling the number of

requests leads to a 4-5% mean increase of expected accountability, forum expertise, and forum legitimacy.[17] We estimate that changing the likelihood of application of consequences with a single step, for example from unlikely to likely, increases the expected accountability and forum expertise with 7%, and forum legitimacy with 16%. We consider the effect of the likelihood of consequences on forum legitimacy particularly strong.

Furthermore, having measurable targets does not increase the expected accountability. These targets do have an effect on forum expertise and forum legitimacy. Increasing the autonomy, lastly, has no effect on felt accountability.

---

[17] Note that the predictor variable was log-transformed.

7.   Some concluding observations

This report has aimed to provide some first, descriptive and comparative insights in some of the results from our survey. Further analysis, thought, and reflection will need to go in to this. Additional information will also be added to the analysis, relating to organizational size for instance and formal tasks, which may have an impact on our analyses. We hope that, by providing this overview, we can stimulate some of that reflection already and provide a first overview for those interested in governance, accountability and agencies in western democracies.

In a comparative sense, it was clear that there were many *similarities* across the responses from the various countries. For instance, in all countries our respondents indicated that they take the many institutions and entities in their environment into consideration when taking decisions regarding their most important task. This suggests that they all see their broader strategic environment and web of accountability as relevant. Simultaneously, however, the strategic link with the parent department is still most important. The responses on organizational autonomy were also by and large comparable: agencies perceive to be more autonomous with regard to policy and implementation decisions than in financial terms. Patterns of contact and information provision are also fairly similar across countries. And finally, more from a research perspective, it was clear that the measures for felt accountability worked very well across the different countries.

The survey however also showed many bigger and smaller differences between and within the responses from different countries. On average, it seemed the responses from Switzerland stood out somewhat from the responses from the other countries, and could be seen as one end of the continuum on many measures, a little further than the UK and Australian responses. On the other hand, responses from Norway and Sweden were often quite comparable and may form the other end of the continuum. This makes sense theoretically and also on the basis of prior research and is suggestive of the descriptive relevance of our research findings. It is interesting for further analyses to explore these types of country differences in more detail.

We believe that felt accountability is a crucial cog in the causal chain between external accountability and organizational performance. The design and policy question related to this, is: how can felt accountability, and its various sub dimensions, be strengthened? For now we have some negative and some positive preliminary conclusions pertaining to this question.

To start with the negative conclusions: organizational autonomy, often the object of the policies relating to agencies, seems to be unrelated to felt accountability. The appropriate level of organizational autonomy of agencies may be relevant for many reasons, yet not for felt accountability.

The same goes for contact frequencies. Often, management consultants and other advisors have suggested that parent departments and agencies should engage in frequent bilateral contact in order to grease the wheels and to establish a good working relationship. Somewhat surprisingly, our results now suggest that this is perhaps not relevant when it comes to felt accountability, including the expectation to be held accountable and the legitimacy of the accountability relationship.

On the other hand, what *is* relevant for felt accountability is the frequency of information requests. By asking for information via various routes, the parent depart signals its interest in what the agency actually does, something which is not always the case, as the variance in our results testify and as has been observed before (Schillemans & Busuioc 2015). For felt accountability, the expressed interest through requests for information now seem to matter a lot.

Finally, the perceived likelihood of use of sanctions is also strongly connected to felt accountability. This may run counter to some expectations; the use or threat of negative sanctions is often thought of as undermining of a good professional working relationship. On the other hand, there are also studies suggesting that *some* threat of sanction is important, although they should hardly be used. It is this latter interpretation that is suggested by our results.

Table A: List of items felt accountability

Expected accountability (Spearman Brown reliability: .71)	I am held very accountable for our most important task.	
	The parent department holds me accountable for all of my decisions.	
Forum expertise (α = .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.	
Forum legitimacy (α = .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.	

Expected accountability: Hochwarter (2007), shortened version. Forum expertise and legitimacy:

Overman et al. (2018).

Table B: Answer percentages

Consequence	Not possible	Possible, unlikely	Possible, somewhat likely	Possible, highly likely
Adapt budget	28.1	46.5	18.8	6.7
Adapt tasks	26.7	37.4	28.3	7.6
Issue directives	32.5	32.7	26.7	8.1
Adapt autonomy	41.2	38.9	16.4	3.5
Appoint or dismiss staff (incl. executives)	60.2	29.4	7.7	2.7
Praise or reprimand	12.8	66.0	16.7	4.5
Enhance or restrict career of the executive	48.2	34.1	14.6	3.1

Appendix: further descriptives single questions

This appendix gives an overview of the mean scores on most single questions with very short explanations.

1. Personal characteristics respondents

The survey is not a personal survey; we have addressed the leadership of agencies and have asked them to respond on behalf of their organization. We nevertheless asked for some personal, professional descriptors.

To begin with, most respondents indicated that they had always had managerial positions in their current organization. This means that hardly any of the managers has emerged as a professional from the primary process.

Respondents have various educational backgrounds, with fairly equal numbers of backgrounds in social sciences & humanities, law, management & economics and natural & technical sciences.

Length of tenure however varies considerably, as the graph below shows. On the one hand we can see that many respondents have worked for longer than ten years for the agency, yet, there are also many respondents with much shorter spans of experience.

20. How long have you been working for your current agency (including legal predecessors of the agency)?

In full years
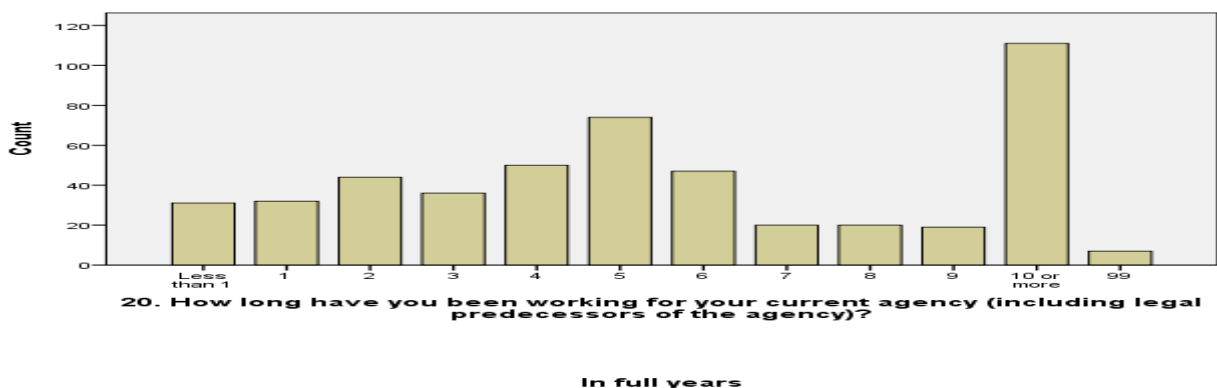
Professional antecedents also vary considerably. Most respondents have earlier experience in other public sector organizations or agencies, both in the same as in different policy areas. Half of the respondents worked for another public agency before joining their current agencies. Others come from government departments (18%), the private sector including consultancy (17%), and NGO's (10%).

23

2. Most important task

Respondents were asked to indicate which category of governmental tasks most adequately described the most important task of their organization. The categories were based on prior research, were formulate quite broadly and divided in 11 separate tasks. Even then, quite a number of respondents opted for the 'other' category, suggesting that their task is different from these standard categories. The list below is now ordered in order of prevalence, with the most prevalent option on top.

1. Supervision, regulation and control (e.g. market-regulation, inspection, rule-enforcement)
2. Providing professional public services (e.g. health, education, etc.)
3. Payment or collection of money (e.g. loans, subsidies, benefits, grants, fines)
4. Research (e.g. statistics, meteorology, public health, environment)
5. Policy task (e.g. policy advice, policy design, policy evaluation)
6. Tribunal (e.g. disciplinary boards professions; Equal treatment commission)
7. Quality assessment, certification and licensing
8. Construction and maintenance (e.g. roads, buildings, infrastructures)
9. Information and communication (e.g. collecting, filing, disclosing and disseminating information for the public)
10. Shared services (e.g. shared service tasks for the public sector such as HR or ICT)
11. Registration (e.g. drivers licenses, vehicles, property rights)

3. Mono- or co-production?

Respondents were then asked to indicate whether or not they performed this task on their own or whether they collaborated with other organizations, which could have an impact in terms of accountability. 61% of the respondents indicated that their organization performed their most important task mainly on its own while 39% of respondents indicated that they collaborated with others. When agencies collaborated with others, this was in slightly more than a third of the cases legally required, in slightly more than a third of the cases informally demanded by government and in somewhat less than a third of the cases voluntarily.

4. The Web of Accountability

In the main body of this report we explain that agencies operate in a web of accountability where our respondents find many external parties relevant in relation to their most important task, although the

parent-department is almost universally considered as the most important partner in this web. We then asked respondents to indicate whether they in general believed that the different partners held largely similar or largely different views in relation to their most important task. Responses were evenly spread; respondents indicated that views were mostly slightly different or fairly similar, although there were quite some respondents pointing at either strong differences in opinion or strong convergence in opinion.
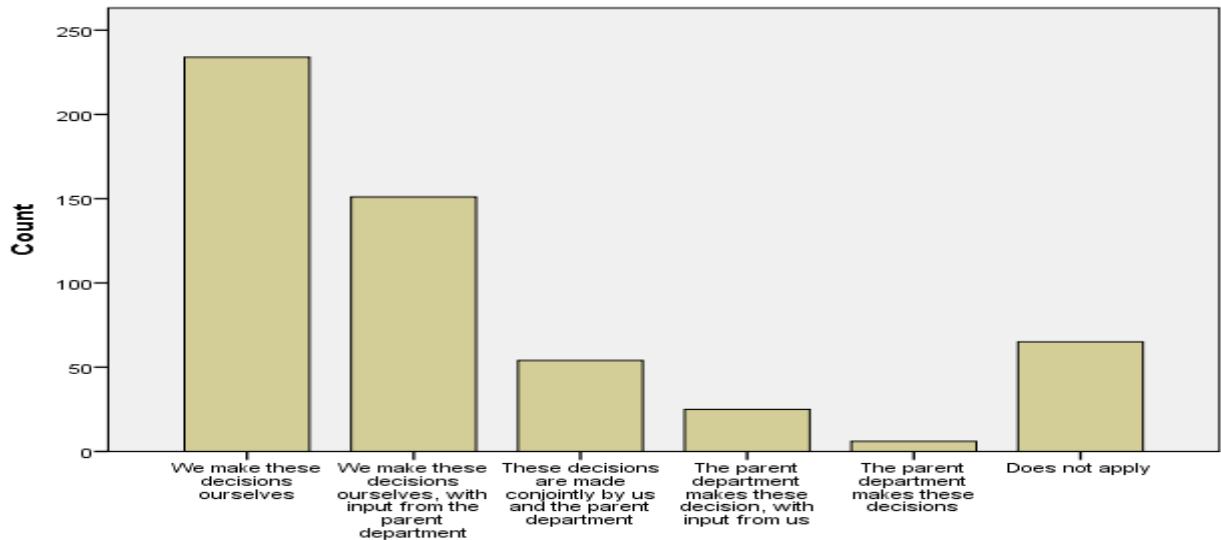
5. Policy Autonomy

Respondents have answered a number of questions relating to the policy autonomy of their organization. The overarching conclusions on autonomy are given in the main body of this descriptive report. As there was quite some variation in the responses, the graphs below provide an overview of the separate dimensions of policy autonomy.

Autonomy re overall goals …
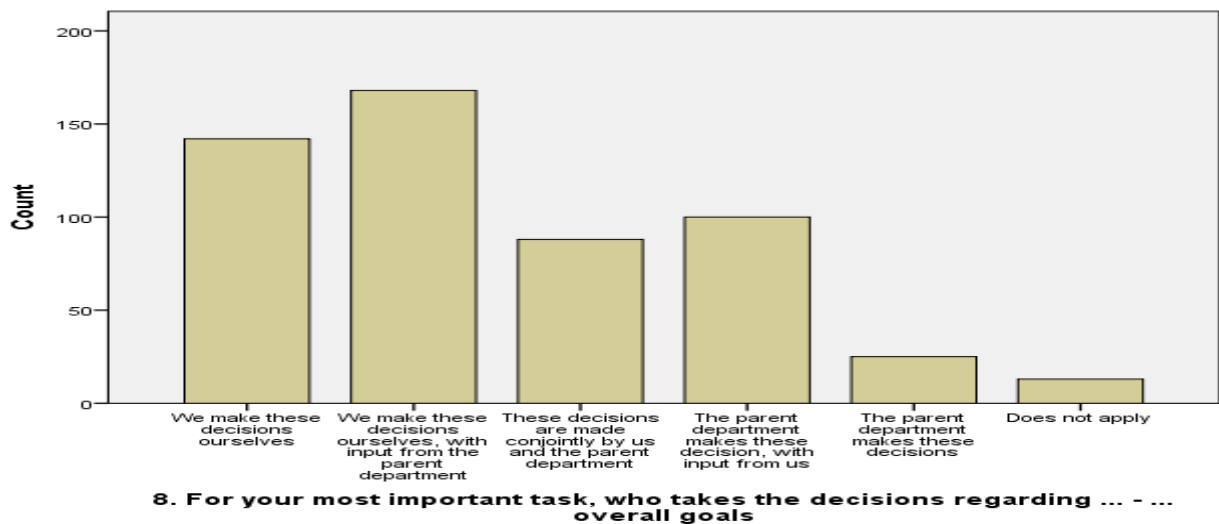
Autonomy re identification target groups…

8. For your most important task, who takes the decisions regarding ... - ... identification of target groups

Autonomy re choice between policy instruments…
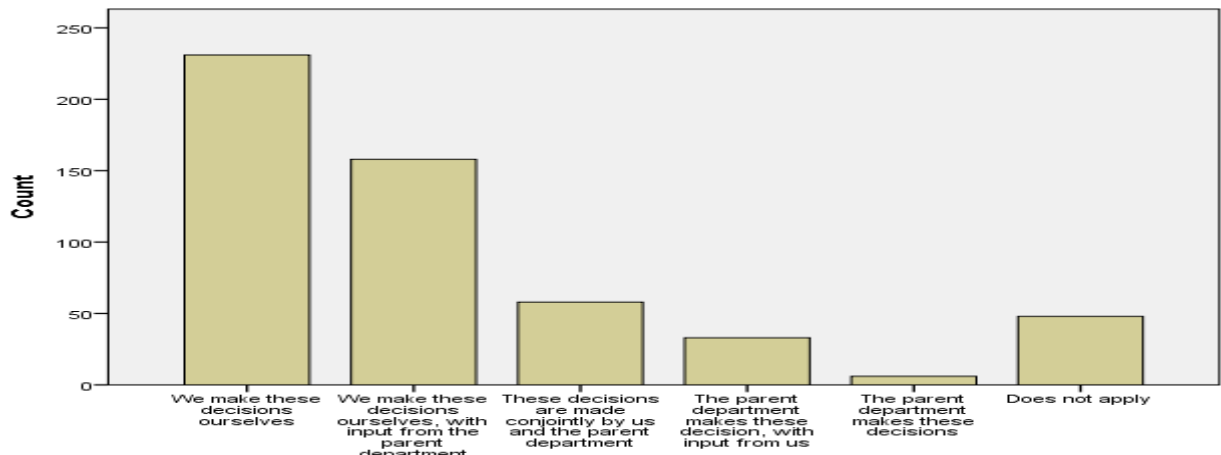
8. For your most important task, who takes the decisions regarding ... - ... choice between different policy instruments

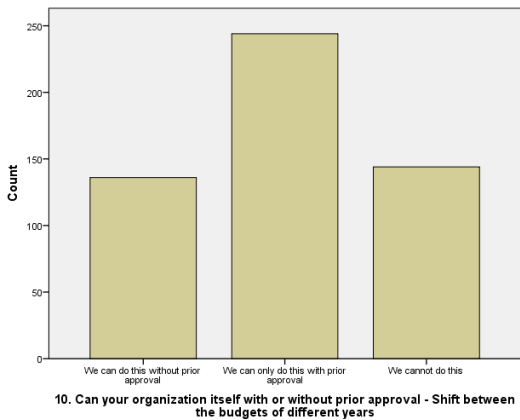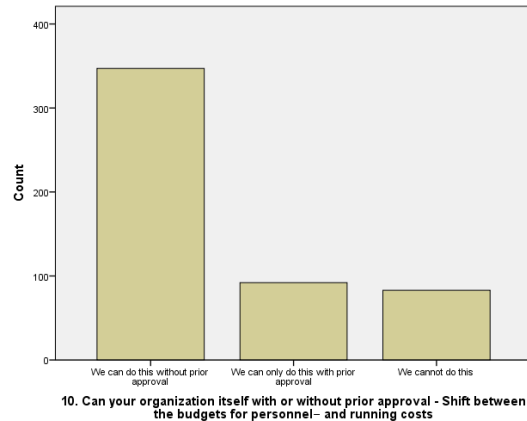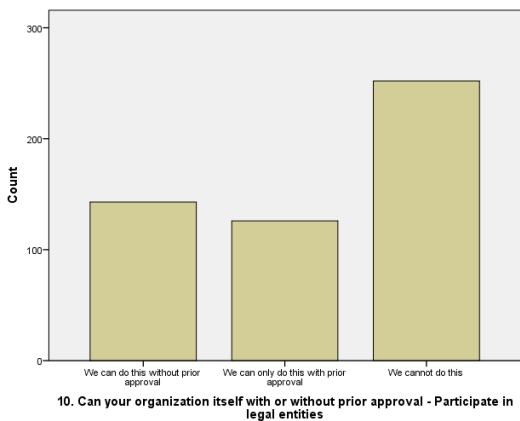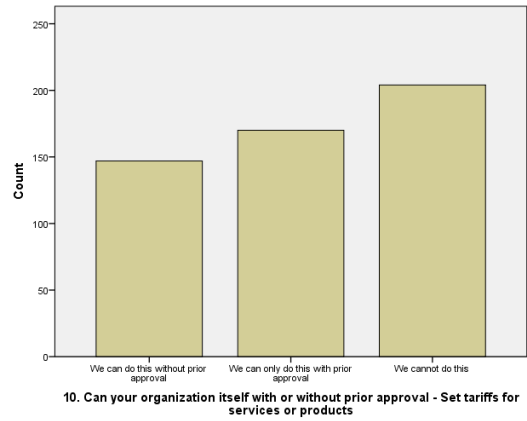*Autonomy re public communication…*

8. For your most important task, who takes the decisions regarding ... - ... public communication about the policy
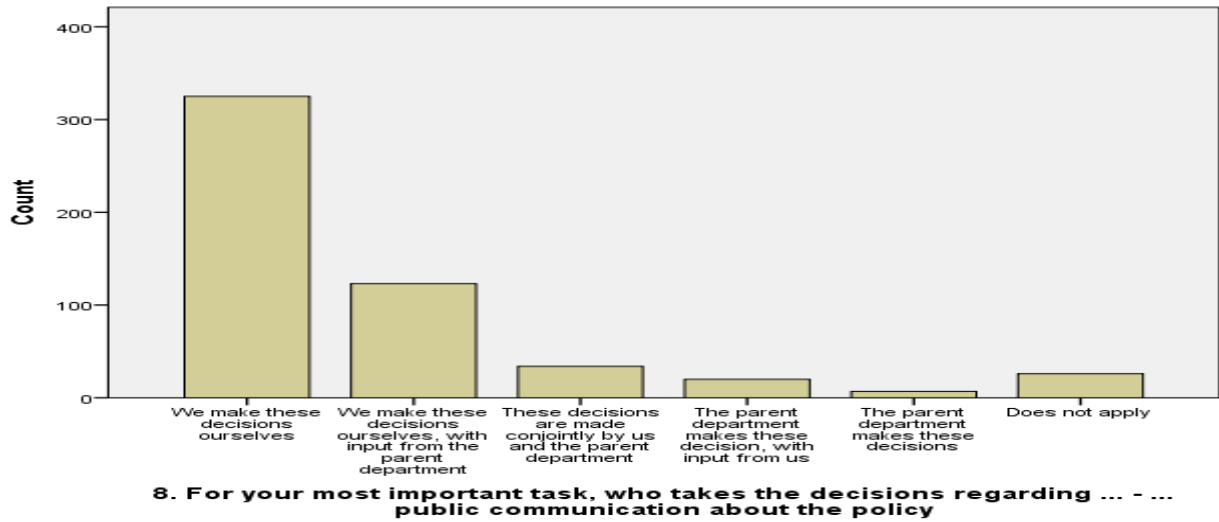
Autonomy re prioritization relative to other tasks…

8. For your most important task, who takes the decisions regarding ... - ... prioritization relative to other tasks
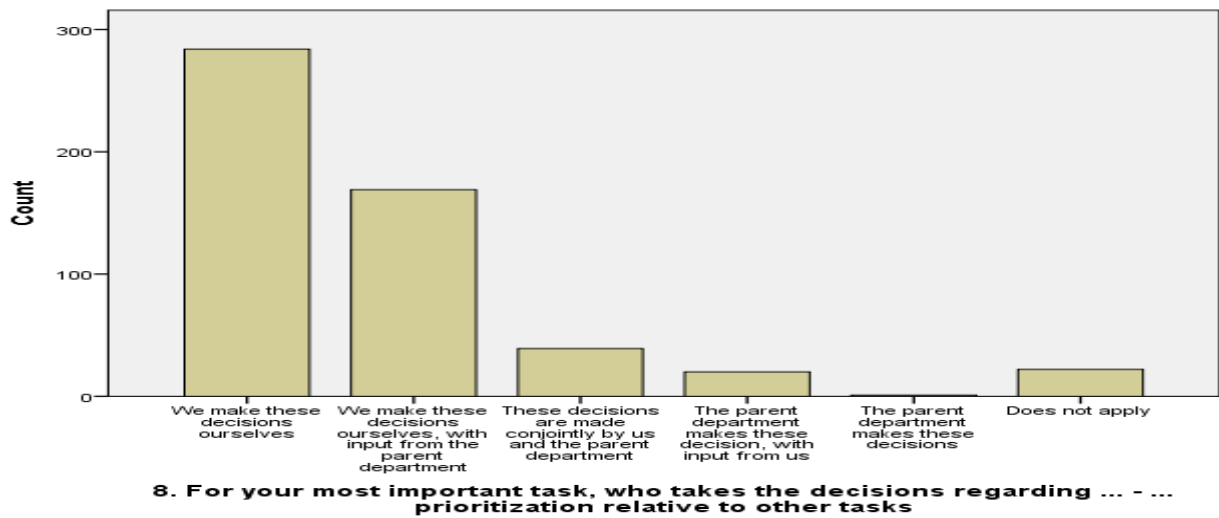
6. Financial Autonomy

Respondents have also answered a number of questions relating to the financial autonomy of their organization. The overarching conclusions on autonomy are given in the main body of this descriptive report. As there was quite some variation in the responses, the graphs below provide an overview of the separate dimensions of financial autonomy. Compared to policy autonomy the overarching conclusion is that agencies are more autonomous regarding policy choices than in terms of financial autonomy. Also, the variations in the sub-dimensions of financial autonomy, are more pronounced.

27

10. Can your organization itself with or without prior approval - Take loans for investments

10. Can your organization itself with or without prior approval - Set tariffs for services or products

10. Can your organization itself with or without prior approval - Participate in legal entities

10. Can your organization itself with or without prior approval - Shift between the budgets for personnel– and running costs

10. Can your organization itself with or without prior approval - Shift between the budgets of different years

7. Contact with parent-department

Our respondents almost universally indicated that their parent department(s) was their most important external stakeholder. The relationship with the parent department is generally formalized. Almost all respondents (more than 91%) indicated that there was at least once a year a formal meeting where the CEO discussed the most important task with the parent department. Relationships are also

institutionalized in organizations, often via designated units. Respondents indicate that it is most common that both parent department and agency work with designated units for their professional relationship (39%), followed by the case that it is only the parent department (30%), neither (29%), or only the agency (2%).
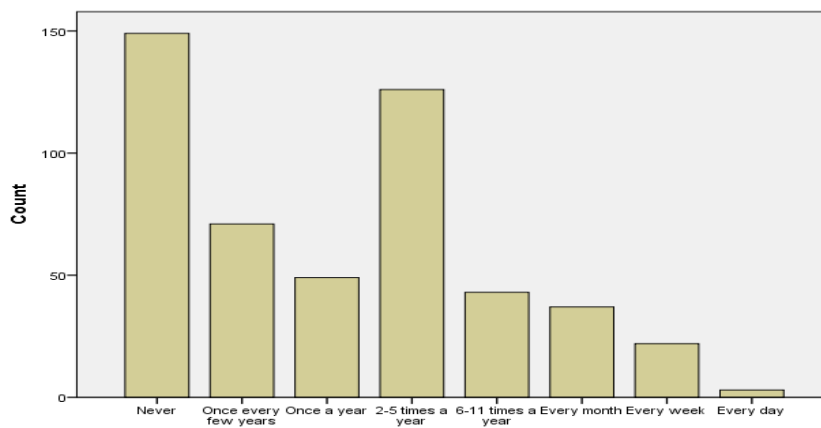
8. *Information Phase of Accountability*

We asked our respondents to indicate how often they submitted information to their parent department about their most important task via a number of potential routes. The overall findings here have been discussed in the main body of this report. We limit ourselves here to the graphs related to the various means of information provision. The graphs clearly show variations between the different modes in which information can be provided to parent departments, which was expected.
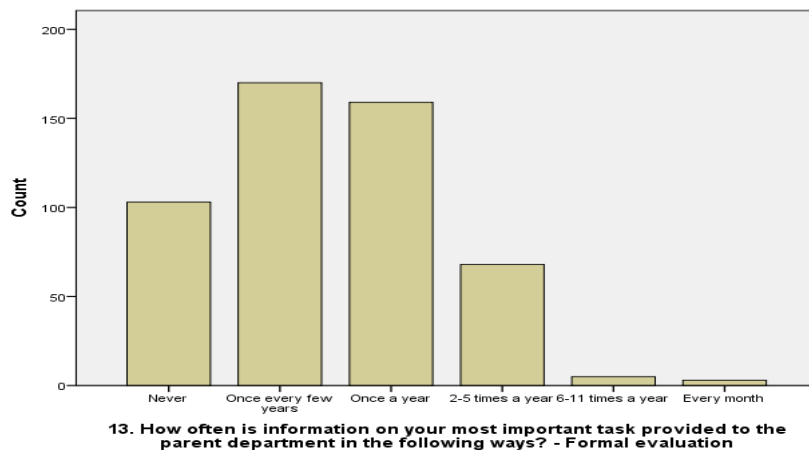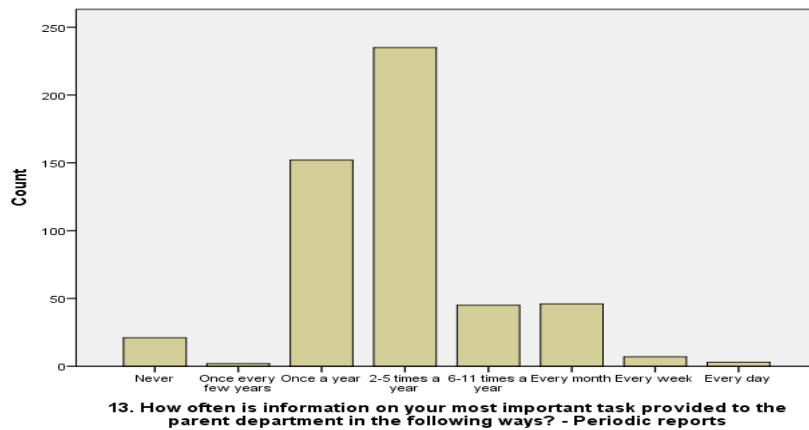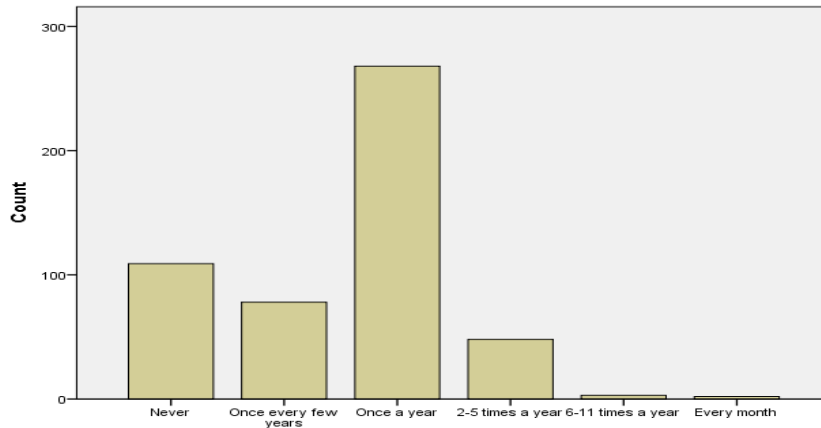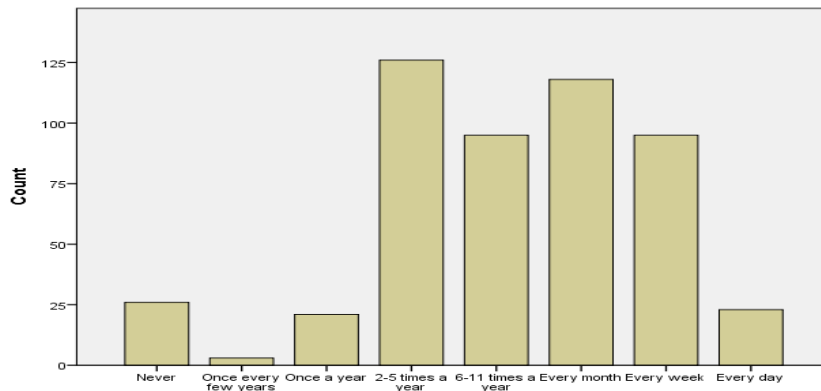

13. How often is information on your most important task provided to the parent department in the following ways? - Periodic reports

13. How often is information on your most important task provided to the parent department in the following ways? - Formal evaluation

13. How often is information on your most important task provided to the parent department in the following ways? - External audit


13. How often is information on your most important task provided to the parent department in the following ways? - In writing after a formal query
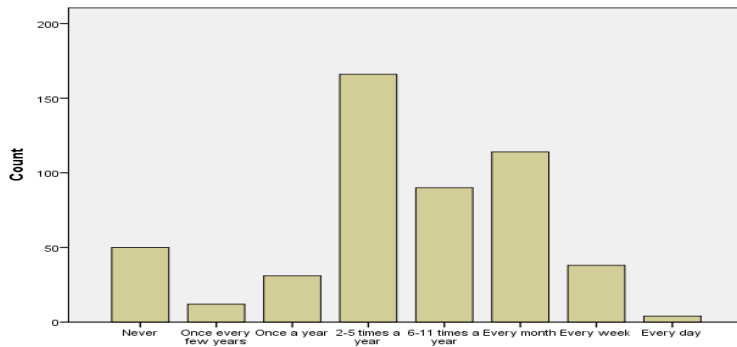

13. How often is information on your most important task provided to the parent department in the following ways? - Informally

9. Debating Phase of Accountability

We asked our respondents to indicate how often their most important task was discussed with the parent department, both at the highest level as well as at less senior levels and in formal meetings and informally. The overall findings here have been discussed in the main body of this report. We limit ourselves here to the graphs related to the individual questions asked.
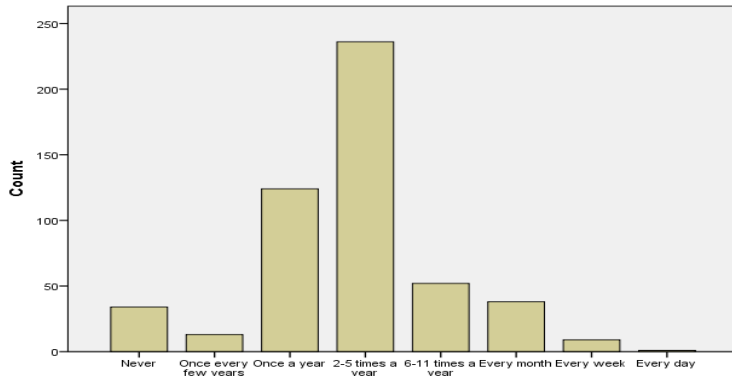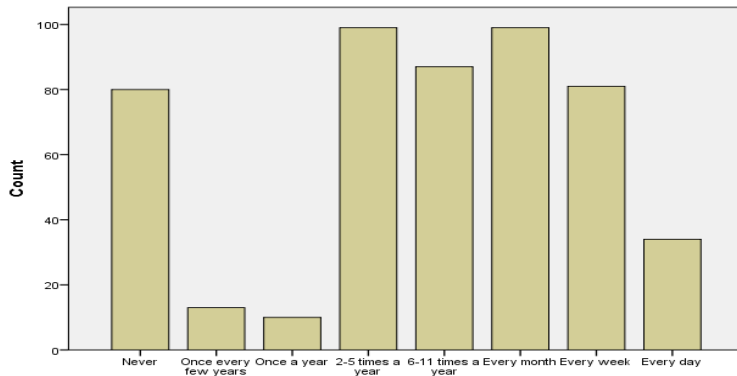
30

14. How often is your most important task discussed with representatives from the parent department? - By the CEO in a formal meeting

14. How often is your most important task discussed with representatives from the parent department? - By the CEO informally
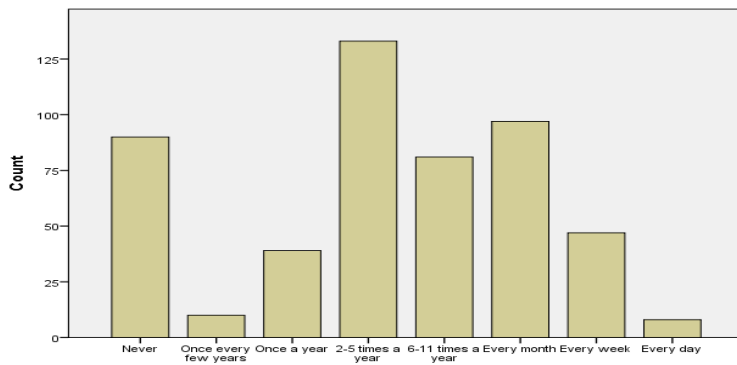
14. How often is your most important task discussed with representatives from the parent department? - By others at lower levels in the organization in formal meetings

14. How often is your most important task discussed with representatives from the parent department? - By others at lower levels in the organization informally

10. Sanctions Phase of Accountability

We finally asked respondents about the prevalence of potential sanctions that parent departments could use when they were not satisfied by how the most important task was performed. It was in fact a double question, where respondents were asked to indicate whether parent departments could or could not use various potential sanctions and, if they could, whether it was (highly) likely or not that they would actually use them. As in other research on sanctions: the overall finding is that sanctions are rarely used but, as explained in the main body, the existence and perceived realism of sanctions does in fact have an impact. Below we provide *additive* perceptual scores on this question for all sanctions surveyed. The red bar is the line between existence and non-existence of a given sanction. Where the overarching finding is that sanctions are believed to be seldom applied, we can also see that the prevalence of specific sanctions in the first place (red bar) is quite different for different potential sanctions.

	Increase / decrease budget	Add / reduce tasks	Issue binding directives	Grant or reduce your operational autonomy	Appoint or dismiss staff	Public praise or a public reprimand	Enhance or constrict future career choices
Yes + highly likely	6,4	7,4	7,8	3,4	2,6	4,2	2,8
Yes + somewhat likely	24,3	34,7	33,3	19,1	10,0	19,8	15,8
Yes + unlikely	68,6	70,7	64,5	56,4	38,2	81,4	46,1
No it can not	95,4	96,4	95,6	95,8	96,2	93,4	89,0
I do not know	100,0	100,0	100,0	100,0	100,0	100,0	100,0

11. Strategic external orientation

The survey ended with some questions tapping into the managerial perspective on how tasks are performed. These findings should be used with great caution: they are based on limited sets of questions and are difficult to relate to other dimensions of the survey, because of the risk of single source bias. These findings should thus be used with great care and interpreted hesitantly. Having said that, it is nevertheless interesting to see what the respondents have indicated.

We start with what we coined a 'strategic external orientation', representing CEO's and agencies actively reaching out to their environment for support and also to demonstrate what the agency does. Two statements were used, the first relating to whether it is important to make performance results publicly available to all stakeholders and the second about regularly meeting with societal stakeholders at a strategic level. Both questions signify a 'horizontal orientation' on the policy network beyond the hierarchical relationship to central government. In earlier questions on the relevance of all sorts of stakeholders, it was already clear that our respondents take the position of many significant others into consideration when taking decisions about their most important task. They now also to a very large degree suggest that it is important to inform stakeholders publicly about results and, although to a somewhat less overwhelming degree, to meet with stakeholders at a strategic level.

12. Compliance focus

Theoretically, operational and managerial decisions can be judged on the basis of two types of criteria (Schillemans 2016). The first could be termed accuracy or compliance focus. With this we mean that decisions can be judged by pre-defined criteria. In the public sector, two of those pre-defined criteria could be the traditional focus on norms for legality and the more modern focus on measurable quality and production targets. It is very clear that our respondents find both types of criteria highly relevant for their most important task. The 'support' for the relevance of norms for legality is extremely high, opinions on the relevance of measurable targets vary somewhat more and are less extreme. Yet on average it is safe to say that respondents find both criteria of the utmost relevance.

13. Reflection focus

The second type of criterion relates to the information- and reflection-process feeding into decisions. This is based on research in psychology and on decision-making, where it is claimed that it is difficult to assess whether decisions are right or not substantively, yet that it is possible to look at the quality of decision making processes. Two dimensions have been surveyed: the extent to which our respondents believe it is important to understand different sides of issues and the extent to which time is spent to collect and analyze information from various sources. Here the findings are more mixed and less outspoken than for the preceding criterion, as can be seen in the graphs below.
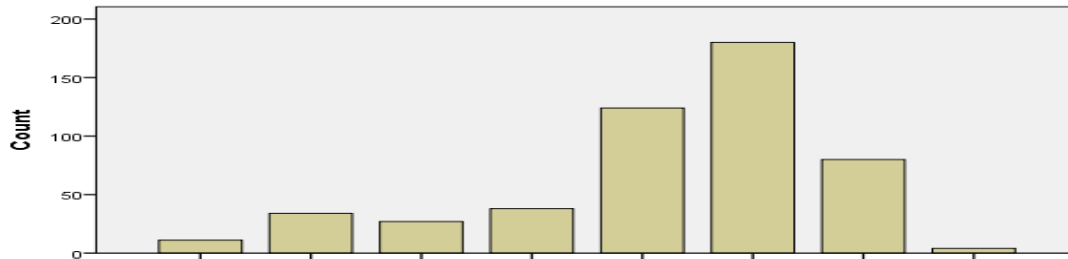
18. Please indicate to what extent you agree with the statements below:

For key decisions on our most important task ... - ... we spend much time collecting and analyzing information from various sources
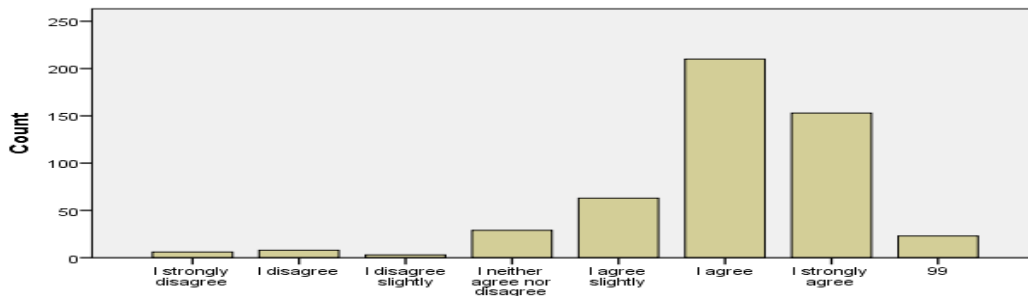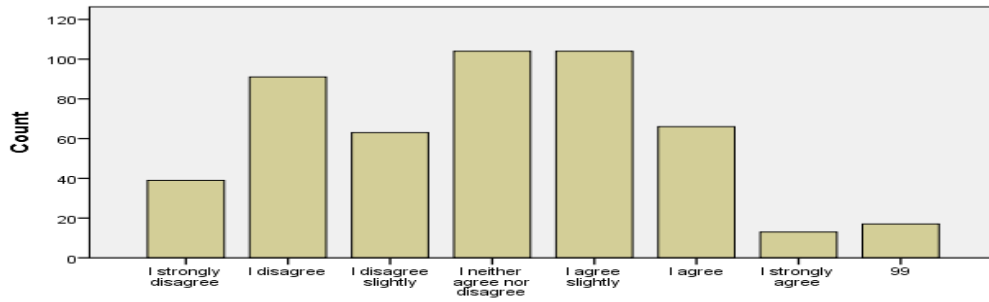

18. Please indicate to what extent you agree with the statements below:

For key decisions on our most important task ... - ... it is imperative that we understand different sides to issues

14. Glimpses of Impressions Management

Finally, the survey so far assumes that there can be a positive relationship between external accountability and managerial decisions and organizational performance. There is good reason to assert such a connection. However, it is also well-known that accountability may invoke all sorts of less salutary responses. One common response could be that accountability leads to impressions management: when one is asked to explain one's conduct and behavior, it is only natural to present a positive image of one's self. We asked two questions relating to impressions management. Responses varied widely, as the two last graphs below signify. These types of questions are always vulnerable to social desirability bias, so they should be understood and interpreted with care. They do however suggest that elements of impressions management are manifest, which was only to be expected, which seems to be stronger in terms of pre-empting potential later criticisms than in terms of collecting unnecessary information.

18. Please indicate to what extent you agree with the statements below:

For key decisions on our most important task ... - ... we collect more information than we really need to cover all the bases
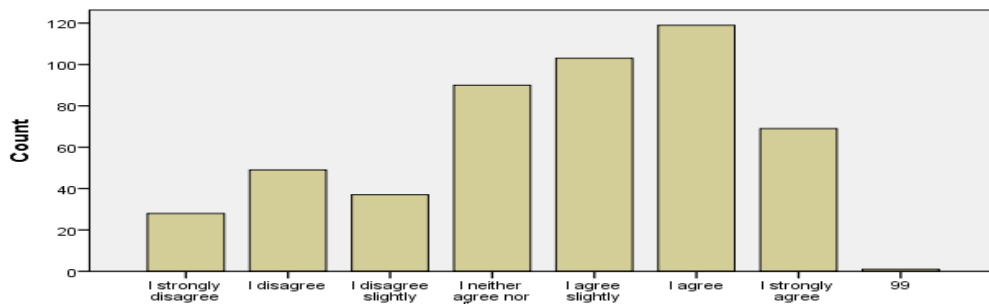
18. Please indicate to what extent you agree with the statements below:

For key decisions on our most important task ... - ... we also collect information to show that we do a good job in case something goes wrong later on

References

Bovens, M. (2007), 2007. "Analysing and Assessing Accountability: A Conceptual Framework." *European Law Journal* 13 (4):447–468.

Bovens, Mark, Robert E. Goodin, and Thomas Schillemans. 2014. *The Oxford Handbook of Public Accountability*. Oxford: Oxford University Press.

Brandsma, G. J. (2010). *Backstage Europe: Comitology, accountability and democracy in the European Union*. Utrecht University.

Christensen, T., and P. Lægreid. 2011. *The Ashgate Research Companion to New Public Management*. Farnham: Ashgate.

Coleman, J. S. (1994). *Foundations of social theory*. Harvard university press.

Frink, Dwight D., and Richard J. Klimoski. 1998. "Toward a Theory of Accountability in Organizations and Human Resource Management." In *Research in Personnel and Human Resources Management*, Vol. 16, 1–51. US: Elsevier Science/JAI Press.

Hall, A. T., Frink, D. D., & Buckley, M. R. (2015). An accountability account: A review and synthesis of the theoretical and empirical research on felt accountability. *Journal of Organizational Behavior*.

Hochwarter, Wayne A., Gerald R. Ferris, Mark B. Gavin, Pamela L. Perrewé, Angela T. Hall, and Dwight D. Frink. 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):226–39. https://doi.org/10.1016/j.obhdp.2006.09.003.

Olsen, J. P. (2013). The institutional basis of democratic accountability. *West European Politics*, 36(3), 447-473.

Overman, S., S. Grimmelikhuijsen, and T. Schillemans (2018), Measuring felt accountability in the public sector: expectation, legitimacy and expertise. *Forthcoming.*

Page, S. (2006). The web of managerial accountability: The impact of reinventing government. *Administration & Society*, *38*(2), 166-197.

Schillemans, T. (2008). Accountability in the shadow of hierarchy: The horizontal accountability of agencies. *Public Organization Review*, *8*(2), 175.

Schillemans, T. (2012). *Mediatization of public services: How organizations adapt to news media*. Frankfurt am Main: Peter Lang.

Schillemans, T. (2016), Calibrating Public Sector Accountability: Translating experimental findings to public sector accountability, *Public Management Review*. 18(9), pp. 1400-1420.

Schillemans, T., & Busuioc, M. (2015). Predicting public sector accountability: From agency drift to forum drift. *Journal of Public Administration Research and Theory*, 25(1), 191-215.

Scott, C. (2000). Accountability in the regulatory state. *Journal of law and society*, 27(1), 38-60.

Verhoest, K., Van Thiel, S., Bouckaert, G., Lægreid, P., & Van Thiel, S. (Eds.). (2012). *Government agencies: practices and lessons from 30 countries*. Springer.

Warren, M. E. (2014). Accountability and democracy (pp. 39-54). In: Bovens, Mark, Robert E. Goodin, and Thomas Schillemans. 2014. *The Oxford Handbook of Public Accountability*. Oxford: Oxford University Press.