

TOMI RAJALA

Use and Non-Use of Performance Information in the Public Sector

Scratching beneath the surface

TOMI RAJALA

Use and Non-Use of Performance Information
in the Public Sector
Scratching beneath the surface

ACADEMIC DISSERTATION

To be presented, with the permission of
the Faculty of Management and Business
of Tampere University,

for public discussion in the Paavo Koli auditorium of the Pinni A building,
Kanslerinrinne 1, Tampere,
on 11 December 2020, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Management and Business
Finland

*Responsible
supervisor
and Custos* Professor Jarmo Vakkuri
Tampere University
Finland

Pre-examiners Wouter Van Dooren Åge Johnsen
University of Antwerp Oslo Metropolitan University
Belgium Norway

Opponent Wouter Van Dooren
University of Antwerp
Belgium

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Copyright ©2020 author

Cover design: Roihu Inc.

ISBN 978-952-03-1769-0 (print)
ISBN 978-952-03-1770-6 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
<http://urn.fi/URN:ISBN:978-952-03-1770-6>

PunaMusta Oy – Yliopistopaino
Vantaa 2020

Acknowledgements

Stories always have a starting point. The journey toward completion of this dissertation began in 2014 while I was at Statistics Finland. My phone rang, and I walked down the corridor to an empty conference room where time was frozen in the furniture that reflected the stylistic preferences of previous decades. I remember the situation, including the furniture, vividly, as it was a significant turning point in my life. That phone call initiated the dissertation process and my life as researcher. I would like to thank my supervisor, Jarmo Vakkuri, for that rather surprising call and the numerous conversations that followed, without which there would be no dissertation, and for lending me so many fascinating books that I may have sometimes forgotten to return in time. Thanks are also given to my pre-examiners, Wouter Van Dooren and Åge Johnsen, for their helpful comments that significantly helped to propel the work forward toward the end. Special thanks to Wouter for agreeing to act as an opponent in my thesis defense.

As a researcher, it became evident that writing a dissertation was a long and winding process. It was not always enjoyable, but sometimes the most painful moments afford the best life lessons. During moments of pain, individuals face questions such as “Who am I” and “What do I want to become?” When recalling warmer memories, the road to completion of my dissertation involved many “sunny moments” and beautiful sunsets as I tended to work in the evenings and night shifts. At best, conducting research transported me into a flow-like state where my excitement was mixed with an insatiable curiosity about the way in which the world worked. On some days, I rushed to my computer because I was so motivated to write down new thoughts. Now, when reflecting on the process, it seems that my writing operated as a gateway to humanity and my inner world. It invariably involved a rollercoaster of emotions. Ultimately, I learned the most about myself and my limits.

My new life as a researcher included meeting interesting members of the work community who helped me to familiarize myself with academic life and its practices. Various members of this community deserve my gratitude; they include Anna, Lotta-Maria, Eija, Lasse, Jaakko, Harri, Ulriika, Kirsi, Kristiina, and Anniina, and the friendly faces at the Faculty of Business and Management, such as Selene, Heli and Aino. Harri is thanked for organizing a dialogue project, which was both fun and productive; house band, JKK, is thanked for its seamless groove and Lotta-Maria for our mutual hobby of observing the behavior of local politicians in their natural surroundings.

My scientific work would not have been half as good if I had not met such fine and interesting people. It was especially rewarding when Mikko, Joni, Petri and their families

pulled me out from the world of research to do something completely different from time to time; it promoted my well-being. I would also like to thank Paavo, Jussi, and Eero for long tennis court rallies over the years. Recognition must be given to Netu for its considerable team spirit over the years. I received great tips and enjoyed many jokes about municipalities from Erki, Hööri, Setä, Ylli, Jesse, and Eino. Having a laugh every now and then was extremely relaxing.

Maria, who walked with me in my journey to adulthood and academic life, deserves similar acknowledgement. Besides the importance of peer support, I learned about the importance of holistic well-being from Maria. I probably owe her forever as it is difficult to offer equally valuable advice in return. I hope she does not mind. Thanks to Viljo for his charming friendship, Hugo for reminding me of the value of time spent together, and Kirsti for her offer of a fantastic place to rest in beautiful scenery near Lake Näsijärvi. I would also like to thank Hessu, Jutta, and the Kirjavainen family for breaking bread with me on many occasions. As one cannot live solely on scientific knowledge, it was wonderful to eat delicious food in good company.

My older sister paved the way for my attendance at university, and I inherited my interest in municipal affairs from my grandfather, Arvo. I am very grateful for my sister for helping me out with my Bachelor's and Master's theses and for giving me invaluable tips on life at university. In many ways, she was a remarkable role model who showed me how to take greater responsibilities. It was a pleasure to follow the growth of Ella and Linnea and see how their board game skills developed! Karri needs to be thanked for his tips on state-of-art technology as they inspired me to adopt productive software applications for my academic work, and Kamu is thanked for waving his tail every time we met.

I am grateful to my mother who provided me with numerous opportunities to exercise my creativity. I do not take for granted that I had multiple opportunities to play different instruments at home because my mother encouraged me to do so. The pre-examiners stated that a strong point of my work was its creative approach. There is no doubt that my creativity was sourced in my musical practice sessions at home, years before I even dreamed of writing a dissertation. I would also like to thank my father for always helping me out in times of trouble and for introducing me to different hobbies when I was young. Taking part in those hobbies taught me invaluable lessons about team spirit and the “ups” and “downs” of life.

Inkeri's family members have my gratitude for taking me into their lives. Lastly, I would like to thank Inkeri; your eyes are the sunshine in my life early in the morning and late at night. The warmth of your smile supported me throughout the writing process. Without doubt, your positivity is rare and intoxicating and is immeasurably invaluable to me.

As a final remark, I would like to dedicate this dissertation to those who are willing to learn lessons about the world from it. This work was written for you, wherever you may roam.

In Tampere,

Tomi Rajala

Esipuhe

Tarinat alkavat aina jostakin. Tämän väitöskirjan tarina alkoi vuonna 2014 Tilastokeskuksessa. Puhelimeni soi ja kävelin käytävältä tyhjillään olevaan neuvotteluhuoneeseen, jossa aika oli pysähtynyt kalusteiden perusteella menneille vuosikymmenille. Muistan tilanteen ehkä liiankin elävästi huonekaluja myöten, sillä se oli iso käännekohta elämässäni. Tuo puhelu aloitti prosessin kohti tämän väitöskirjan valmistumista. Tuon puhelun mukana tuli tutkijan elämä, josta olen sittemmin huomannut pitäväni kovasti.

Aikoinaan tuon puhelun minulle soitti väitöskirjani ohjaaja Jarmo Vakkuri. Kiitokset tuosta yllättävästä puhelusta ja sen jälkeisistä lukuisista keskusteluista! Ilman niitä ei oltaisi tässä hetkessä laittamassa viimeistä pistettä väitöskirjaani. Ja kiitokset myös kiehtovista kirjoista, jotka olen saattanut toisinaan unohtaa palauttaa ajallaan takaisin. Väitöskirjan valmistumisesta on syytä lisäksi kiittää esitarkastajiani Wouter Van Doorenia ja Åge Johnsenia, jotka ovat avittaneet merkittävästi työn eteenpäin viemistä loppumetreillä. Erityiskiitokset Wouterille suostumuksesta toimia vastaväittäjänä.

Tutkijan elämän mukana tuli myös kiinnostava työyhteisö, joka on auttanut minua perehtymään monipuolisesti akateemisen maailman saloihin. Kiitokset tästä kuuluvat Annalle, Lotta-Marialle, Eijalle, Lasselle, Jan-Erikille, Jaakolle, Harrille, Ulriikalle, Kirsille, Kristiinalle, Anniinalle sekä muille ystävällisille kasvoille johtamisen ja talouden tiedekunnan henkilöstössä, kuten Selenelle, Helille ja Ainolle. Olen oppinut teiltä paljon tällä matkalla ja toivon mukaan tulen vielä oppimaan paljon. Erityiskiitokset Harrille mukavasta ja naurun täyteisestä dialogihankkeesta ja JKK house bändille svengaavasta menosta. Lotta-Mariaa täytyy kiittää yhteisestä havainnointiharrastuksesta, jossa on bongailtu kunnanvaltuutettujen touhuja.

Mitä tulee itse tutkijan työhön, niin ainakin väitöskirjan tekeminen on ollut pitkä ja tuulinen matka. Aina ei ole ollut kivaa, mutta toisinaan kivuliaimmat hetket ovat parhaimpia oppitunteja elämästä. Noina hetkinä usein joutuu kysymään, kuka oikein on ja mitä elämältään haluaa. Käännettäessä katse mukavampiin muistoihin matkan varrelta tulvii mieleen aurinkoisia hetkiä ja hienoja auringonlaskuja ikuiselle ilta- ja yövuorolaiselle. Parhaimmillaan tutkimuksen tekeminen on ollut tiedehurmasta, jossa sekoittuu pienen ihmisen innostus ja kyltymätön uteliaisuus maailmanmenoa kohtaan. Olenpa joskus juossut töiden pariin ajatusten pakottaessa nopeasti kirjoittamaan. Nyt, kun katson ajassa taaksepäin prosessia kokonaisuutena, näyttäätyy kirjoitustyö tunteiden vuoristoratana, ikkunana ihmisyyteen ja peilinä ennen kaikkea omaan sisimpään. Pitkällä tutkimusmatkallani eniten taisin lopulta oppia itsestäni ja omista rajoistani.

Aikojen saatossa olen huomannut, että tieteellinen työni ei olisi puoliksikaan niin hyvää, ellei minulla olisi ollut ympärilläni niin hienoja ja omalaatuisia ihmisiä. Erityisen mukavaa on ollut, että Mikko, Joni ja Petri perheineen ovat aika ajoin vetäneet minut pois tutkimuksen syövereistä aivan muihin puuhiin. Se on ollut tervettä. Vuosikausia kestäneistä palloralleista on syytä muistaa kiittää Paavoja, Jussia sekä Eeroa. Netulle on syytä antaa tunnustusta hienosta joukkuehengestä vuosien varrelta. Parhaat kuntanaurut ja -vinkit on jo pitkään toimittanut Erki, Hööki, Setä, Ylli, Jesse ja Eino. Nauru tekee hyvää ihmiselle.

Marialle kumarran kiitokseksi yhteisestä taipaleesta aikuisuuteen ja akateemisuuteen. Tärkeän vertaistuen lisäksi omaksuin sinulta vaalimisen arvoisia ajatuksia siitä, mistä rakentuu ihmisen kokonaisvaltainen hyvinvointi elämässä. Noista opeista jäänen velkaa sinulle, ja en taida noin arvokkaita neuvoja pystyä koskaan maksamaan takaisin. Toivottavasti se ei haittaa. Kiitokset Viljolle valloittavasta kaveruudesta ja Hugolle ajan arvokkuuden oppitunnista. Kirstille iso kiitos hengähdyshetkistä kuvan kauniin Näsijärven rannalla ja Hessulle, Jutalle sekä Kirjavaisille yhteisistä leivänmurtamisen hetkistä. Pelkällä tieteellä ei elä, joten välillä on hyvä syödä eväitä hyvässä seurassa.

Isosiskoni Kati aurasi minulle tietä yliopistoon siinä missä Arvo-papalta taisi tulla verenperintönä kiinnostus kunta-asioihin. Olen hyvin kiitollinen siskolleni avusta opinnäytetöiden kanssa ja pelivinkeistä akateemiseen maailmaan. Olet ollut monella tapaa hyvä esikuva maailmalle ponnistamisessa. Ella ja Linnea, on ollut ilo seurata teidän kasvamistanne ja lautapelitaitojen kehittymistänne. Karrille kiitos aina ajankohtaisista teknologiavinkeistä, joiden takia olen innostunut käyttämään erilaisia varsin hyödyllisiä applikaatioita työssäni. Kamulle kiitos lukuisista hännän heilutuksista. Äitiäni haluan kiittää luovuuteni ruokkimisesta jo pienestä pitäen. Ei ole itsestäänselvyyys, että saa soitella kaiken maailman soittimia aamusta iltaan saman katon alla. Arvioitsijoiden mukaan väitöskirjan luovuus oli yksi sen vahvoista puolista, joten erilaiset luovuusharjoitukset ovat selkeästi valmentaneet minua tätä väitöskirjaprosessia varten jo kauan ennen kuin se alkoi. Isääni kiitän siitä, että hän on aina auttanut minua ja vienyt minua nuorena tärkeisiin harrastuksiin, joissa on ollut oiva sauma oppia ihmiselon käännteitä ja joukkuepelaamista. Inkerin perhettä haluan kiittää heidän maailmansa avaamisesta minulle.

Viimeisenä haluan kiittää Inkeriä. Inkeri, sinun silmäsi ovat olleet päivänsäteeni aamuisin ja ilta-aurinkoni iltaisin. Hymysi lämpö on toivottanut minulle sekä hyvää yötä että huomenta pitkin tätä ajatus- ja kirjoitusprosessia. Positiivisuutesi on eittämättä harvinaista ja juovuttavaa eliksiiriä. Se on ollut mittaamattoman korvaamatonta minulle.

Näin lopuksi haluan todeta, että työni omistan niille, jotka haluavat antaa sille mahdollisuuden opettaa heille jotakin maailmasta. Kaikkia teitä varten tämä työ on kirjoitettu, missä sitten ikinä olettekaan.

Irjalan Karibialla

Tomi Rajala

Abstract

Life in the information age includes encountering vast amounts of information. The era reflects how much time is being used for creating, processing, sharing, storing, and using information. In this flood of information, separating facts from fiction is not an easy task; neither is deciding what information should be used. Indeed, there is counter knowledge or fake news, and then there is knowledge that is either usable—or not. In the context of the public sector, fact checking and defining usable information is also part of everyday life, and problems in information use are as evident as they are in the surrounding society. The problematic nature of information use is the main source of inspiration for this thesis, which narrows its focus to the following research question: What reasons stimulate the use and non-use of public sector performance information?

This research question is not a new one, but it deserves revisiting because current theories of performance information use in the field of public management and administration are lacking in terms of coverage and cohesion. The coverage is lacking because the current research has largely neglected the non-use of performance information and the ambiguity associated with factors driving performance information use. Moreover, the studies have not dealt in depth with how intertwined these factors are. Therefore, this work attempts to broaden the current theory by examining the neglected research areas. Additionally, cohesion in the study field is lacking because there is no agreed-upon common theoretical ground for studies investigating performance information use in the public sector. Consequently, the researchers have adopted a wide variety of theories in a rather idiosyncratic manner, and this has created a disconnected stage where studies on performance information use do not have a common theoretical background that would be easy to identify. This disconnected stage affects the ability to aggregate the results and makes it difficult to see the big picture describing the current knowledge on performance information use in the public sector.

The thesis at hand is based on pragmatism, and it applies a mixed-method and interdisciplinary approach to construct a theoretical framework joining the disconnected studies. According to this framework, the reasons for performance information use relate to individual factors (i.e., demographic attributes, mental models, and power) and contextual factors (i.e., social pressures, information provider, information channel, and features of information). The four articles that are part of the thesis confirm the applicability of the framework and expanded the current theories of performance information use. As an implication, this thesis offers an archive of reasons that practitioners can use to understand and enhance performance information use. For scholars, the study opens new doors to expand the current theories of performance information use. These opportunities relate to non-use and its value, as well as ambiguous, intertwined reasons for use and non-use.

Tiivistelmä

Elämä tietoyhteiskunnassa pitää sisällään suurien tietomassojen parissa toimimista. Aikakautta kuvastaa se, että tiedon luomiseen, prosessointiin, jakamiseen, säilyttämiseen ja käyttöön uhrataan paljon aikaa. Tässä informaatiotulvassa faktojen ja fiktion erottaminen ei aina ole helppoa. Vaikeita ovat myös päätökset siitä, mitä tietoa tulisi käyttää. Vaihtoehtoiset faktat, valeuutiset ja vaikeudet erottaa hyödyllinen tieto hyödyttömästä hankaloittavat omalta osaltaan tiedonkäyttöä. Ei olekaan ihme, että julkisen sektorin toiminnassa faktojen tarkastuksesta sekä käyttökelpoisen tiedon määrittelystä on tullut osa arkea, kuten myös hankaluuksista määrittää hyödyllinen tieto. Tiedonkäytön hankaluudet ovat tämän väitöskirjan suurin inspiraationlähde, sillä työ keskittyy tarkastelemaan sitä, mitkä syyt motivoivat tulostietojen tiedonkäyttöä ja tiedon käyttämättä jättämistä julkisella sektorilla.

Tutkimuskysymys ei ole uusi kansainvälisessä tutkimuksessa, mutta siihen on hyvä palata, koska nykyteoriat aiheesta kärsivät sekä koheesion puutteesta että kattavuuden katvealueista. Koheesion puuttumisella viitataan siihen, että tuloksellisuuden johtamisen tutkimuskentässä ei ole selkeää yhteistä ja yleisesti hyväksyttyä teoriapohjaa, johon tehdyt tutkimukset voisivat kiinnittyä. Tutkimusten teoreettiset taustat ovatkin olleet idiosynkraattisia. Koheesion puute on heikentänyt tutkijoiden kykyä aggregoida ja integroida tutkimustuloksia, jolloin kokonaiskuva tuloksellisuustietojen käytöstä julkisella sektorilla on jäänyt pimentoon. Teorioiden kattavuudessa on parantamisen varaa, koska tiedon käyttämättä jättämisen syyt ovat saaneet liian vähän huomiota ja monitulkintaisuus sekä päällekkäisyys tiedonkäytön syissä on jäänyt suurelta osin tunnistamatta. Tämä tutkimus yrittää laajentaa tulostietojen käytön tutkimusta näiden laiminlyötyjen osaluoiden tutkimisen avulla.

Väitöskirja on luonteeltaan poikkitieteellinen monimenetelmätutkimus, joka pohjautuu pragmatismiin tieteenfilosofisena suuntauksena. Työn tarkoituksena on luoda holistinen teoreettinen viitekehys tuloksellisuustietojen käytön ja käyttämättä jättämisen syistä. Tutkimuksessa luotu teoreettinen viitekehys ehdottaa tiedonkäytön tai käyttämättä jättämisen syiksi yhteen kietoutuneita yksilö- ja kontekstitekijöitä. Tarkemmin jaoteltuna yksilötekijät muodostuvat demografisista muuttujista, mentaalimalleista ja vallasta. Kontekstitekijät viittaavat sosiaaliseen paineeseen, informaation tuottajaan, informaatiokanavaan ja informaation ominaisuuksiin. Väitöskirja koostuu teoreettisen viitekehysten luovasta johdannosta sekä neljästä artikkelista, jotka testaavat ja kehittävät luotua viitekehystä. Tutkimustulokset tarjoavat käytännön-toimijoille kattavan kokoelman

syitä, joiden avulla voidaan paremmin ymmärtää sekä edistää tulostietojen käyttöä julkisella sektorilla. Tutkijoille väitöskirja tarjoaa uusia tutkimuspolkuja, joita kulkemalla voi tulevaisuudessa pyrkiä laajentamaan ymmärrystämme tulostietojen käytöstä. Nämä tutkimuspolut liittyvät tiedon käyttämättä jättämisen arvoon sekä monitulkintaisuuteen tiedonkäytön syissä.

List of original publications

- I Rajala T., Laihonen H., Vakkuri J. (2018) Shifting from Output to Outcome Measurement in Public Administration-Arguments Revisited. In: Borgonovi E., Anessi-Pessina E., Bianchi C. (eds) *Outcome-Based Performance Management in the Public Sector*. System Dynamics for Performance Management, vol 2. (pp. 2–23). Springer, Switzerland: Cham. https://doi.org/10.1007/978-3-319-57018-1_1
- II Rajala, T., Laihonen, H. & Vakkuri, J. (2020) Exploring challenges of boundary-crossing performance dialogues in hybrids. *Journal of Management and Governance*. 24, 799–820. <https://doi.org/10.1007/s10997-019-09485-x>
- III Rajala, T. (2020) Blame avoidance strategies in governmental performance measurement. *Financial Accountability & Management*. 36, 278–299. <https://doi.org/10.1111/faam.12225>
- IV Rajala, T. (2019), “Numerical performance information in presidential rhetoric: Comparing Estonia and Lithuania”, *Journal of Accounting in Emerging Economies*, 10 (3), 321–344. <https://doi.org/10.1108/JAEE-10-2018-0119>

Contents

1	Introduction	17
1.1	Setting the table for the dissertation	18
1.2	Structure of the thesis	23
2	Background for the thesis	24
2.1	Why performance information use became fashionable	25
2.2	Variations in use: A premise for studying reasons for use and non-use	31
2.2.1	Taxpayers, consumers of services, and other resource providers as performance information users	32
2.2.2	Public managers and employees of the public sector entity as performance information users	35
2.2.3	Elected officials and oversight bodies as performance information users	37
3	Why use or not use performance information: The question in different disciplines	40
3.1	Accounting studies and performance information use and non-use	41
3.2	Performance information use and non-use in economics	42
3.3	Perspectives of psychology on information use and non-use	45
3.4	Drivers of performance information use and non-use according to political science	46
3.5	Factors stimulating performance information use and non-use in public administration and management	47
3.6	Purposes of performance information use as competence-related mental models	53
4	Research method	55
4.1	Research method in the synthesis	60
4.2	Research methods used in the articles	62
4.3	Justification for the chosen methods	67
5	Shifting from output to outcome measurement in public administration: Arguments revisited	69

6	Exploring the challenges of boundary-crossing performance dialogues	71
7	Blame avoidance strategies in governmental performance measurement	73
8	Numerical performance information in presidential rhetoric—comparing Estonia and Lithuania	76
9	Conclusion and discussion	79
9.1	The coverage of important factors	80
9.2	Ambiguity in factors driving performance information use	83
9.3	Conceptual relationships between reasons for performance information use and non-use	84
9.4	Theory of non-use	89
9.5	Limitations of the thesis	92
9.6	Where to go from here? Topics for future researchers	93
	References	96
	Appendices	121
	Appendix 1.	
	Table 5. Individual attributes associated with performance information use 123	123
	Appendix 2.	
	Table 6. External environment factors as contextual factors associated with performance information use	125
	Appendix 3.	
	Table 7. Organizational structures as contextual factors associated with performance information use	127
	Appendix 4.	
	Table 8. Management structures and integration of organizational structures as contextual factors associated with performance information use	129
	Appendix 5.	
	Table 9. Reforms and legitimacy of organizational structures as contextual factors associated with performance information use	131
	Appendix 6.	
	Table 10. Organizational culture and performance capacity and results as contextual factors associated with performance information use	132

Appendix 7.

Table 11. Information producer, information channel, and information as contextual factors associated with performance information use	133
Original publications	135

List of figures

Figure 1. Individual and contextual factors associated with the use and non-use of information	20
---	----

List of tables

Table 1. Purposes of performance information use	54
Table 2. Summary of the research methods used in the thesis	58
Table 3. Search rounds and their results	61
Table 4. Intertwined reasons for performance information use (i.e., the chicken-and-egg dilemma)	87

1 Introduction

It has been argued that information is the most valuable commodity in human life. Of course, some may rightfully disagree with this bold statement, but most individuals are likely to acknowledge the importance of information in modern day society, characterized as it is by high-technology solutions created to rapidly disseminate knowledge around the globe. With this in mind, the purpose of this thesis was to evaluate how the theory of performance information use and non-use in the public sector could be improved. The thesis reveals how issues relating to research methods and problems with conceptual and theoretical sophistication in the field of performance management justified the research objective. Different performance movements over the past century are described, along with the reasons why performance information use and non-use has been topical for over a hundred years and will continue to be so. Using a historical perspective, it was important to first construct a theoretical framework that took into account past research results as this would help with an understanding of current knowledge of drivers of use and non-use. The four articles included to this thesis provide multiple theoretical ideas on how to improve the theory of performance information use and non-use. The concluding remarks of the thesis summarize the theoretical contributions made and explain the implications of the results for academic communities and society.

If performance management is a type of performance in the public sector, then this dissertation provides performance information about performance management. The author of this dissertation hopes that people will use the information provided in this thesis. If they do not, the thesis will help the author to understand and learn why this is the case. I could teach this lesson to my students who are the public managers of the future. In this way, the tax money used to produce this thesis would not go to waste.

1.1 Setting the table for the dissertation

While denialism, misinformation, disinformation, and pseudoscience raise concerns currently, they make information use a hot topic. However, the topicality of information use is hardly surprising, considering that information has the power to change human history by shaping how we think. Consequently, it comes as no surprise that information has economic value (Stigler, 1961), and people have a strong desire for data and facts in general. This fascination with information can be seen clearly in the everyday operations of individuals. For example, employees in different companies use approximately 19% of their average work week on information searching and gathering (McKinsey & Company), while Google's information search engine is the most used website today (Alexa, 2019). Simultaneously, we live in the post-truth era, where ignoring information has become fashionable to some extent. The use and non-use of information seems to be everywhere. The public sector makes no exception to this situation, as the use and non-use of performance information has raised many discussions among scholars and practitioners.

Systematically gathered, produced, and shared information on the public sector's performance is considered useful, important, and valuable for myriad reasons (Behn, 2003; Johnsen, 2005). People demand public sector performance information because it describes the inputs, processes, workloads, outputs, outcomes, productivity, and cost-effectiveness of the public sector (Hatry, 2006), and performance information can be used for many things, such as learning, the rationalization of chosen actions, or performance improvement (Burchell et al., 1980; Moynihan, 2009). As information on the public sector's performance can be important and valuable, many report the use of such information (Askim, 2007; Saliterer & Korac, 2013). At the same time, many researchers have provided evidence showing that this type of information is not valued highly and it is either underused or not used at all (Andrews, 2004; Pollitt, 2006; Van Helden et al., 2008; Schmidle, 2011; Raudla, 2012; Grossi et al., 2016). Talbot (2000) summarizes the problem with non-use in a theatrical manner by stating:

If much performance reporting has not, in practice, been meant to change anything why has it been done? If the audience is inattentive, the directors uninterested and the actors apathetic, what's the point? In one sense, this can be seen as a classic "implementation failure."

To understand the use of performance information better, this thesis attempts to answer the following research question: What reasons stimulate the use and non-use of public sector performance information? This research question is further divided into two sub-questions, as follows:

1. How does an interdisciplinary perspective help one to understand the reasons for performance information use in the field of public administration and performance management?

2. What would be the characteristics for a more holistic theory of performance information use and non-use in the field of public administration and management?

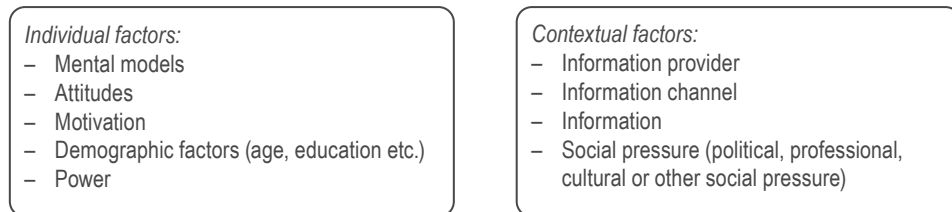
The main research question adopted for this thesis is not new, considering that evidence-based policy making has addressed very similar questions from the 1960s to the present (see Wilensky, 1967; Weiss, 1979; 1983; Harries, Elliott, & Higgins, 1999; Sanderson, 2002; Cairney, 2016), and the New Public Management (NPM) movement has driven public management researchers to ask these same questions during recent decades (see Franklin & Carberry-George, 1999; de Lancer Julnes & Holzer, 2001; Kroll, 2015a).

Regrettably, regardless of the field of study, previous research has primarily been limited by a lack of attempts to conceptually understand the ways in which different factors identified in the literature have simultaneously contributed to the use and non-use of performance information because such factors are intertwined. Current studies mainly offer answers on performance information use and limitations to information use by highlighting a few important factors at a time while neglecting others (e.g., Raudla, 2012; Lu & Willoughby, 2015). Thus, the research has oversimplified reality, generated “tunnel vision” regarding the search for solutions improving the information base in decision-making (e.g., Ammons & Rivenbark, 2008), and enhancing what Mayston (1985) refers to as “decision relevance.” In addition, studies have neglected to consider how the factors are embedded in one another (see Lu et al., 2009; Kroll & Moynihan, 2015). For example, explaining the use of performance information with the existence of performance-based laws neglects the fact that laws are mental models rendered explicit, and laws as mental models create organizational structures and cultures driving performance information use and new laws. Thus, it is justifiable to question whether the mental model, law, organizational structure, or culture induce performance information use, regardless of the findings of statistical analysis. Unfortunately, statistical analysis or artificial categorization in qualitative studies cannot address problems pertaining to a conceptual understanding that arises from the mental models of people, be these people researchers or study responders. The literature talks about internal and external validity, but the point being made about conceptual understanding relates to construct validity, a topic that is rarely addressed in the field. With construct validity, the key question is whether, for example, the effects of social pressure are measured correctly in studies that evaluate performance information use.

What is also missing is a broad synthesis of the results attempting to comprehend how all the factors identified in the past literature can be summarized into a holistic theory of reasons for performance information use and non-use. This study argues that a more holistic theory would include factors seen in Figure 1, as these arise from the fields of accounting, economics, psychology, and political science, and these factors are umbrella concepts that can be used to categorize the results seen in public administration and management. More nuanced knowledge can be built upon the individual and contextual factors seen in Figure 1, and using this framework would provide a common point of reference for the disconnected theoretical approaches used in the current literature. The lack of synthesizing

research and in-depth analysis on intertwined factors in the current research justify the sub-questions stated. However, why has research been inadequate in understanding how intertwined the reasons for non-use and use are, and why has the coverage of reasons been limited? There are many reasons for this, which are described subsequently.

Figure 1. Individual and contextual factors associated with the use and non-use of information



In studies addressing performance information use, the current trend is to use data collection and analysis methods that focus only on some factors, although these factors represent only part of the set of factors influencing performance information use (e.g., Kroll & Proeller, 2013; Askim, 2009; Andersen & Moynihan, 2016). This type of research approach to performance information use can cause omitted variable bias in both quantitative and qualitative studies. According to one of the arguments presented in this thesis, the risk relating to omitted variables has materialized in the field of public administration and management, as all the studies in the field focus on some factors in figure 1 while neglecting others (e.g., Bourdeaux, 2006; Raudla, 2012; Abdel-Maksoud et al., 2015). Because of the omitted variable bias, academics are to some extent unaware of the contribution of the important individual and contextual factors incentivizing the use of performance information as overlaps, dynamics, and relationships occurring between all the factors are being ignored. The article and book chapter formats with word limits may endorse this type of approach where some important incentives of use according to past results are being omitted. Meanwhile, ignoring the important factors associated with performance information use may prevent a holistic understanding of how the use and non-use of performance information occurs. It may also hamper the ability to comprehend the value of the use and non-use of performance information to users. Fortunately, the pitfall arising from word limits can be avoided, and the conclusion and discussion sections of this dissertation offer insights into how to achieve this.

In conclusion, the current literature teaches us disconnected lessons on how few variables associate with each other in performance information use, but the big picture remains rather blurred. The analysis habit of researchers to exclude some important factors associated with using has led to this situation, where a comprehensive theory on performance information use is underdeveloped. It is true that myriad reasons for performance information use have been investigated and the principle of parsimony has been applied (e.g., Kroll, 2015a).

However, the principle of parsimony has been used in a strange way in the theory building occurring in the field of public administration and management because it simplifies the measured phenomena by leaving out important factors driving use. In fact, the classic idea behind the principle of parsimony encourages researchers to describe all the reasons that are found to influence performance information use in the simplest way. The current version of parsimony seen in performance information use studies limits factors driving use while not considering how to cover past results in a holistic manner in their studies. A lack of holistic understanding has meant that the cohesion between past results and newer studies could have been better (e.g., Moynihan & Pandey, 2010; Hammerschmid et al., 2013; Lu & Willoughby, 2015). The currently used version of parsimony may have resulted from pressures to publish novel findings, which is difficult if one connects the new findings to past studies too closely so new ideas start to seem just like the old ones. Another possible reason for the selective approaches adopted by the researchers may be that no one has conducted an extensive literature review on the topic.

Besides the limited focus of researchers on some factors, there is another reason why past research has had problems in addressing all the relevant factors influencing performance information use. The research literature has examined causal relationships, but it has done so mostly by using surveys and experiments built into surveys (e.g., Saliterer & Korac, 2013; Nielsen & Moynihan, 2017). The problem with this approach is that surveys and experiments have a limited focus chosen by researchers and they report how respondents understand and conceptualize the world, not necessarily how the world actually is (e.g., Folz et al., 2009; Askim, 2009; Taylor, 2011). The limited focus chosen by the researcher may lead to situations where respondents report simplistic heuristics to use performance information because the survey format guides them to do so (e.g., Yang & Hsieh, 2007; Johansson & Siverbo, 2009; Moynihan & Landuyt, 2009). Survey responders may also conceptualize matters differently than the designers of the survey and overemphasize some relationships while neglecting others if these are, for example, latent ones that allow performance information use and make it possible, though these relationships are not seen as important use performance information in the current situation by the responders. As surveys and experiments used in the performance management field do not provide opportunities for the researcher to interact with informants, problems in respondents' understanding and conceptualization remain unseen. Interviews do offer chances to explore respondent's conceptualization in depth, but so far these opportunities have not been used.

If surveys or interviews are used to make claims about the world, it is assumed that the respondents can describe the world with the help of survey questions (e.g., Melkers & Willoughby, 2005; Moynihan, Pandey, & Wright, 2012). Otherwise, these studies would only report how the responders saw the factors driving performance information use, and they would not make any claims on what factors influence performance information use in reality. In the field of performance management, the researcher using interviews, surveys

and experiments usually seems to assume that respondents are capable of reporting the world as it is (e.g., Dull, 2009; Masal & Vogel, 2016). This thesis questions this assumption, which is held very dear among scholars studying performance information use and is not alone doing so (e.g., Van Helden, 2016). One of the main arguments of this thesis is that ambiguity in reasons and intertwined factors has been largely ignored in past studies, as scholars usually do not consider these issues. The thesis also argues that asking questions about limited set of drivers of use guides the respondent to simplistic answers that generate tunnel vision among scholars and practitioners. This tunnel vision jeopardizes our ability to see solutions to problems related to performance information use and non-use arising from intertwined factors incentivizing use.

In general, this thesis is not alone with its criticism of surveys and experiments based on surveys, as many economists have been skeptical of the survey-based techniques used extensively in behavioral economics. Some economists favor revealed preferences over stated preferences from surveys in the determination of economic value (Hirschey, 2008), and the common view is that experiments and surveys are at risk of strategic behavior and systemic biases (Furham, 1986). Thus, this thesis uses other methods than surveys to study performance information use. However, as respondents can bias interview results too, it is necessary to control the possible common source bias. Therefore, this study utilizes variety of empirical data taken from separate sources, and different qualitative and quantitative techniques to confirm the results.

In addition to limiting the studied factors, the past studies have mostly devoted their time to understanding how and why performance information is being used, not why performance information is not used (e.g., Moynihan et al., 2012; Kroll, 2015b; Sinervo & Haapala, 2019). The research questions used have mostly focused on the following research question: What motivates performance information use or is associated with it (e.g., Kroll, 2015a; Saliterer & Korac, 2014)? The chosen direction has meant that, currently, we know very little about the reasons used to justify the non-use of performance information (see Angiola & Bianchi, 2015). Thus, this thesis attempts to broaden the understanding of the field by studying what factors can provide reasons for the non-use of performance information.

In summary, non-use is under-examined topic and the simple heuristics used to understand what factors incentivize performance information use arise mainly from two things: the survey and interview methods focusing on only some factors while neglecting many others and the overly simplistic theoretical and conceptual views of the scholars and responders. Thus far, this introduction has been very critical of past studies on performance information use. However, not to make the matter black and white, the past results have also been very useful, as they have provided results describing what factors motivate performance information use. This is a fertile starting point for this research.

1.2 Structure of the thesis

The rest of this thesis is structured in the following way: The chapter Two explains why performance information as a topic is relevant to the field of public administration and management and why one should study more the reasons for use and non-use. After reviewing the background providing the justification for this study, the thesis moves into the theoretical section presented in chapter Three. By creating a conceptual system of reasons for performance information use in chapter Three, this theoretical segment attempts to generate a holistic theoretical framework describing factors affecting performance information use in the public sector. Here, studies on accounting, economics, psychology, and political science are used to create a theoretical framework to categorize the reasons for use and non-use seen in public administration and management. This theoretical framework is then modified later on in chapter Three based on the categorization exercise with the public administration and management literature. The research method is described in the chapter Four of the thesis. The results of research articles and their contributions to performance information use studies are presented in Chapters Five, Six, Seven and Eight. In Chapter Nine, the synthesis ends with a discussion and conclusion. Consideration is given to the theoretical processes and interdisciplinary perspectives required when devising a holistic theory of performance information use and non-use.

2 Background for the thesis

The concept of performance is central to performance information use studies. Traditionally, performance has been defined in two ways (van Dooren, Bouckaert, & Halligan 2010). According to the first way, performance is seen as the result of a production process whereby inputs are allocated to output production and used to generate outputs leading to desired outcomes. The process where inputs are transformed into outputs and outcomes is tracked by using the well-known 3 Es model that measures the economy, efficiency, and effectiveness of public sector service production. When performance refers to the 3 Es model, performance information uses then concentrate on information on economy, efficiency, and effectiveness. In the second way, the concept of performance refers to the realization of public values. However, there are two different interpretations of how the realization of public values occurs. The first interpretation comprehends *performance* as a synonym for public value, while the second understands performance and values as two distinctive concepts (van Dooren, Bouckaert, & Halligan 2010). In the second interpretation, public values provide the point of reference against which performance is measured. *Performance* can mean not only *Economy*, *Efficiency*, and *Effectiveness* but also extensiveness, acceptability, equity, fairness, and many other concepts (Pollitt, 1986). Nonetheless, performance does not have to cover all these aspects. Indeed, performance information use can concentrate only on some public values or all of them. As a field of study, public administration is interested in knowing what performance information is being used and, thus, what public values are being monitored and pursued for.

Past studies have scrutinized whether performance information has been used and, if so, how and why it has been used and what have been the impacts of use (e.g., Askim, 2007; Raudla, 2012). Concepts such as soft and hard use, opportunistic use, demand- or supply-driven use, and formal and informal use describe how performance information has been used in the past (Van Dooren & Van de Walle, 2011). By demonstrating how performance information use varies in different contexts, studies investigating variations in use have also contributed to how dimensions of performance information use. Why performance

information is being used has listed purposes for use and identified factors associated with use (e.g., Kroll, 2015a; Lu & Willoughby, 2015). Research focusing on the impact of performance information use has reported whether performance information contributes to successful performance (e.g., Van Thiel & Leeuw, 2002). This thesis mostly contributes to the studies examining the why question, although some new knowledge is produced on whether performance information is being used and how it is used. However, the thesis does not study the effects of performance information use because this would require another type of theoretical framework describing the different impacts of performance information use. Moreover, due to the space restrictions in current article formats, combining all the research topics seen in the performance information use literature would not be possible in the research articles that are part of this thesis.

Next, the importance of the topic is described. First, a short overview is provided addressing the question of why performance information use has been a relevant research topic in recent history. Second, past studies are used to demonstrate why it is essential to study the reasons for performance information use and non-use.

2.1 Why performance information use became fashionable

The road leading to studies of performance information use in public administration could have started at different points in history. In this review, the focus is, however, limited to the twentieth and twenty-first century movements in administration relating to performance information use. Van Dooren (2008) adopted similar approach in his work and the review in this thesis is partly based on the one conducted by Van Dooren. However, some additions and extensions are added to the review which were not part of Van Dooren's work indicating that more performance movements are identified in this thesis.

The social survey movement was one of the first movements of this type, and it was based on the notion that social reformers need information on social problems and quantified data on governments' performance results (Bulmer, Bales, & Sklar, 1991). What sparked the social survey movement in the early parts of the twentieth century was Booth's argument claiming that the debate on poverty was unsophisticated because it lacked facts about the poor and the reasons leading to poverty (Booth, 1988). According to Booth, one needs information about societal outcomes of government actions, and this information should be used in discussions addressing the poverty problem. The aim of the social survey movement sounds familiar when compared to current topics in the field of public administration and management addressing the outcome orientation of the government or the lack of this orientation (e.g., Ferlie et al., 2005). From Booth to the current debates on outcome orientation, the focus has remained on performance information use and non-use.

Besides the social survey movement, the scientific management and science of administration represented another type of effort driving the use of performance information in the early parts of the twentieth century (Van Dooren, 2008). This

movement relied heavily on the notion of a rational manager who has planned objectives and measures tracking organizational development and who is motivated to find the gold standard in operations and activities by using performance measures (Mosher, 1968). Scientific management and the science of administration as a movement aimed to turn administration into a profession and science based on the measurement and use of measured information (Van Dooren 2008). One of the assumptions driving the movement was that performance information is used in rational management (e.g., Wagner, 1991). This line of thinking is strongly still present in today's public administration, as well as among scholars examining performance information use in the public sector.

The development of cost accounting has also provided a push toward performance information use in the public sector (e.g., Ansari & Euske, 1987). Cost accounting aims to improve the cost awareness of organizations by tracking, documenting, and evaluating costs. Simultaneously, cost accounting offers control and openness to the public sector (Previts & Merino, 1979). In the performance management literature, information about costs is known as *input information*, and it is used to monitor productivity and efficiency (i.e., the relationship between outputs and inputs) as well as cost effectiveness (i.e., the relationship between costs and outcomes) (Hatry, 2006). Although cost accounting has been around a while, it is not fully developed in the public sector and a lot needs to be done to improve the cost awareness of public organizations around the globe (e.g., Pollitt & Bouckaert, 2011). One main agenda driving cost accounting was to enhance input information use and, through it, the productivity and efficiency of public organizations. By producing and sharing cost information, the pressure to use such information usually increases. However, as the role of accounting information in societies seems to change over time or depending on the context and users; such information is not created solely to improve decision-making (Burchell et al., 1980).

To further increase performance information use in the government, performance-based budgeting approaches were adopted in several countries (Jones & McCaffery, 2010; Mauro, 2016). Interest in performance budgeting dates all the way back to the 1870s in the U.S. (e.g., in New York City); however, most significant developments in this area occurred in the 20th century (Jones & McCaffery, 2010, p. 43–67). As budgeting practice, performance-based budgeting relies on the measurement of performance, integrating performance measurement into budgeting, and using performance information during the budgeting process. The aim of performance-based budgeting is to improve the efficiency and effectiveness of public expenditures by linking the funding of public sector organizations to the results they deliver (Robinson & Last, 2009). These linkages between expenditures and results are achieved if performance information is being used systematically in the budgeting process. Thus, the performance-based budgeting movement highlights the importance of performance information use in the budgeting context so more rational and informed decisions can be made.

The performance budget movement facilitated the first wave of evaluations in the 1960s and 1970s. The evaluation movement used social science-based methods to assess the outputs and outcomes of policies (Wollmann, 2003). Group experiments were used to examine the most optimal means of achieving goals that would then be adopted by decision-makers (Vedung, 2010). According to Vedung (2010), scientific evaluation lost momentum in the 1970s, and evaluation practices started to utilize participatory and non-experimental methods that incorporated different stakeholders participating in dialogue about performance. During the neo-liberal evaluation wave, deregulation, privatization, contracting out, efficiency, and customer influence became key concepts, and evaluation served accountability, customer satisfaction, and, most importantly, value for money. The latest evidence-based evaluation wave seems to be a scientific experimentation-based renaissance (Vedung, 2010). The series of evaluation waves is an example of the promotion of the use of performance data in the public sector as it has been argued that evaluation is a form of non-recurring performance measurement.

Parallel to the already mentioned performance initiatives, the social indicator movement attempted to gain momentum in societies (Innes, 1990). Problems in economic growth alongside the creation of the welfare state incentivized the demand for social data (Dowrick & Quiggin, 1998) and provided justification for the social indicator movement (e.g., Rose & Miller, 2008). Governments started to collect more statistics on social phenomena, and this increased the supply of governments' outcome information in societies (Innes, 1990). The aim of the movement was to capture more comprehensively the characteristics of the population by developing measures tracking health, education, leisure activities, social security, wealth, and crime, to name a few. Social indicators make matters more governable through the provision of information that can be used in decision-making. Performance information use was a central part of the social indicator movement as it allowed control over societies and boosted governmentality (e.g., Rose & Miller, 2008).

Whereas the social indicator movement focused on outcomes, total quality management was interested in both the internal world of organizations and their external operating environments. Indeed, total quality management concentrated on customer satisfaction and input, process, and output quality (Reddy, 2012). In general, performance data played a central role in total quality management (Kaufman, 1992) because the quality of inputs, processes, and outputs can be improved by using performance data (Sinclair & Zairi, 2000). Once again, the idea that performance information should be used was restated, as performance information use was embedded into the quality management movement.

The performance movement that truly kickstarted performance information use studies in the field of public administration was the New Public Management reform, as this reform highlighted the need to change performance information use practices in the public sector. Because performance information use studies are linked so closely with the adoption of New Public Management reform, this reform is dealt with here in more depth than the other performance initiatives promoting performance information use.

The roots of New Public Management are in the criticism faced by the traditional model of bureaucracy (e.g., Osborne & Gaebler, 1992). The model of public administration known as traditional bureaucracy (Hughes, 1998) was based on classical public administration (Wilson, 1887) and Weberian bureaucratic principles. In traditional bureaucracy, politics are separated from administration; there exists a hierarchical command and centralized bureaucracy that utilizes central planning and requires compliance with rules. According to the critics, the problem with traditional bureaucracy was that it could not be used to address today's problems (Mauro, 2016). The criticisms of traditional bureaucracy opened doors for novel reforms, such as New Public Management (Hood, 1991). Essentially, New Public Management was a mixture of reforms affected by several theories, including the public choice theory, classical and neo-classical public administration theory, principal agent theory, and management theory (Gruening, 2001). Although different countries applied New Public Management differently, the main agendas of New Public Management were quite universal and included the following:

1. Add managerial elements, such as performance management (and performance information use with it), from private sector management models to the public sector governance mode to provide more focus on results.
2. Redefine the boundaries of government by using privatization.
3. Change the structure of the public sector to a more decentralized one.

Many managerial reforms have been adopted around the globe to make the public sector more businesslike and results-oriented (Jansen, 2008). Simultaneously, NPM has both increased and decreased the "publicness" of organizations in society. According to Bozeman (1987, p. xi), "publicness refers to the degree to which [an] organization is affected by political authority". On the one hand, privatization has increased the politicians' political influence on companies that provide public sector services. Performance measurement and information use have also offered opportunities to increase political control over service production in the public sector through the promotion of transparency. On the other hand, the political influence on organizations that provide public services is sometimes weakened because companies may opt to use commercial confidentiality to protect their information concerning the production of public services.

Performance information use as managerial practice has been in the center stage of this movement from traditional bureaucracy to private sector practices (e.g., Van Dooren & Van de Walle, 2011). In New Public Management reform, the rational manager measures performance to see how goals are achieved and uses this measurement information to steer the organization toward goal achievement. Because New Public Management has advocated performance management consisting of performance measurement and performance information use, performance measurement itself cannot guarantee effective performance management. Here, the activities designed to quantify and categorize performance refer to performance measurement (e.g., Modell, 2004; Hatry, 2006), and

these measurement activities should be used in management practices if these measures are to add value to the operations of the public sector. The use dimension is, therefore, an important step in performance management practices. When evaluating the success or failure of a performance initiative emphasized by the New Public Management movement, it is necessary to examine performance information use (van Dooren, 2008) because real performance management cannot occur without performance information use (Vakkuri, 2010).

Of course, public managers' performance information use requires one to determine what the concept of performance means. Therefore, the New Public Management movement promoted the adoption of the concept of performance as a managerial innovation important to the reform. Although each wave of the New Public Management reform has had particular themes and specific tools, performance has been a vital part of every one of these waves (Talbot, 2005). As a concept, *performance* is an old term that is multi-dimensional, multifaceted, and ambiguous at best (Vakkuri, 2010). Moreover, performance can even be "the great unknown" (van Dooren, Bouckaert, & Halligan, 2010, p. 16). The conceptual nature of performance complicates performance information use and can make New Public Management reform unsuccessful in making the public sector more results oriented. Sometimes, performance information has failed due to the conceptual ambiguities associated with the concept of performance. Besides emphasizing performance information use, the possible complications in the public sector performance information use arising from the New Public Management reform generate demand for studies that investigate use and non-use.

Performance measurement also creates pressures to use information and study this use. Of course, performance measurement is nothing novel, but its role has been emphasized in the New Public Management movement (Cavalluzzo & Ittner, 2004; van Helden, Johnsen, & Vakkuri, 2008). Because of New Public Management, performance measurement has been adopted on an international and government-wide scale, and in all management functions (van Dooren, 2008), and measurement has developed into its own industry (Johnsen, 2005). The result orientation in New Public Management has also meant a shift from inputs to results, and this shift increased performance information, changed the data produced in public administration, and altered the type of evidence used in public management (Kristensen, Groszyk, & Buhler, 2002, p. 10).

Increased data production and changes in the information used to govern the public sector create the need to study public sector performance information use because additional information and renewed management practices based on performance information use should make the public sector more efficient. However, the more efficient public sector is not guaranteed, as there are difficulties in both measurement and the use of measured results (e.g., Smith, 1995; 1996). For instance, the different ways to define *performance* make performance measurement difficult. *Performance* can be defined differently in various levels of administration, such as macro (i.e., government-wide performance), meso (policy

sector, a network, or a chain of events), and micro-level (individual organizations), which complicates measurement (Bouckaert & Halligan, 2008). The produced performance information can also cause perverse behavior and adverse information use, leading to inefficient operations in the administration. Problems in performance measurement and information use provide interesting starting points for researchers investigating performance information use.

The results concerning the success of performance management in the public sector were divided in practice. On one hand, New Public Management and its assumptions about performance information use have received criticism, as the public sector is not identical to the private sector. These differences between the two sectors mean that the adoption of private sector practices in the public sector context can create symbolic managerial practices (Lapsley, 1999). In reality, the ideas presented by the New Public Management movement have suffered from poor implementation (Arnaboldi, Lapsley, & Steccolini, 2015). Hence, New Public Management has lost much of its support (Lynn, 1998). On the other hand, successful performance management practices based on performance information use have generated benefits to the public sector. The divided results on the success of performance management have motivated more studies that aim to understand why performance information use as management practice succeeds or fails.

In addition to New Public Management, evidence-based policymaking (Sanderson, 2002; Cairney, 2016) and knowledge management in the public sector context are among the most recent developments that attempt to increase performance information use in the public sector (Laihonen & Mäntylä, 2018; Lönnqvist, Käpylä, Salenius, & Yigitcanlar, 2014). Evidence-based policymaking promoted information use in policy making because societies should be improved in the name of utilitarianism (Solesbury, 2001). Most typically, knowledge management is described as the process of creating, sharing, using, and managing the knowledge and information in an organizational context. Basic questions addressed in knowledge management ask what type of performance information should be used in management and how we gather, share, and use this information in the most beneficial way (Syed-Ikhsan & Rowland, 2004). These questions show that knowledge management is one of the latest attempts to understand and generate performance information use in the public sector.

The performance movements seen in recent history have inspired governments to generate their own performance models to promote information use. The “New Zealand model” (Boston et al., 1996), the Canadian “La Relève” model (Bourgon, 1998), the Belgian “Copernicus” model, and the German “slim state” model are examples of performance models created by the governments (Pollitt & Bouckaert, 2011). Besides these models, the academic community has also implemented various performance frameworks, including the Neo-Weberian State model (Drechsler & Kattel, 2008), the French model (Bartoli, 2008), Napoleonic model (Ongaro, 2009), Nordic model (Veggeled, 2007), and the digital

era governance model (Dunlevy et al., 2006) and new public governance model (Osbourne, 2010).

Despite various efforts, different performance movements and models have faced challenges achieving regular performance information use, and the research results concerning the usefulness of performance information use are mixed. Consequently, the factors stimulating performance information use are the focus of the present study. A great amount of research has focused on performance information use and drivers of such use in the fields of public administration and performance management (de Lancer Julnes & Holzer, 2001; Jansen, 2008; Taylor, 2009; Moynihan & Pandey, 2010; Liguori, Sicilia, & Steccolini, 2012; Saliterer & Korac, 2013; Grossi, Reichard, & Ruggiero, 2016). However, there is a lack of theoretical work that attempts to theorize the factors driving performance information use and non-use into holistic theory. The theoretical section of this thesis aims to synthesize and theorize the past results describing the use and non-use of performance information.

2.2 Variations in use: A premise for studying reasons for use and non-use

Reviewing the past literature on performance information users around the globe and variations in their performance information use demonstrates the international need to understand the factors driving performance information use. The main problem with the current research is that models explaining performance information use have not been compared and analyzed concerning conceptual overlap and missing aspects. Because this type of comparison has not been conducted, there is no established theory of performance information use covering all the performance information use models and the factors included in these models. By showing that scholars have used limited and selective theoretical approaches previously in disparate parts of the world, this section examining the performance information use habits of diverse user groups reveals the research gap that this thesis is attempting to cover. Knowing the research gap and what generates it provides justification for this thesis, which attempts to generate comprehensive theory on the factors driving the non-use and use of performance information.

In the literature, there are several performance information user categories identified (e.g., Mayston, 1985; van Dooren & Van de Walle, 2011; van Helden & Reichard, 2019). One of the most comprehensive identifications was conducted by Mack (2004), who identified ratepayers (i.e., taxpayers), consumers of goods or services, other resource providers (i.e., suppliers of goods and services, suppliers of finances, donors of funds or voluntary donors of time), representatives or employees of another public sector entity, oversight bodies (i.e., members of oversight bodies, regulators, or auditors), management (i.e., public administrators and managers), and elected officials (i.e., local politicians, state legislators, and members of parliament or congressional representatives). Although performance information user groups could be defined differently based on, for example,

their age, education, or work background, this thesis uses Mack's categorization to structure the literature review on the results of the studies addressing public sector performance information use.

2.2.1 Taxpayers, consumers of services, and other resource providers as performance information users

According to research, taxpayers have been shown to use input, process, output, outcome, efficiency, and cost-effectiveness information. However, the use seems to vary depending on individual attributes and the context (e.g., Lu, 2011; Baekgaard, 2015; Brusca & Montesinos, 2016; Chen, 2017). In these studies, the taxpayers have dual roles, as they also usually act as consumers of public goods and services (e.g., James, 2011; Kool & Bekkers, 2016). Therefore, the results regarding taxpayers often describe how consumers of public sector goods and services use performance information. Here, taxpayers are also understood as consumers of services.

Studies investigating citizens as performance information users have been conducted in Sweden (Murray, 2002), Hong Kong and Singapore (Taylor, 2006), Spain and Italy (Brusca & Montesinos, 2016), the United States of America (Charbonneau & Van Ryzin, 2015), England (James, 2011), China (Lam & Wang, 2014), Denmark (Baekgaard, 2015), and the Netherlands (Kool & Bekkers, 2016). Geographically, studies from Africa and South America are missing while most studies have been conducted in United States. The studies concentrating on the performance information use of citizens have focused on three levels of governance: local (e.g., Baekgaard, 2015; James, 2011), state (Lu, 2011), and central government (Joyce & Levy, 2008). Therefore, there are no studies focusing on how citizens use performance information relating to international relations. Variations in performance information use have been found in local (Charbonneau & Van Ryzin, 2015), state (Lu, 2011), and central government (Taylor, 2006) and in different countries. These reported variations in assorted countries and government levels emphasize the need to understand the factors driving performance information use.

Unfortunately, studies examining citizens as performance information users have only considered some factors associated with performance information use while neglecting other important factors (e.g., Taylor, 2006; James, 2011; Kool & Bekkers, 2016). In fact, studies have been intentionally limited, as this directs focus on novel matters that have not been studied before. This has been the case regardless of what country has been investigated. For example, Baekgaard (2015) investigated whether information content affects attitude formation. Nevertheless, he did not examine how, for example, educational or work background and information channels affect such use. Baekgaard (2015) reported that randomized survey experiments are often used to address problems related to endogeneity (i.e., omitted variables, such as work background and information channels). Nobel prize

winner, Angus Deaton, along with Nancy Cartwright (2018:2), highlighted the problems associated with this type of thinking in relation to the effect of randomization:

Randomized controlled trials (RCTs) are increasingly popular in the social sciences ... We argue that the lay public, and sometimes researchers, put too much trust in RCTs over other methods of investigation. Contrary to frequent claims in the applied literature, randomization does not equalize everything other than the treatment in the treatment and control groups; it does not automatically deliver a precise estimate of the average treatment effect ..., and it does not relieve us of the need to think about (observed or unobserved) covariates.

Based on this perspective, it is difficult to determine the extent to which past RCT results, in conjunction with theoretically limited perspectives that ignore a vast amount of research evidence in the field, can be trusted. Assumptions relating to randomization are widely adopted in the field of performance management as many studies have produced their findings while assuming that from randomization will equalize everything other than the treatment. As a consequence of these assumptions, overconfidence in the published results can be problematic. Overconfidence arises from unawareness of relevant theories and overlooking the effect of chance in the study setting. Deaton and Cartwright (2018:2), describe this problem as follows:

Finding out whether an estimate was generated by chance is more difficult than commonly believed. At best, a RCT yields an unbiased estimate, but this property is of limited practical value. Even then, estimates apply only to the sample selected for the trial, often no more than a convenience sample, and justification is required to extend the results to other groups, including any population to which the trial sample belongs, or to any individual, including an individual in the trial. Demanding “external validity” is unhelpful because it expects too much of an RCT while undervaluing its potential contribution. RCTs do indeed require minimal assumptions and can operate with little prior knowledge. This is an advantage when persuading distrustful audiences, but it is a disadvantage for cumulative scientific progress, where prior knowledge should be built upon, not discarded. RCTs can play a role in building scientific knowledge and useful predictions, but they can only do so as part of a cumulative program [in combination] with other methods, including conceptual and theoretical development, to discover not “what works” but “why things work.”

As there currently exist no systematic literature reviews that attempt to map out concepts and conceptual relationships relating to performance information use, creating a comprehensive theory of the incentives of performance information use is a challenging task for an article or book chapter format. Therefore, the current state concerning models of performance in formation in studies examining citizens is not surprising.

What comes to other resource providers as performance information users, performance information use also seems to vary in divergent contexts and among users. The performance information use of suppliers of goods and services has been mostly studied in the context of hybrid organizations (e.g., Agostino & Arnaboldi, 2015; 2018). Typical examples of hybrid organizations are public–private partnerships (e.g., Liu, Love, Smith, Regan, & Sutrisna, 2014) or public service networks (e.g., Provan & Milward, 2001). In public–private partnerships, performance information concerning inputs, processes, workloads, outputs, outcomes, efficiency, and cost-effectiveness is being used. Limited, moderate (e.g., Agostino & Arnaboldi, 2015; 2018) and high use (e.g., Liu, Love, Smith, Regan, & Sutrisna, 2014) has been reported in studies examining the performance information use of suppliers of goods and services operating hybrids. As public–private partnerships have been studied extensively on each continent and in developed as well as developing countries (Roehrich, Lewis, & George, 2014; de Castro e Silva Neto, Cruz, Rodrigues, & Silva, 2016), the variations in performance information use are well documented (e.g., Bäckstrand, 2008). However, the variation has been explained by using incomprehensive models of performance information use in hybrid contexts, as disparate studies neglect aspects considered important in accounting, economics, psychology, and political science. The incomplete models in the hybrid context justify one aim of this thesis, pursuing a holistic model of performance information use.

Donors of time as performance information users have been examined if one considers the studies examining hybrids, including voluntary sector organizations. Again, the limited, moderate, and high use of performance information has been found in these studies (e.g., Rajala et al., 2018; Agostino & Arnaboldi, 2018; Seppänen et al., 2019). These studies have focused on local, central (e.g., Rajala et al., 2018), and regional government (e.g., Agostino & Arnaboldi, 2018). Moreover, use varies on each level depending on the information type, user, and context. Unfortunately, the theoretical frameworks have not been comprehensive in these studies. As an illustration, Rajala et al. (2018) examined mental models, motivation, power, information, information systems, organizational culture, organizational structure, and social rule systems as possible factors driving the non-use of performance information while they ignored aspects relating to information producers and fiscal stress, among other things.

The use of performance information concerning public sector projects has also varied among suppliers of finance and donors of funds when social impact bonds initiated by the government have been examined (Calderini, Chiodo, Valeria, & Michelucci, 2018). Social impact bond reporting usually focuses on input, output, outcome, efficiency, and cost-effectiveness measures. However, these studies have not addressed what factors serve as incentives for performance information use among suppliers of finance or donors of funds (e.g., Del Giudice & Migliavacca, 2019). According to Hyndman (1990), non-financial information is considered most interesting by the donors of funds. Social impact bonds have been studied on every continent and in the countries listed below. Use has varied

depending on the context, highlighting the global need to further study what drives performance information use (e.g., Del Giudice & Migliavacca, 2019):

1. Austria
2. Belgium
3. Colombia
4. Germany
5. Kenya
6. Uganda
7. Mali
8. Nigeria
9. Democratic Republic of Congo
10. New Zealand
11. Peru
12. Sweden
13. Switzerland
14. Finland
15. France
16. India
17. Israel
18. Japan
19. South Korea
20. Canada
21. Portugal
22. Australia
23. Netherlands
24. United States
25. United Kingdom

2.2.2 Public managers and employees of the public sector entity as performance information users

Public managers are the most studied performance information user group in the field of public sector accounting and performance management (ter Bogt, 2004). In the literature, public managers have been demonstrated to use input, process, workload, output, outcome, efficiency, and cost-effectiveness information (Ho & Ni, 2005). The performance information use of public managers has been studied at the local (e.g., Folz et al., 2009; Kroll & Vogel, 2014; Yang & Hsieh, 2007), regional/state (e.g., Taylor, 2009; Moynihan & Hawes, 2012), and central government levels (Moynihan & Lavertu, 2012). However, public managers operating at the international level have not been examined (e.g., United nations or EU). Whether the focus has been on the local (e.g., Ho & Ni, 2005; Ammons & Roenigk, 2015), the regional/state (e.g., Andrews, 2004; Taylor, 2011), or the central level (e.g., Newcomer, 2007; Jones & McCaffery, 2010), performance information use

has varied. The variations in performance information use among public managers have been reported on all continents and several countries within disparate continents. As an example, the performance information use of public managers has been studied in the following countries, and the intensity of use has varied:

1. Taiwan (Yang & Hsieh, 2007)
2. China (Lu, 2013)
3. New Zealand (Norman, 2002)
4. Australia (Taylor, 2011)
5. United States (Moynihan & Landuyt, 2009)
6. Canada (Abdel-Maksoud et al., 2015)
7. Austria (Saliterer & Korac, 2013)
8. Germany (Kroll & Proeller, 2013)
9. Sweden (Johansson & Siverbo, 2009)
10. Zimbabwe (Zinyama & Nhema, 2016)
11. Kenya (Korir, Rotich, & Bengat, 2015)
12. Chile, Colombia, Costa Rica, and Uruguay (Ospina et al., 2004)

In the studies conducted in the listed countries, the theoretical models used to explain the variation seen in the public managers' performance information use have been selective in terms of what factors have been included and excluded from the models. For example, Moynihan and Pandey (2010) tested whether public service motivation, job attributes, organizational factors (e.g., developmental culture and information availability), external factors (e.g., citizen participation), population size, income per capita, population homogeneity, government size, and region explain performance information use. However, Moynihan and Pandey (2010) ignored information systems, information providers, fiscal stress, laws, and regulations as contextual factors. Overall, reviewing the studies on public managers' performance information use in different countries reveals the need for a holistic model of performance information use, as the study field in its current form lacks theoretical coherence and is disconnected in many parts.

Public employees' performance information use has been examined at the local (e.g., Moynihan, Pandey, & Wright, 2012), regional/state (e.g., Moynihan & Ingraham, 2004), and central government level (Moynihan & Lavertu, 2012; Taylor, 2015). However, few studies investigate this topic. In addition, public employees working at the international level have not been researched. In the literature, public sector employees have been demonstrated to use output, outcome, and efficiency information (Taylor, 2015; de Boer et al., 2018). Performance information use has varied in the local (e.g., Moynihan, Pandey, & Wright, 2012), regional/state (e.g., Moynihan & Ingraham, 2004), and central level (e.g., Moynihan & Lavertu, 2012; Taylor, 2015). Again, the theoretical models used to explain this variation have been selective in terms of the included and excluded factors. As an example, Moynihan and Ingraham (2004) test whether leadership, political factors, and professionalism explain performance information use while ignoring, for instance, the

quality of information and information systems. The studies focusing on public sector employees as performance information users have been conducted in Europe, North America, and Oceania and in developed countries, such as Australia, the United States, and the Netherlands. Although Africa and Asia are not covered by the current studies, the evidence showing variations in use is strong. Similarly, researchers use incomplete models of performance information use in different countries because comprehensive literature reviews are missing on the topic. Studies can be conducted without a comprehensive theoretical framework, as developing hypotheses or proposing research questions does not currently necessitate systematic literature reviews on the topic.

2.2.3 Elected officials and oversight bodies as performance information users

When considering what elected officials have been studied as performance information users, research has concentrated on presidents (e.g., Rajala, 2019a), members of parliament or congressional representatives (Rhee, 2014), state legislators (Lu & Willoughby, 2015), and local politicians (Sinervo & Haapala, 2019). However, the research properly covers only three administrative levels: local, regional (or state), and central (or federal) government. Although Rajala (2019a) has studied presidents before, there is currently a void in research concerning performance information use at the fourth governance level, consisting of forums of international relations, such as the United Nations. The void exists because Rajala (2019) only used brief examples from presidential speeches to demonstrate that presidents do use performance information for problem solving in their rhetoric. Therefore, research on presidential performance information use in international relations is lacking.

Thus far, research results have shown that politicians at the local (e.g., Johansson & Siverbo, 2009), regional (Lu & Willoughby, 2015), central (Raudla, 2012), and federal government (Posner & Fantone, 2007), and the international level (Rajala, 2019a) do use performance information. However, performance information is usually used in limited fashion in local, regional, and central government, as this is the most typical finding (e.g., ter Bogt, 2004; Ospina, 2004; Moynihan, 2005b; Joyce, 2011; Raudla & Savi, 2015; Grossi et al., 2016). In addition, high use (Moynihan, 2005a; Saliterer & Korac, 2014; Ellul & Hodges, 2019) and moderate use (Posner & Fantone, 2007; Lu & Willoughby, 2015) have been reported in these three governance levels.

Information on inputs, processes, outputs, outcomes, efficiency, and cost effectiveness has been used by elected officials (e.g., ter Bogt, 2001; Ho, 2005; Kroll & Proeller, 2013). Quite often, input information is the most used one in the political sphere (e.g., Moynihan, 2005b). Many studies state that performance information use varies depending on politicians' attributes and the context where the elected official operates (e.g., Ho, 2005; Askim, 2009; Charbonneau & Bellavance, 2015; Giacomini et al., 2016). However, important factors associated with performance information use have been omitted from all these studies. For example, Askim (2009) examined how formal roles, municipality

size, the search for other information, education level, and political experience explained performance information use, but he did not test whether information and information system quality or organizational performance explained performance information use. Other researchers limit the scope of their research in the same manner as Askim (2009) does. This omits some important factors while the focus is on others.

Most studies on politicians' performance information use are conducted in the United States. Geographically, studies do not cover politicians' performance information use in Africa, Asia, or Oceania. However, the evidence on variations in performance information use is considerable, and a lack of theoretical cohesion is well documented. The studies conducted in the countries listed on the next page provide examples of how common theoretical bases are missing in the field, as all studies neglect important factors driving performance information use according to studies in accounting, economics, psychology, and political science:

1. Netherlands (ter Bogt, 2001)
2. Chile, Colombia, Costa Rica, and Uruguay (Ospina et al., 2004)
3. United Kingdom (Ezzamel et al., 2004)
4. United States of America (Moynihan, 2005b)
5. Norway (Askim, 2007)
6. Lithuania (Nakrošis, 2008)
7. Sweden (Johansson & Siverbo, 2009)
8. Estonia (Raudla, 2012)
9. Austria (Saliterer & Korac, 2014)
10. Canada (Charbonneau & Bellavance, 2015)
11. Italy (Giacomini et al., 2016)
12. Germany (Grossi et al., 2016)
13. Denmark (Bjørnholt et al., 2016)
14. Belgium (Buylen & Christiaens, 2016)
15. Finland (Sinervo & Haapala, 2019)
16. Malta (Ellul & Hodges, 2019)
17. Portugal (Jorge et al., 2019)

Mack (2004) has reported that oversight bodies do use performance information. Financial information (i.e., balance sheet, cash flow statement, notes to financial statement, and auditors' report) is especially under scrutiny. This indicates that oversight bodies are interested in inputs and outcomes in financial matters but do follow other performance indicators, such as productivity indicators. However, Kroll (2015b) has shown that some oversight bodies might not always use input, output, or outcome information, and performance information use varies among oversight bodies. Again, the research reveals variations in performance information use but has tried to explain these variations with incomplete models of factors driving performance information that neglect many reasons for use and non-use shown in the fields of accounting, economics, psychology, and political science. As an example, Bourdeux (2006) tested whether initial legislative engagement and

disposition to support the reform, institutional structure and capacity, and environmental factors influence performance information use. However, Bourdeux (2006) did not investigate whether the reputation of information producers or information systems affects performance information use. As performance information use has varied in all continents in the studies addressing oversight activities, performance information use variations are well documented. Moreover, there is evidence that important factors driving information displayed in Figure 1 have been neglected in studies conducted in the following countries:

1. Australia (Guthrie & English, 1997)
2. OECD (Organization for Economic Co-operation and Development countries (Curristine, 2006)
3. South Africa, Thailand, Tanzania, Bolivia, and Ghana (Andrews, 2006)
4. United States (Rabovsky, 2014; Bourdeaux, 2008)
5. United Kingdom (Wimbush, 2011)
6. Ireland (McGeough, 2014)
7. Germany, Austria, and Switzerland (e.g., Kroll, 2015b)
8. Norway (Jantz et al., 2015)
9. Australia, Belgium, Italy, Netherlands, Hong Kong, Portugal, Romania, and Sweden (e.g., Wynen & Verhoest, 2016)
10. Zimbabwe (Zinyama & Nhema, 2016)
11. China (Ye & Ni, 2016)
12. Denmark (Nielsen & Moynihan, 2017)

3 Why use or not use performance information: The question in different disciplines

Based on the history of science, many relevant study fields can be explored to find reasons for the use or lack of use of performance information in the public sector. Already the classics of philosophy pondered why people use or do not use information. For example, epistemological arguments concerning the criteria for methods, validity, and the nature and scope of knowledge (Steup, 2005) could well be used to explain the use and non-use of performance information in the public sector. The connection between epistemology and performance information use and non-use just points out how far into history the topic of this thesis could go. As it was not the aim of this thesis to map out the study fields that could provide insight into performance information use in public sector, informed choices were made to limit the scope of the study fields examined.

For the purposes of this thesis, it is more useful to understand the more recent background for studies addressing performance information use, as the connections to recent studies are more evident. As a research area, four fields of study have mainly influenced performance information use studies: accounting, economics, political science, and psychology. Next, a short overview from each of the four study fields is presented to make it evident why studying performance information use has been inspired and affected by these four fields of study. Moreover, investigating these four study fields provides the first opportunity to theorize the factors incentivizing public sector performance information use and non-use. Thus, each overview ends with concluding notes concerning factors that can explain use and non-use.

3.1 Accounting studies and performance information use and non-use

Financial reporting aims to support decision making (Bergmann, 2012), and it is required for accountability and transparency (Barton, 2004; Yamamoto, 2008). Accounting literature has examined the same aspects of performance information use in the public sector as studies of performance information use conducted in the field of public administration (e.g., Noqueira & Jorge, 2017). However, performance information has often been called financial or accounting information in the accounting literature (Brusca, 1997; van Helden, 2016). Various user groups, information needs, performance information usability, performance information types used, reasons for performance information use (i.e., purposes for use and antecedents of use), and how performance information is being used have been investigated in the accounting literature (van Helden & Reichard, 2019). Thus, the accounting literature has contributed significantly to current knowledge on performance information use in the public sector. Accounting scholars examining public sector performance information use have drawn many of their research ideas from accounting studies conducted in the private sector context. Therefore, it is reasonable to argue that accounting literature has had important influence on performance information use studies.

After the implementation of a private sector model of financial reporting for the public sector, many scholars have doubted the usefulness of public sector performance reporting to information users (e.g., Daniels & Daniels, 1991; Carvalho et al., 2006; Lee, 2008; Cohen, 2008; Andriani et al., 2010). Empirical evidence has also shown how performance information has suffered from limited usefulness in the public sector (Barton, 2004; Carnegie & West, 2005). As a counterargument to limited usefulness, more recent accounting studies have demonstrated a degree of usefulness of the performance reporting for several users (Andriani et al., 2010; Kober et al., 2010). Thus, the results of accounting studies have highlighted the need to understand why performance reporting is sometimes used. Therefore, it came as no surprise that considerable effort has been made to determine the reasons for the use and non-use of performance information. In the accounting literature, reasons for use and non-use of financial information have been associated with the following factors: information quality (Chewning & Harrell, 1999; Melville, 2008; Oulasvirta, 2014), mental models (Bourmistrov, 2017), attitudes (Jones & Dewing, 1997), motivation (Luft, Shields, & Thomas, 2016), social context (Chatman, 1989; Saliterer & Korac, 2013), demographic factors (i.e., education level) (Haas & Speckbacher (2017), information channel (Lepistö, 2014), and information provider (Lewis, 2016). In the theoretical section of this thesis, it is tested whether these reasons for the use and non-use of performance information are also in seen in public administration and performance management literature.

At this point, it is necessary to clarify what is meant by mental models, attitudes, motivation, demographic factors, the social context, information providers, information channels, and information quality in the accounting literature. Here, mental models are mental representations showing how something works in the real or hypothetical world

(Johnson-Laird, 1998). Explicitly expressed perceptions, explanations, values, and other knowledge are all examples of mental models (Rajala et al., 2018). *Attitude* is defined as “an enduring set of beliefs about an object that predisposes people to behave in particular way toward the object” (Weigel, 1983, p. 257). Motivation as a concept describes what incentivizes a person to action or inaction (Pervin 2003), whereas demographic factors point to age, sex, income level, religion, marital status, occupation, education attainment level, and other matters of that sort.

Information is, in this thesis, understood as propositional knowledge that is stored, sent, received, or manipulated in any medium (Dretske 1985). *Quality of information* means that information matches users’ qualitative and quantitative expectations for it (Rajala, 2019b). *Information channels*, conversely, refer to “conveyance devices that collect information from a source or sources, repackages it and then disseminates it” (Dunwoody & Griffin, 2014, p. 222). *Social context* points to social pressures arising from social rules, cultural habits, socialization, institutional logics, or peer pressure that guides the use of performance information (Van Maanen & Schein, 1979).

3.2 Performance information use and non-use in economics

From the perspective of economics, performance information use as a topic relates to discussions debating whether humans can act like *homo economicus* in a market context or if their rationality is bounded in decision making. Here, *homo economicus* refers to an economic person who acts like a rational agent in economic models. *Homo economicus* is a forward-looking agent who serves his or her self-interest and has well-established preferences (Zak, 2010). In decision-making situations involving different options, *homo economicus* considers all the relevant information and chooses the alternative with the maximum expected value. The concept of *homo economicus* relates closely to the rational choice theory as it assumes that human actors have stable preferences and engage in maximizing behavior (see Becker, 1976). In the debate on *homo economicus* and rational choice theory, it was questioned whether the assumptions relating to *homo economicus* were realistic, and, if not, how important this finding was to theories of economics (e.g., Stickers, 2003; Levine, Chan, & Satterfield, 2015).

As a response to the concept of *homo economicus*, Simon (1957) proposed the concept of bounded rationality. Of course, Simon operated in multiple study fields, including psychology, economics, computer science, and public administration. However, his work related to bounded rationality is placed under economics because he received the Nobel Prize in Economic Sciences for that work (Simon, 1978). According to Simon (1979: 502), “rationality is bounded when it falls short of omniscience... [a]nd the failures of omniscience are largely failures of knowing all the alternatives, uncertainty about relevant exogenous events, and inability to calculate consequences.” In general, bounded rationality limits information searching and processing because people have cognitive limitations and lack

time (Simon, 1957). Thus, the public sector decision maker as *homo economicus* would use all the relevant performance information to maximize the utility generated by public sector service production. However, decision makers whose rationality is bounded will not use all the relevant performance information. Therefore, they cannot make optimal decisions. The theory of bounded rationality has inspired the field of study known currently as behavioral economics.

In counterarguments to the rational choice theory, behavioral economics studies have claimed that people are predictably irrational (Ariely, 2008). By using psychological experimentation to identify cognitive and emotional bias in human decision-making, scholars in the field of behavioral economics have attempted to change the way that economists understand the expressed preferences and value perceptions of people (e.g., Thaler, 1985). In fact, behavioral economics have shown that people do not purely serve self-interest and individuals are not benefits maximizing and cost-minimizing actors with stable preferences. By contrast, people are subject to insufficient knowledge, problems with processing capability, and inadequate feedback (Baddeley, 2017).

Incomplete knowledge can result from several types of bias, for example, people's tendency to resist change, otherwise known as status quo bias. Status quo bias reduces the extent to which individuals search for information, thus leading to insufficient knowledge (see Samuelson & Zeckhauser, 1988). Besides status quo bias, an individual's emotional state and memories can affect his or her knowledge levels. According to Shiller (2015), the powerful emotions of investors have been known to easily override any suspicion about the real value of an investment. Memories are also problematic because the recollection of past events is likely to be non-representative (Morewedge, Gilbert, & Wilson, 2005), and people evaluate past events based on the most pleasing moments and those at the end of the experience rather than on the average of every moment of the experience. This is known as the peak-end rule (Kahneman & Tversky, 1999). Finally, a lack of salient information in the environment can also lead to a lack of knowledge (Coulter & Coulter, 2005; Thorndike et al., 2012).

The prospect theory is often used to describe the phenomena pertaining to information processing challenges. For example, according to this theory, risk-averse and risk-seeking behaviors reveal how people do not always determine the best possible outcome of a decision in a fully rationally manner when dealing with information (Kahneman & Tversky, 1979). Another example is the Law of Least Effort that posits that lazy thinking leads to errors in information use and decision-making (Kahneman, 2011). Inadequate feedback loops are associated with information avoidance, which involves disregard for important but unpleasant feedback, leading to long-term negative utility (Karlsson, et al., 2009).

An important message of behavioral economics is that uncertainty and context affect decision-making. The current view seems to be that people are poor predictors of uncertain future behavior as present events carry greater weight than future ones (Frederick, Loewenstein, & O'Donoghue, 2002). Indeed, studies that utilized the intertemporal choice

theory demonstrated how the preferences of information users for smaller gains achieved in the short term resulted in a considerable loss in well-being over the long term (Hampton, 2018). When considering the information use context, behavioral economics scholars have argued that people are social creatures who regularly adhere to social norms and desire self-consistency. Social norms are here understood as implicit or explicit behavioral expectations or rules established within a group of people (Dolan et al., 2010). These norms are important part of identity economics according to which economic actions are motivated by monetary incentives and self-concept (Akerlof & Kranton, 2010). The need to preserve a positive view of who he or she is increases a person's likelihood of accepting feedback about social norms.

The effects of social norms were evident in a study demonstrating that information users were not always in conscious control of their actions because they were primed by certain social and cultural conditions (Vohs et al., 2006). To understand how social rules affect people, it is important to acknowledge that adherence to social rules relates to social preferences, such as trust and fairness. Dishonesty destroys trust, and a lack of social norms promotes dishonest behavior (Mazar & Ariely, 2006); therefore, the association between trust as social preferences and social norms is evident. However, fairness involves peoples' desire for reciprocity. The principle of reciprocity requires that a person's action is returned with another equivalent action. Thus, as a social rule, people respond to positive actions kindly and use a punitive response to a negative action (Fehr & Gächter, 2000).

Currently, information use and non-use in economics have been explained by information quality (Greenwald & Stiglitz, 198), education level (as a demographic factor) (Ionciă, Petrescu, Ionciă, & Constantinescu, 2012), mental models (Kahneman & Tversky, 2013), attitudes (Soper & Walstad, 1983), identity (Akerlof & Kranton, 2010), motivation (Epley & Gilovich, 2016), the social context (Bursztyn & Jensen, 2017), information content (Diclemente et al., 2001), information channels (Svensson & Yanagizawa, 2009), and information providers (Akerlof, 1995). As these findings are in line with those in the field of accounting, they reinforce the idea that the framework in figure 1 could be suitable for an analysis of performance information use and non-use in the public sector context.

Overall, many studies in the field of behavioral economics have argued against the accuracy of *homo economicus* because heuristics and biases cause difficulties with the concept of rationality used in the expected utility theory (Kahneman, 2011). Thus, imperfections in human characteristics and the decision-making context elucidate the choices made by people that are not based on careful deliberation. The “nudge” theory was proposed to address limitations related to heuristics and cognitive biases, (Thaler and Sunstein, 2008). Various debiasing and rebiasing tools in behavioral economics have been proposed to address challenges with decision-making (Larrick, 2004). The field of public administration has also investigated the nudge theory and debiasing tools (Nagtegaal et al., 2020).

Thus, several behavioral economics approaches have been used to provide reasons for the use and non-use of information, including bounded rationality (Simon, 1957), the prospect theory (Tversky & Kahneman, 1979), intertemporal choice theory, the *homo economicus* concept, and the nudge theory (Thaler & Sunstein, 2008). Recent studies on behavioral public administration have applied behavioral economics theories to assess performance information use in the public sector (e.g., Baekgaard, 2017; Bellé, Cantarelli, & Belardinelli, 2018).

3.3 Perspectives of psychology on information use and non-use

In psychology, decision-making models include information use as part of their process. In general, information can be used to find solutions to problems requiring actions (Mann, Harmoni, & Power, 1991). Moreover, information is gathered to see whether previous decisions have been successful (Plous, 1993). In problem solving, collecting and using information concerning the problem is also essential for finding solutions and monitoring how the chosen solutions are working (Davidson & Sternberg, 2003). What comes to performance information in public administration, it is often used for problem solving (Rajala, 2019a) and decision making (Moynihan, 2008). Thus, similar ideas about problem solving and decision making exist in the fields of psychology and public administration. As psychology examines information use in decision making, it is not surprising that theories from psychology have been adopted and tested in research examining performance information use in public administration (e.g., Olsen, 2016).

The use and non-use of information due to cognitive biases has also been studied in psychology (Dandeneau & Baldwin, 2004). Here, a *cognitive bias* refers to “systematic selectivity in information processing that operates to favor one type of information over another” (MacLeod & Matthews, 2012, p. 191). Myriad cognitive biases and their effects on information use have been examined. Cognitive biases in decision making have also been studied in management studies (Duhaime & Schwenk, 1985), accounting research (Bol & Smith, 2011), and in the field of public administration (George et al., 2018).

Studies in psychology have shown that the cognitive, motivational, and personal attributes of the decision maker influence the ability to evaluate information diagnosticity and integrate information into a high-quality solution (Taylor & Dunnette, 1974). Motivated reasoning (i.e., prior beliefs) also affects information processing (Taber & Lodge, 2006). In addition, correcting people’s existing misbeliefs is often challenging (Nyhan & Reifler, 2010). Indeed, information users are sometimes motivated to defend themselves against information, creating cognitive dissonance in one’s thinking (Festinger, 1957). For example, it is possible for people to neglect any data or choose only those parts of the information that fit their existing belief systems. According to Festinger (1957), a person may even modify the content of information so it can be better incorporated into current cognitive structures. Dweck (1999) adds that individuals can respond in divergent ways

and can give divergent explanations to the same information because they have dissimilar worldviews. It has also been recognized that people differ in their capabilities to learn from experiences, use different reasoning, comprehend multifaceted concepts, and figure out the correct solutions to problems while adapting to the environment.

It is also known that demographic factors, such as age, can deteriorate cognitive capabilities to use information (Schaie, 2001) and alter attitudes affecting the recall of information (Levine & Bluck, 1997). The features of information have been associated with the use and non-use of information in psychological studies (Malhotra, 1984), along with social pressure as a contextual factor (Asch, 1956). Indeed, it was shown how social pressure changed how people used information (Asch, 1956). Information systems and evolutionary psychology also play an important role in explaining witnessed use and non-use (e.g., Kock, 2010). Moreover, the information provider has offered reasons for both non-use and use (Watkins & Terrell, 1988).

Whether one is talking about motivated reasoning, cognitive dissonance, or the mental models on performance information use, these all have inspired new research investigating the use of performance information. Motivated reasoning (Baekgaard & Serritzlew, 2016), cognitive dissonance (van der Kolk, 2018), problems of various mental models, information, information systems, and social pressures (Rajala et al., 2018) have all been studied in the context of performance information use. Thus, psychological studies have provided multiple ideas for performance information use studies seen in the field of public administration and management. This indicates that reasons for use and non-use seen in psychology could provide a suitable framework for analyzing the results of performance information use studies in the field of public administration and management. When comparing results from psychological studies explaining the use and non-use of information to the ones seen in economics and accounting, they seem to be in line.

3.4 Drivers of performance information use and non-use according to political science

Information is part of decision making in the political world (Calvert, 2013) in the same manner as it is part of decisions according to economics, accounting, public administration, and psychology studies. Because the search for information is closely connected to finding solutions to societal problems, information is at the heart of politics (Baumgartner & Jones, 2015). In political science, researchers have examined who uses information and how (De Figueiredo, 2002; Bimber, 1991), why and how groups share information in politics (Carpenter, Esterling, & Lazer, 2004), and in what contexts is information being used and why (Rich & Oh, 2000).

From political studies, blame avoidance as a research topic is the most influential one seen in the public administration and management field examining performance information

use (e.g., Charbonneau & Bellavance, 2012; Nielsen & Baekgaard, 2015). Blame avoidance as a concept describes actions that attempt to minimize the blame faced by public sector actors when undesired events occur in the public sector domain (Weaver, 1986; Hood, 2014). Otherwise, political theories have not been used very much in performance information use studies, even though many of these studies have examined politicians' use of performance information (e.g., ter Bogt, 2004; Askim, 2007; Raudla, 2012). In general, theories adopted from economics, accounting, and psychology have been more popular in performance information use studies.

The theoretical foundations of most political studies have been rational choice, bounded rationality, blame avoidance, incomplete information, and asymmetric information models (e.g., Hood, 2011; Calvert, 2013). Thus, information, contextual factors, or social context (i.e., social pressures), mental models, motivation, attitudes, and information systems have been used to explain information use. However, topics addressing distrust in politics also demonstrate that the reputation of the information provider matters in information use (Shiratori, 1995). Reasons for information use in politics support similar findings in accounting, economics, and psychology. Thus, these themes will operate as a framework that guides the categorization of factors motivating performance information use and non-use according to the public administration and management literature. The literature in political science does add one theme to those already identified: power. Indeed, political science does show clearly how power over others can make people use information and how the power to use and not use is a significant determinant of use (Hood, 2011). Power is the ability to use or not use information and the capability to make other people use the information.

3.5 Factors stimulating performance information use and non-use in public administration and management

Figure 1 summarizes the theoretical framework describing factors associated with information use and non-use according to studies conducted in the fields of accounting, economics, psychology, and political science. In Figure 1, all the reasons for the use and non-use of information were placed under two broader categories: individual factors and contextual factors. This indicates that the framework is information user-centric, as it separates information users from external objects (i.e., contextual factors) surrounding the information user in the information use context. The individual factors point to the reasons for use and non-use relating to the information user (e.g., Simon, 1957; Taber & Lodge, 2006). The contextual factors are objects in the environment where information user operates. The contextual factors have related to the producer of information (Schul, Mayo, & Burnstein, 2004), information channel (O'Reilly, 1982), context of the information use, and information itself (Borgida & Nisbett, 1977, p. 258; Lindblom, & Cohen, 1979). Next,

this thesis evaluates how reasons for the use and non-use found in public administration and management literature and listed in appendixes 1–7 fit the categories of the theoretical framework.

In general, understanding information use requires understanding the information user. It is essential to realize that sometimes information users might not have the motivation to examine the information at hand, which leads to non-use (e.g., Festinger, 1957). Other times, information may not be used because it does not support the mental models of people (McGrath, 1999). However, it is also known that mental models promote the use of particular information when confirmation bias occurs (Nickerson, 1998). Age, education, and other demographic variables can both promote or inhibit use, as can the use of power. The social pressures set limitations on how freely the individual attributes can dictate performance information use.

Overall, the literature review on factors driving performance information use and non-use confirmed that broad concepts taken from past studies on economics, political science, psychology, and accounting capture the factors associated with performance information in the public administration and management literature. However, the literature on public administration and management provided a much more nuanced view on information use than the other examined fields. Thus, the theoretical framework was moderated. This is very typical for hermeneutic analysis, in which analyzing theoretical parts provides new knowledge on the whole theoretical system and analyzing the whole system creates insights from the parts. The public administration and management literature offered new ways to analyze both the parts and the whole theoretical system describing the factors used to explain performance information use. The conceptual hierarchy in this theoretical framework was formed from five levels. In the following list, this conceptual system is described. The highest level is marked with Arabic numbers (1, 2, 3, etc.), whereas letters (a, b, c, etc.) are used to point out the second-highest level. Lowercase Roman numerals point out the third-highest level. Finally, uppercase Roman numerals show the fourth-highest hierarchy level, and bullet points reveal the lowest level. Overall, the conceptual framework created in this literature review is the following (compare the list to Figure 1):

- I. Individual attributes
 - a. Demographic attributes
 - b. Mental models
 - i. Attitudinal mental models
 - ii. Mental models of identity
 - iii. Competence-related mental models
 - iv. Motivational mental models
 - c. Power

2. Context of information use
 - a. Social pressures
 - i. External environment
 - I. Service users
 - II. Political environment
 - III. Economic environment
 - IV. Social rules of society
 - ii. Organizational factors
 - I. Organization's structure
 - Task structures (i.e., division of labor)
 - Management structures
 - Integrating structures
 - Reforming structures
 - Legitimacy of structures
 - II. Organizational culture
 - III. Organization's performance
 - b. Information provider
 - c. Information channel
 - d. Information

Next, the newly constructed theoretical framework is explained category by category. The first of these subcategories under individual factors was named *demographic attributes*. In this category were variables such as age, education, and work background (see Figure 1 and Appendix 1). The second category (see Figure 1) was altered from attitudes to attitudinal mental models. It included attitudes promoting or inhibiting performance information use. To name a few attitudinal mental models, Pro-NPM attitude, cynicism, and negativity bias affected use. The decision to call attitudes *attitudinal mental models* was based on the fact that attitude is one type of belief and beliefs are often described as mental models in psychological literature. In the modified theoretical framework presented in the list, the third category was called *mental models of identity*. Identity types, such as managerial identity, affected the use of performance information. *Identity* here refers to a complex psychological construct that is an integral part of an individual's self-concept (e.g., Cuéllar & González, 2000; Phinney, 1996). Under the fourth category labeled *competence-related mental models*, one can see various intellectual capabilities, be they emotional intelligence or the ability to understand performance information.

The fifth new category under individual factors holds motivational mental models driving for use. As the attentive reader notes, the category of motivation was renamed to *motivational mental models* (see Figure 1). The reason for this renaming was the author's view that motivation represents one type of mental model because, essentially, motivation consists of world views and beliefs that describe what is worth doing and when. People can be intrinsically motivated, which means that beliefs about what is worth doing and when may come from within the individual (Frey & Osterloh, 2001). Alternatively, they can be extrinsically motivated in which case motivating beliefs are obtained from the

environment. The last category in Figure 1 and Appendix 1 points out that power to use or not use performance information can also explain performance information use. The competence-related mental model contained most factors studied, but this thesis does not calculate whether these factors are statistically significantly more important than other factors.

Besides individual attributes, the context of use also affects information use (e.g., Asch, 1956). Therefore, the research results in public administration and management are divided into four contextual categories describing contextual factors motivating use and non-use: social pressures (i.e., external environment factors listed in Appendix 2 and organizational factors listed in appendixes 3–6), information provider, information channel, and information (see Appendix 7). External and organizational factors are subcategories under the category of social pressures.

The external environment is further divided into many subcategories of service users, service production environment, political environment, economic environment, and social rules of society (see Appendix 2). These categories reflect the suggested drivers of performance information use in the literature. The category of service users had several subcategories, such as service user diversity, intensity of service use, and dissatisfaction with service among the citizens. These subcategories described in more detail what factors associated with service users can incentivize performance information use. The service production environment had subcategories describing various aspects of the environment, be they complexity of the environment or benchmark organizations. Political influences and competition as performance information use drivers are examples of subcategories under the political environment. In the political science literature, Hinterleitner and Sager (2015) have argued that political stakeholders and their information use affect how data about public sector activities are being used by public sector actors. According to Weaver (1986), political competition between an opposition and the government makes information about blameworthy actions valuable.

In the category of the economic environment, there are, for example, subcategories called *market competition* and *economic downturns*. These can also influence performance information use and non-use. For example, performance information can gather attention in times of fiscal stress (e.g., MacKuen, 1983; Ostrom & Smith, 1993; Edwards, 2003). Moreover, economic upturns and downturns affect government actors' actions (Pierson, 1996). The last subcategory was the social rules of society. It included legislation, coercive isomorphism, and professional influences, among other things, incentivizing performance information use. As can be seen from Appendix 2, factors explaining the use of performance data were found most in the social rules of the society category. In addition, the importance of service users and the political environment to performance information use was emphasized in many studies (e.g., Williamson & Snow, 2014).

As subcategories of social pressure, organizational structures and culture can dictate the performance information use of an organization. Organizational factors are divided

into seven subcategories: task structures (i.e., division of labor in public administration); management structures (i.e., plan, budget, measure, analyze, control, manage, and the roles and authority required to do these things); the integration of an organization's structures; reforming an organization's structures; the legitimacy of current organization structures; organizational culture; and an organization's performance results (i.e., input, process, output, outcome, productivity, and cost-effectiveness results). The research has shown that the organization and its structures affect the information use of its personnel (e.g., Mortimer & Lorence, 1979; Chatman, 1989, pp. 459–464; Goh, 2002). Here, *organizational structure* defines how tasks serving the main goal of the formal organization are divided among divisions, departments, sections, positions, and jobs. The structure of the organization also represents a relationship pattern demonstrating the tasks and roles needed in the organization. These tasks and roles require either activities from computers or behaviors expected to be performed by members of the organization. As Flamholz (1996, p. 116) states:

[R]oles refer to the jobs people occupy in organizations, and to the sets of behavioral requirements expected to be performed by people in those jobs. An organization's structure refers to the pattern of arrangements of the sets of jobs comprising the organization. Thus, there are two major elements of structure: 1) roles and 2) their patterned arrangements in relation to one another.

Overall, the literature review revealed that most drivers of performance information use related to organizational structure. In fact, the integration of organizational structures, reforming organization structures, the legitimacy of current organization structures, as well as current task and management structures included many subcategories driving performance information use. This indicates that organizational structure is used in many theoretical frameworks and research settings to explain performance information use. Many studies have also identified organizational structures as an important factor affecting performance information use behavior (e.g., Moynihan, 2005a; Askim, 2007; Kroll, 2015a).

The level of performance and organizational culture were the last two concepts describing the organizational factors affecting use. Past and current performance results as well as predicted future performance drove information use. *Current performance* refers to the social, intellectual, and other capital and organizational capacity obtained. In this thesis, organizational (sub)culture points to a set of norms, values, and beliefs shared by members of the organization. In general, organizational culture influences actions and thinking (Van Maanen & Schein, 1979). Organizational culture is established in the ceremonies and rituals of clans within the organization (Malmi & Brown, 2008). Trust, peer exchange, and staff devotion to evaluating performance information use demonstrated the shared values and norms of the personnel affecting group actions. These exemplified the cultural aspects associated with performance information use in the literature.

As performance information use is an organizational habit, it represents organizational culture. It can seem circular to explain organizational culture with organizational culture. Indeed, many studies have used some form of circular logic when they have stated that performance information as a manifestation of organizational culture arises from the organizational culture. One is here simply saying that performance information use exists because performance information use exists in the culture. The problem with this circular logic often arises because organizational culture as a concept is vaguely defined. Thus, the definition can include performance information use. Clearly, this is not an ideal approach, but luckily, some studies have also explained what aspects of the organizational culture drive performance information use, which avoids circular logic. For example, if trust in the organizational culture drives performance information use, then we can use trust as part of the organizational culture to explain performance information use, which is another part of the culture.

When the category of information is the focus, factors incentivizing use relate to features of the information, which include form, essence, quality, and quantity of information (Rajala, 2019b). Indeed, the form of information can be incorrect according to the information user. Consequently, the information remains unused (Terhune & Kennedy, 1963). As an example, knowledge of a particular event may be used instead of statistical knowledge explaining how such events usually occur because statistical information does not interest the user (Colarelli et al., 2002). Moreover, sometimes, the information is not used because it is either too complex or too simple (Lindblom & Cohen, 1979). Thus, how information preserves the complexity or simplicity of real life according to the information user seems to be one way to justify information use (e.g., Rajala, 2019b).

There can be too much information, as shown in studies addressing information overload (Speier et al., 1999, p. 339). In information overload, some information remains unused, as it is not possible to process all available information (Hahn et al., 1992). Information asymmetries describe situations where information is lacking and thus incomplete because some data cannot be used, as it is not available to the information user (Akerlof, 1995). Before information use occurs, the information needs to be sufficient in quantity to satisfy the expectations of information users (e.g., Elliott & Elliott, 2007). Concerning information quality, Bingman (2006) notes that improving information quality would lead to more accounting information use (see also Boyne et al., 2002). Information as a category had more subcategories than the categories of information provider and information channel.

Access to information, effective communication, information brokers, simple reporting, and integration problems between information channels were all examples of subcategories of information channels. These subcategories as characteristics of information channels have been claimed to be capable of influencing performance information use. Similarly, auditor expertise, credible information sources, and the ability to take part in information production are subcategories of the information provider, and they can boost performance information use.

When reflecting critically on the theoretical framework, it is complicated to separate individual attributes from each other. As an example, mental models are embedded into demographic variables, such as education and work experience. Mental models also connect to power because using power means using mental models on many, if not all, instances. In addition, the conceptual separation between context factors and mental models is rather artificial. This artificiality becomes evident when one understands that information or information systems are also mental models (i.e., knowledge about what type of information and information systems exists in the world). If one wants to change the information system, it requires changes in mental models concerning what type of information systems works best in the world. There are also expectations for information and information systems, and these expectations are mental models representing those features the information or the information system has when it is considered worth using. Individuals also compose mental models from external stakeholders, the political environment, the social rule system, organizational performance, organizational structure, and culture. For example, the social rule system in the form of laws and regulations is also a mental model rendered explicit through legal texts (Rajala et al., 2018).

However, performance information use is not only attributable to mental models. In fact, arguing that mental models are the cause of performance information use or non-use is difficult because service users, the political environment, the social rule system, information providers, information systems, information, and both organizational culture and structure can create mental models. Defining the causes of performance information use leads to a chicken-and-egg dilemma in which it is difficult to determine whether mental models influenced, for example, the social rule system—or was it the other way around? Therefore, searching for causes among the factors driving performance information use takes scholars and practitioners to debates where the root cause is only a proposition that can be infinitely questioned. This endless debate arises from the following premises: Mental models and various contextual factors may portray different phenomena, but they are in many ways overlapping and intertwined in practice. Although it is useful to know these concepts, as they may help to improve performance information use, it is still necessary to know how these concepts are layered together, no matter what the regression analyses and other methods of statistical inference say about the important factors driving performance information use. Thus, while testing the theoretical framework, the articles included in this thesis attempt also to describe how factors are intertwined.

3.6 Purposes of performance information use as competence-related mental models

In the field of public administration and management, many purposes for performance information use can be found (de Lancer & Holzer, 2001; Behn, 2003; Melkers &

Willoughby, 2005; Van Dooren, 2006; Bouckaert & Halligan, 2008, p. 28–144; Van Dooren et al., 2010). Table 1 collects the most typical purposes of performance information use seen in past studies. All these purposes in Table 1 are competence-related mental models describing how performance information can be used in the managerial or political world to do one’s job. As mental models, they can drive performance information use by providing examples of how to use performance information. A noteworthy point in Table 1 is that it does not include purposes for the non-use of performance information, as use is considered status quo and a rational act. This lack of purpose for non-use is a shortage in the current knowledge that this thesis attempts to remedy by examining how blame avoidance as a purpose drives non-use in the public sector.

Table 1. Purposes of performance information use

<i>Author</i>	<i>Purpose(s)</i>
Burchell et al. (1980)	Answer, learning, ammunition, and rationalization
Rossi & Freeman (1985)	Instrumental, conceptual, and persuasive use
Johnson (1998)	Instrumental, conceptual, process, and symbolic use
Kirkhart (2000)	Instrumental, conceptual, and symbolic use
de Lancer & Holzer (2001)	Decisions, strategic planning, resource allocation, program management, monitoring, evaluation, and reporting to internal management, elected officials, and citizens or the media
Behn (2003)	Budgeting, controlling, promoting, evaluating, motivating, celebrating, learning, and improving
Melkers & Willoughby (2005)	Reporting, accountability, assessment, evaluation, budgeting, planning, oversight activities, managing operations, establishing or changing policies, personnel decisions, contracting, benchmarking, and improvement
Van Dooren (2006)	Resource allocation or budgeting, changing work processes, formulating and monitoring contracts, rewarding, strategic planning, communicating, reporting, monitoring, motivating, evaluating, reducing duplicate services, adopting new program approaches, setting program priorities, cost saving, setting individual job expectations, cost-benefit analysis, triggering further investigation and action, enabling consumers to make informed choices, improving responsiveness to customers, creditor and grantor informing, cost accounting, capital management, managerial incentive schemes, management by objectives, staff motivation, strategic human resource management, clarifying objectives, quality models, sanctioning, organizational development, and coordination
Newcomer (2007)	Instrumental and conceptual use
Moynihan (2008)	Deterministic and interactive use
De Lancer Julnes (2008)	Instrumental and non-instrumental use
Moynihan (2009)	Purposeful, passive, political, and perverse use
Hatry (2011)	Accountability, budgeting, and improving
Giacomini et al. (2016)	Reassuring, ammunition (legitimizing and de-legitimizing), answering, and learning uses

4 Research method

Table 2 summarizes the research methods used in this thesis. As one can see from Table 2, the research method of the thesis is based on pragmatism (e.g., Dewey, 1948; 1920) and mixed methods (e.g., Creswell, 2014). According to James (1995), the pragmatic method is adopted when the aim is to settle metaphysical disputes that might otherwise remain unsolved. This is also the case in this thesis, which is based on the mixed-method approach. The pragmatic method states that instrumental value of any claim made in research is determined by the practical consequences of that claim (Murphy, 1990). Therefore, this study tests theoretical claims in empirical settings to confirm their practical relevance. If the methods used in this study lead to the same theoretical conclusions concerning the theory addressing the reasons for performance information use and non-use in practice, the qualitative and quantitative methods used provide stronger evidence for accepting the theory compared to using only the quantitative or qualitative approach.

The theory proposed in this study has instrumental value in problem-solving. The practical consequences of theoretical conclusions, confirmed using qualitative and quantitative techniques, are evident in practical problem-solving that attempts to address the limited use and non-use of performance information in the public sector. The current thesis identified factors driving use and non-use, and it is hoped that the findings will assist the development of incentives for performance information use. Therefore, the study adopted the perspective of Dewey (1990) who argued that the utility of a theory pertains to its problem-solving power. However, the complex framework presented in chapter Three may be impractical if it causes a condition known as paralysis by analysis. Paralysis by analysis means that the problem identification process takes a long time owing to the “chicken and egg” dilemma described in the same chapter Three. The chicken-and-egg dilemma makes it challenging for practitioners to determine the reasons associated with non-use if it causes infinite regression. Here, paralysis by analysis relates to the difficulty in identifying the primary factors in a set of reasons for non-use. Similarly, the “chicken and egg” dilemma could also impede the creation of incentives for performance information

use because it can be difficult to determine what factors associated with non-use should be changed to incentives that would promote use.

As a philosophical approach in this thesis, pragmatism (adopted from ideas presented by Johnson & Onwuegbuzie, 2004):

1. Attempts to find a middle ground between philosophical dogmatism and skepticism
2. Rejects dualism (e.g., rationalism vs. empiricism, subjectivity vs. objectivity, and qualitative vs. quantitative) and reductionism (i.e., reducing culture and beliefs to nothing more than neurological processes)
3. Assumes that knowledge is both constructed and based on the experienced reality
4. Accepts that warranted evidence provides answers that are ultimately tentative, but eventually, the use of scientific epistemology enables one to get closer to larger truths
5. Endorses fallibilism meaning that research conclusions are rarely, if ever, viewed as perfect
6. Sees theories as instrumental (i.e., how well they predict and are applicable in practice)
7. Favors theory that informs effective practice
8. Endorses eclecticism and pluralism (e.g., different research methods are useful in gaining understanding of the world)
9. Uses warranted assertability to justify arguments

Indeed, the thesis attempts to find a middle ground between positivism, constructivism, and hermeneutics. Therefore, the chosen approach rejects dualism and accepts the notion that knowledge is constructed and is based on experienced reality. In this thesis, the extensive literature review in the theoretical section of the synthesis relies on the notion that scientific epistemology can progress toward larger truths. Simultaneously, warranted assertability is used to justify the arguments. This indicates that the conclusions of this thesis are not considered perfect, and fallibilism is accepted. Theories in this thesis are seen as instrumental, as the theoretical framework created from accounting, economics, psychology, and political science well predicts the reasons for performance information use and non-use in the field of public administration and management and in the articles included in this thesis. The theory explaining performance information use and non-use also has practical relevance (e.g., Murphy, 1990), as it helps practitioners to understand and develop performance information use in the public sector. The choice to use a multidisciplinary approach and mixed methods also demonstrates eclecticism and pluralism in the research design.

This thesis is based on the view that both quantitative and qualitative studies have many benefits and costs (e.g., Johnson & Onwuegbuzie, 2004). Although the qualitative approach is seen as more appropriate in some situations and the quantitative approach is perceived as more suitable in others, putting together the two approaches is seen as the most useful approach for this thesis. As a mixed-method approach, this study can be categorized as having a partially mixed sequential dominant status design (e.g., Leech & Onwuegbuzie,

2009). The thesis is partially mixed, as most parts of it apply only qualitative methods and one part applies quantitative methods (e.g., Onwuegbuzie & DaRos-Voseles, 2001). The study is sequential because qualitative and quantitative methods are used in a sequential rather than congruent manner (Senne & Rikard, 2002). Finally, *dominant status* refers here to the fact that most of this thesis applies qualitative methods, which makes the qualitative approach more dominant (e.g., Collins, 2000). To summarize, because the thesis is a mixed-method approach, it:

1. Combines different analysis and data collection methods
2. Applies both qualitative and quantitative research designs
3. Utilizes both theoretical and empirical approaches
4. Employs inductive, abductive, and deductive reasoning in different parts of the study
5. Operates in cross-sectional and longitudinal settings
6. Uses a variety of sampling methods

The articles in this thesis covered politicians, public managers, private sector actors (i.e., resource provider and service producer), third-sector actors, oversight bodies, citizens, and media. The studies were conducted in three countries, all in Europe. Subsequently, the methods used in the synthesis and the research approaches adopted in the four articles are presented.

Table 2. Summary of the research methods used in the thesis

What data were used, and who was studied?	What is the sampling method?	Cross-sectional or longitudinal study?	Qualitative, quantitative, or mixed methods?	Theoretical or empirical study?
Past research (n=410) and research subjects were public managers and politicians	Criterion sampling	Cross-sectional	Qualitative	Theoretical
Interviews, documents, and research subjects involved public and private sector managers	Extreme case and purposive sampling	Cross-sectional	Qualitative	Empirical
Archive data about productivity, reports, documents, financial statements, and news articles and research subjects were public managers, citizens, and media	Theoretical sampling	Cross-sectional	Qualitative	Empirical
Presidential speeches (observational data) and research subjects were presidents	Homogenous sampling	Cross-sectional	Mixed methods	Empirical
Past research and research subjects were public sector managers and employees, politicians, citizens, oversight bodies, service producers, and resource providers	No sampling. As the study is a population study, it includes all articles meeting certain criteria	Longitudinal	Qualitative	Theoretical
Data from the articles, synthesis, and research subjects were the ones named in articles and synthesis	Sampling methods of the articles and synthesis	Both	Mixed methods	Both

Part of the thesis	Philosophy of science	What research question was used?	What inference method was used?	What analysis method was used?
Shifting from output to outcome measurement in public administration—arguments revised	Constructivism	What contingent decision-making arguments stimulate output instead of outcome measurement in public management?	Abductive	Argumentative literature review
Exploring challenges of boundary-crossing performance dialogues in hybrids	Constructivism	What kinds of challenges are in a hybrid organization's boundary-crossing performance dialogues?	Inductive	Content analysis and secondary data analysis
Blame avoidance strategies in governmental performance measurement	Positivism	How can blame avoidance strategies be embedded in performance measurement?	Deductive	Theoretical literature review and content analysis
Numerical performance information in presidential rhetoric Comparing Estonia and Lithuania	Positivism	How do presidents use performance information?	Inductive	Historical literature review, qualitative and quantitative content analysis, group comparisons, and covariance analysis
Synthesis	Hermeneutics	What reasons stimulate the use and non-use of public sector performance information?	Inductive and deductive	Scoping literature review in the fields of accounting, economics, psychology, political science, and systematic literature review applying meta-synthesis in the field of public administration and management
Whole thesis	Pragmatism	What reasons stimulate the use and non-use of public sector performance information?	Inductive, abductive, and deductive	Mixed methods

4.1 Research method in the synthesis

The analysis in the synthesis applies an inductive hermeneutic approach (e.g., Gadamer, 2004). Studies are investigated by examining one discipline at a time while reflecting on how the overall picture changes concerning the factors driving use and non-use. This method provides a view of how the knowledge of reasons leading to the use and non-use of performance information develops when one moves from one discipline to another. As the synthesis also describes the development of understanding over time by comparing results from other disciplines to the ones seen in public administration and management, it can be defined as a longitudinal study. Indeed, results from accounting, economics, psychology and political science were published before the performance information use was studied in the public sector. By using a deductive approach, this thesis investigates how each article fits into the theoretical framework adopted from past studies and how the overall understanding of the reasons for performance information use and non-use changes based on the results of the article. The articles that are part of this thesis are also examined one by one and in a chronological manner to see how they fit into the theoretical framework.

The theoretical section in the synthesis utilizes a systematic literature review method called meta-synthesis (e.g., Walsh & Downe, 2005). A meta-synthesis integrates past research results by using qualitative techniques, and it allows the researcher to interpret past studies. As a research method, meta-synthesis does not test existing theories. Instead, the aim is to generate new theory. This makes meta-synthesis a research approach that expands understanding and provides new ways to develop theories addressing performance information use (e.g., Onwuegbuzie and Frels, 2016). Subsequently, the search method for the systematic literature review is described in detail.

The literature review used three search phrases: performance information use, use of performance information, and use performance information. As there can be sentences in the literature stating that someone did use performance information, the last search phrase was also useful in literature searches. The search procedure focused on accounting, economic, and public management/administration studies, as these are the most relevant disciplines that address public sector performance management (Van Helden, Johnsen, & Vakkuri, 2008). EBSCOHost, Google Scholar, Jstor, and Web of Science were the databases where the literature was searched because these databases have been used in previous literature reviews on adjacent topics related to performance management (e.g., Kroll, 2015a; Mauro et al., 2017). Other parameters limiting the scope of the search results were the following:

1. The study must be written in English to avoid translation issues (language limitation).
2. The research should be published between 1985 and 2018, as this covers the NPM era according to Mauro et al. (2017) (limited time frame).
3. The studies should reflect some aspect of performance information use in the public sector so it would be possible to compare what various studies state about performance information use (limitations on the topic).

4. The review included studies published as book chapters and research accepted in international academic journals and books. Additionally, the word public is included in the name of these books or journals (limited sources). As a quality check, only peer-reviewed studies were used. Additionally, this thesis did not use quality assessment criteria for the studies chosen from international peer-reviewed journals (e.g., Tranfield, Denyer, & Smart, 2003). Moreover, the quality of book chapters that were peer reviewed was not checked in a systematic way.

To gather the data set from relevant research literature, three search queries were used in each database selected for the study. Direct quotations from the selected keywords were searched in each search round, while other parameters limiting the scope of the search results were used, as explained above. Table 3 shows the search results in each search round conducted in different databases.

Table 3. Search rounds and their results

Databases	Search rounds		
	<i>Search one</i>	<i>Search two</i>	<i>Search three</i>
Google Scholar	430	662	448
EBSCOhost	38	33	16
JSTOR	35	72	57
Web of Science	45	58	28
Total number of hits	548	825	549

As a final screening criterion, a word search in the documents was used to find what is being reported in the study about performance information use. Here, word searches were used to confirm whether the article contained information on public sector performance information use. The word search was an efficient way to screen a huge amount of literature that did not focus on performance information use. The word search as a method is also reproducible. After removing duplicates and conducting the word search, the number of articles forming the data for the theoretical section of the synthesis was 486. As there were so many articles addressing performance information use in the public sector, not all the studies are referred to in the theoretical section. However, the theoretical overview provided aims to cover the views presented in all 486 articles by using few articles to demonstrate points made in several other articles not referred to in the text of this thesis. Overall, the conducted literature review is considered a population study because all articles meeting the chosen criteria were examined and relevant results from these studies were described in this thesis by using the investigated research or other research providing similar results.

4.2 Research methods used in the articles

The first article, “Shifting from Output to Outcome Measurement in Public Administration,” was written in collaboration with Jarmo Vakkuri and Harri Laihonen. As the first author, I wrote the book chapter; I was also responsible for creating the theoretical framework and for the practical and scientific contributions that resulted from the framework. However, the chapter was developed from the brainstorming sessions of the three authors. These sessions took place before and during the writing process and also during the revision stage. The first article utilized an argumentative literature review technique to construct contingent arguments incentivizing the use of output indicators while neglecting the use of outcome measures. Argumentative literature reviews look at the literature selectively to develop a body of literature that establishes an antithesis to the thesis (Kennedy, 2007). Criterion sampling was used as the sampling method, as studies with arguments against outcome information use were selected because they matched the criteria.

In this study, the arguments generated present a contrarian viewpoint to the advocates naming the benefits of outcome information compared to output information. These contingent arguments are mental models. The word *contingent* here indicates that the truth value of each created argument is not universal; it is, in fact, contextual. Currently, there are many arguments relevant to this study that have been acknowledged by performance management scholars (e.g., Smith, 1996). However, previous studies have not attempted to systematically and comprehensively gather these arguments to understand why output information is preferred over outcome information on some occasions. This theoretical article fills this research gap by gathering these dispersed arguments. Literature reviews are usually used to map out what we know about the study subject currently, and this study maps out arguments encouraging sticking with output measures while resisting the adoption of outcome measurement in public administration. Overall, abductive inference (Peirce, 1998) and constructivist epistemology (Guba & Lincoln, 1998) describe the research approach, because contingent arguments are constructed in the article by the researchers, and the truth value of these arguments is more uncertain compared to arguments crafted with deductive or inductive inference.

Exploring the challenges of boundary-crossing performance dialogues in hybrids was the second article. It aimed to answer the following research question: What challenges are faced in a hybrid organization’s boundary-crossing performance dialogues? The method of collaboration used by the authors, Tomi Rajala, Jarmo Vakkuri and Harri Laihonen, was similar to the one used in the first article of this thesis. Therefore, the theoretical framework created inductively in the study was my work, as well as the scientific and practical contributions that resulted from the framework.

The research setting was based on a case study approach. An extreme case was chosen, as the purpose of the study was to investigate performance dialogues between heterogeneous organizations providing very different types of services. The case selection served the theory-building objective (e.g., Thomas, 2011), as it was anticipated that the chosen hybrid would

have many difficulties in arranging a working boundary-crossing performance dialogue because the organizations forming the hybrid differed substantially. With the chosen case, it was possible to develop the current theory addressing performance dialogue challenges. The study could also be described as an instrumental case study that creates a general understanding of the problems in conducting boundary-crossing dialogues. The empirical data were formed from interviews (n=11) and documents (n=5). One of the documents used was secondary data by nature (i.e., the term secondary data here refers to research data and the report that were provided by a private company).

Aside from the 11 people who were interviewed, over 100 study subjects were added to the second article of this thesis through the use of secondary data from a well-documented ethnographic study that assessed the same hybrid organization. This ethnographic study was conducted by a private consultation company specializing in field studies, and it involved over 100 study objects. The study report by the company contains detailed descriptions of the research methods, and it explains the research setting. The report also included many quotations that constituted the empirical data set for this study. Our theoretical framework was used to interpret the quotations obtained from the private company report on the study. According to Johnston (2014), the use of empirical data collected by other researchers to meet an alternative research objective is known as secondary data analysis. Secondary analysis is a research method that applies the same basic research principles as studies that examine primary data. Just like any other research method, secondary data analysis involves a number of steps, and the steps followed in this thesis were as follows (Johnston, 2014):

1. The research questions were developed (see the research questions reported in the second article of this thesis).
2. The dataset was identified from the webpages of case organizations.
3. The dataset was evaluated according to the following perspectives: (1) The purpose of the study, (2) collectors of the data, (3) data types collected, (4) time interval for the data collection, (5) methods applied, (6) administration of the primary data and (7) consistency between the secondary data and other data collected as primary data.

The study objective of the ethnographic study was to understand how a hybrid organization affected public services. An evaluation was performed of the work-related well-being of personnel at the hybrid organization, and challenges and opportunities were identified regarding collaboration between the public, private, and third sector actors. The data was collected by the workers of the company. The data collected in autumn of 2017 included observations about the hybrid organization made by workers of the company, written feedback from customers, group interviews with staff, and in-depth interviews with customers, staff members, and partners. As research methods, the study utilized ethnographic study design and thematic analysis. Parts of the primary data were publicly available, and parts were located in the company archives. The data were consistent with

the interviews conducted; this finding was supported by triangulation (reported in the second article of this thesis).

As a primary data, the second article of this thesis utilized semi-structured interviews and internal documents (n = 4) obtained from the organization and the municipality that took part in the hybrid. An inductive content analysis was adopted as a qualitative analysis method because previous frameworks could not capture in detail the challenges of boundary-crossing performance dialogues. Thus, there was a need to create a new analytical framework from the empirical data to describe these challenges of boundary-crossing performance dialogues. The study was cross-sectional because the attempt is to capture the challenges in one time point. Additionally, the aim was not depicting how the challenges of boundary-crossing performance dialogues developed over time. The study provided an opportunity to test how the theoretical framework developed in the literature review could be used as a starting point for the development of new theory.

The third article in this thesis addressed blame avoidance strategies in governmental performance measurement. It used Popper's (1959) hypothetico-deductive method. The article aimed to falsify the universal claim stating the following: Performance measurement does not incorporate blame avoidance strategies. Four hypotheses predicting the use of a certain blame avoidance strategy in performance measurement were created for the hypothetico-deductive method. These are known as falsifying hypotheses according to Popper (1959). In this research setting, accepting even one hypothesis would falsify the claim that measurement in the public sector does not use blame avoidance strategies. Rejecting all four hypotheses would indicate that blame avoidance strategies are not used in performance measurement.

Some might argue that the results of the gaming literature do not justify a hypothesis that states that performance measurement does not incorporate blame avoidance strategies. It is true that, as a phenomenon, gaming addresses blame avoidance, but it also includes credit claiming, which differs to blame avoidance. With blame avoidance, an individual attempts to minimize blame while in credit claiming the objective is to maximize credits (Hood, 2011). Thus, blame avoidance is included in gaming; however, not all aspects of gaming can be included in blame avoidance. From this perspective, the gaming literature results should be considered when the relevance of the universal claim in the third article of this thesis is considered.

Before the universal claim can be justified, clarification is required on what gaming means in the context of performance information design and use. Gaming that relates to performance measurement takes two forms. Firstly, gaming is seen in performance information use behavior. Secondly, gaming is evident in the measurement design. Gaming in terms of information use refers to gaming within an already designed game that has a purpose and a set of rules, whereas gaming in terms of performance design constitutes gaming in the design of the game. This game could be called an accountability game as performance measurement is part of the accountability system and the design of such a

system. Gaming in terms of measurement design refers to gaming strategies that are built into the performance measures defining rules of the accountability game, whereas gaming in terms of performance information use relates to performance information use that reflects gaming behavior (Vakkuri & Meklin, 2006). Creative accounting is an example of gaming within the accountability game (e.g., Vinnari and Näsi, 2008). Designing an accounting system would be example of gaming in the design of the game.

When an accountability game is designed, the rules and purpose of the game are established. As accountability systems are designed to punish those who deviate from the desired behavior or performance results, it makes little sense to use blame avoidance strategies as a design principle when constructing accountability systems and performance metrics as part of that system. If blame avoidance is used as the design principle, well-designed performance measurement in an accountability system, is likely to bring the number of opportunities to blame the account giver to zero. If there are zero opportunities to blame the account giver, these account givers would not have to use blame avoidance strategies during the performance information use stage. As the purpose of accountability systems is to make the account giver accountable, it would be counterintuitive to use blame avoidance strategies that protect the account giver as a design principle in the performance measurement design. If the design of the accountability system and performance measurement were not based on blame avoidance strategies, the ability to observe the use of blame avoidance strategies in performance information use would be more likely.

The public administration literature (e.g., Taylor, 2020) has primarily focused on blame avoidance strategies in performance information use (i.e., gaming after the adoption of measurements), which indicates that blame avoidance strategies are not design principles in accountability systems that are based on performance measurement. This justifies the universal claim that performance measurement does not incorporate blame avoidance strategies. However, blame avoidance and gaming literature provide justification for the falsifying hypotheses. As the main claim can be justified by using accountability theories and falsifying hypotheses can be justified with gaming theory, it is evident that these theories provide contradictive ideas about human behavior in measurement design. The article three attempts to understand what theories better explain behavior in performance measurement design.

The case study in article three utilizes productivity data from 2006–2012 acquired from Statistics Finland. Over 1,900 data points that showed work or total productivity values in central government were collected. In addition, national news articles ($n = 112$), webpages ($n = 10$), and documents ($n = 135$) from central government agencies were used as the empirical data. The research subject was chosen because the chosen productivity measurement methods in the statistics were criticized by the National Audit Office of Finland (2011). The productivity results also generated public blame at the micro, meso, and macro levels. The micro level refers to individual central government agencies ($n = 271$), whereas the meso level reflects the aggregate data that demonstrate the productivity of an

administrative sector ($n = 11$). Here, the total productivity of central government was the sum of the productivity values of the central government agencies and this sum is called as macro level in productivity statistics. Theoretical sampling was chosen because the case was selected owing to its theoretical relevance.

Overall, the empirical data collected from the study subjects provided a comprehensive view of the blame avoidance strategies used at the micro, meso, and macro levels in central government. This added the central government as empirical context to the research setting of this thesis. Investigating productivity information use is especially valuable as there is a scarcity of research on productivity information use in central government owing to difficulties measuring its productivity. The coverage of different administrative sectors and their blame avoidance strategies in performance information use also meant that different service sectors (i.e., judicial and financial services, traffic and communication services, social and health services, environmental and technical services, education and cultural services, services related to agriculture and forestry, governmental research institutes, foreign policy, public security and safety, and national defense) were included in this academic dissertation.

In the third article, content analysis was used as an analytical method. In content analysis, descriptions of blame avoidance behavior described in the four theoretical hypotheses operate as coding categories, and empirical data are coded accordingly. The coding aims to locate public actors' actions that can be identified as blame avoidance behavior. The content analysis was based on a positivist approach because hypotheses suggesting the use of blame avoidance were derived from theory and proven with a rigorous method to be accurate in the examined case. By using quotations and empirical examples, the analysis becomes reproducible and transparent to the reader, which makes the method rigorous.

In the fourth article, called "Numerical Performance Information in Presidential Rhetoric," the purpose was to observe how two sitting presidents use numerical performance information in their speeches. The empirical data consisted of 35 public speeches given by the president of Lithuania and 85 by the president of Estonia. These presidents were chosen because they were considered similar as possible. However, it is possible that the speeches that were evaluated were written by political advisors or speech writers or in collaboration with different ministries. A limitation of the current study was that it did not collect any data on who wrote the speeches, which created problems understanding the relationship between performance information use, the personal attributes of the presidents, and the speech context. This limitation revealed a general problem in relation to observational studies, namely the difficulty of researchers having definitive knowledge of the author of the public statements given by public managers, politicians, or citizens. This makes it problematic to identify who is being observed during the delivery of speeches. The matter becomes even more complicated when the speeches contain ideas adopted from other people. This presents problems with whether the messenger giving the speech or the original source of the information is being monitored. To be sure, the entire speech

writing process must be monitored, and it is important to determine whether the written speech is the same as that delivered at the public event and whose ideas are presented in it. As the current study did not monitor the writing process, caution is advised with the interpretation of the results.

The study concentrates on politicians' performance information use in mass communication. The topics in public speeches ranged from legal matters to foreign and domestic policy issues. Examining these speeches provides an opportunity to produce new knowledge on how performance information use in rhetoric appears in diverse policy contexts. The empirical data were analyzed using quantitative and qualitative content analysis techniques. Both manual and computer coding were used as numbers were searched from speeches by the computer while manual coding as a qualitative analysis method coded the numbers found to disparate performance information use categories. In the quantitative analysis, descriptive statistics and statistical tests were used to test the hypotheses created in the research. The study was cross-sectional by nature, as it focused on the difference between presidents in certain time frames. The study did not describe how the presidents' performance information use developed in time and in relation to one another. The study was based on inductive inference, as it generated novel theory from the empirical data. There also were no previous theories explaining presidential performance information use in rhetoric. The research design builds on a positivist approach that utilizes hypotheses and statistical methods.

4.3 Justification for the chosen methods

An extensive literature review involving 486 studies was performed, and the empirical evidence obtained from these studies was determined to be the optimal available method for mapping a current understanding of the drivers of performance information use. The ability to gain a similar understanding through empirical studies conducted in different parts of the world would require an army of researchers to conduct original research to match the studies that were conducted before this dissertation and included in the literature review in chapter Three of this thesis. This would also require an extensive timeframe. Thus, the use of a literature review was a shortcut to achieving the overview of the drivers of performance information use needed to construct the holistic framework. Similarly, an argumentative literature review is a rapid way of testing the holistic framework in one problematic area of performance management. The challenges involved in establishing outcome-oriented public management in different countries are documented in the 410 evaluated articles, and this provided an interesting and comprehensive context in which to examine the different aspects of the theoretical framework created for this thesis. The examined articles were written in the time interval ranging from 1960 to 2017. In the introduction, the author of this thesis emphasized the importance of having knowledge of past theories on the topic before conducting an empirical study. As is evident, this approach

was incorporated in the research design of this thesis and tested with a view to confirming the usefulness of the recommended practice.

To develop new subcategories for the created theoretical framework, interviews were adopted as the type of challenges that would be encountered with performance information use in a hybrid organization were unclear beforehand. Observations are a less flexible way of exploring unknown challenges as there is no way of directing the study subject to an interesting topic. Surveys were ruled out as they require prior knowledge of the challenges.

The content analysis of performance measurement design is a reliable way of investigating blame avoidance strategies built into performance measures provided that data are available on what the measure quantifies. This approach identifies which blame avoidance strategies are incorporated in the designed productivity measures. However, to determine whether blame avoidance is intentional in performance measurement design, being involved in the measurement design is recommended as it would be difficult to obtain confirmation that the measures were designed to limit blaming opportunities.

The observation of presidential speeches is another way to test the framework because a solid theoretical framework is able to provide relevant theoretical predictions in different contexts and research settings. Studying the influence of demographic factors, power, social pressure, organizational structure, and information is possible when considering the examined speech; however, mental models, information providers, and information systems that drive information use cannot be usually determined comprehensively by observing speeches. Thus, the theoretical framework cannot be fully tested by examining speeches because the researcher cannot influence its content. In general, speeches might reveal something about mental models, information providers, and information systems; however, in the current study, they did not reveal enough. Nevertheless, the framework provided many different perspectives on the speeches although the researchers did not have control over their contents.

Other data collection methods, such as surveys, were used in other studies (e.g., Hammerschmid et al., 2013) to assess factors included in framework created in this thesis, which indicates that various data collection methods are compatible with it. However, this study did not use a survey as there are several such studies in the literature. Similarly, the thesis tests the applicability of the framework to theoretical and empirical studies that utilize a diverse set of analysis methods. A pluralistic approach is able to determine the versatility of the framework. Past studies have utilized various research approaches; thus, pluralism was deemed to be an invaluable approach to testing a framework, as well as being in line with the research tradition.

5 Shifting from output to outcome measurement in public administration: Arguments revisited

The first research article in this thesis is called “Shifting from Output to Outcome Measurement in Public Administration—Arguments Revisited.” This article was a book chapter addressing mental models that drive the use of output information in public administration while motivating the information user to ignore data on outcomes. The main idea behind this article was to demonstrate how the non-use of outcome information can occur because mental models favor the use of output information over outcome information (e.g., Carlin & Guthrie, 2003; Bandy, 2011). The attempt was also to show how certain types of performance information can be used because they have more value to the performance information user than another type of performance information left unused. In a world with limited time and resources (e.g., Kristensen et al., 2002), the information user cannot always examine all information. Therefore, it becomes essential to evaluate what information is more valuable than other information to the information user. The chapter shows that performance information users can have a hierarchy of performance information in divergent decision-making situations. At the top of the performance information hierarchy is performance information that is most likely to be used in decision making, whereas at the bottom of the performance information hierarchy is information about performance that is least likely to be used. Thus, the mental model describing the hierarchy of performance information can explain why the non-use and use of performance information exists in the public sector.

The focus of this study was two performance information user groups: public managers and politicians. As a contribution, the article found a complex network of interrelated contingent arguments that may influence the performance information user’s motivation to use output information while neglecting the use of outcome information. Mental models on information need, loss aversion, controllability of results or information, nature of outputs or outcomes, pressure for legitimacy, conflict orientation, and information systems can explain why outcome information is neglected while output information is used. These

results fit well into the theoretical framework created in the systematic literature review segment of this thesis and broaden it by showing new attitudinal, competence-related, and motivational mental models supporting the use of output information and inhibiting the use of outcome information. Moreover, arguments including information need and information systems show how the concepts of information and information system are embedded into mental models. Pressure for legitimacy and conflict orientation in the organization demonstrate how external stakeholders and organizational culture can be incorporated into mental models. The points made about loss aversion and performance measurement over organizational boundaries demonstrate how mental models inhibiting or promoting use can include an organization's performance and organization structures as concepts.

The article did not assume that all the above-mentioned arguments as mental models are present or assessed at the point of performance information use. Instead, it is probable that the performance information user, according to the argument made in the article, does not consider some of these arguments at all. If at least one of these mental models is acknowledged and considered by performance information users, outcome information use may be inhibited, and output information is used. The article also claimed that all of these contingent arguments can be valued differently by various performance information users, and this valuing most likely varies among situations where performance information becomes relevant and the decision to use output or outcome information may deviate from the decision made purely based on weighting and calculating all the arguments, either favoring or opposing output or outcome information use.

6 Exploring the challenges of boundary-crossing performance dialogues

The second article in this thesis aimed to explore and map out the challenges of boundary-crossing performance dialogue in a case organization that can be described as a hybrid organization. As dialogues are often used to transform conflicts into cooperation, the organizational disparities causing conflicts in hybrid organizations' performance management can be dealt with using boundary-crossing performance dialogues. *Boundary-crossing performance dialogues* are understood here as performance management conversations among agents employed by public- and private-sector organizations forming the hybrid (e.g., Rajala and Laihonen, 2019). Performance dialogue as a concept points to situations where “participants jointly interpret performance information and discuss it while identifying the actions needed to manage the performance according to this information” (Rajala et al., 2018).

As a concept, performance dialogue has more extensions than, its predecessor, a learning forum (e.g., Moynihan, 2005a). A learning forum, a type of performance dialogue, is a concept that was proposed by Moynihan (2005a). However, although there are many types of performance dialogue other than learning forums, they have not been addressed in the literature. As the concept of a learning forum is rather normative and describes the features of successful dialogue about performance, the idea was to create an “umbrella” concept that pertains to different types of dialogue and is based on performance information (Rajala & Laihonen, 2019). The “umbrella” concept was termed “performance dialogue” (e.g., Laihonen & Mäntylä, 2017). Moynihan (2005a) did not use this concept nor the definition of performance dialogue, but Pollitt and Bouckaert (2011) used the term performance dialogue although they did not define it.

By exploring the problems of boundary-crossing performance dialogues arising from mental models, motivation, power, information, information systems, organizational structure, and organizational culture, the article provided valuable insights on drivers of performance information non-use in hybrid organizations. Accordingly, the study

identified relationships between inter-organizational factors that caused difficulties for performance information use in boundary-crossing performance dialogues. The study was meant to test more comprehensively the theoretical framework generated in the literature review section. The results were promising, as the framework provided fertile starting ground for the development of new theory. This new theory was in line with the theoretical framework but expanded its application to a hybrid context. The hybrid context added new theoretical dimensions to the framework, as follows:

1. Language barriers between member organizations (problems with competence-related mental models describing how language works)
2. Conflicting mindsets of member organizations (competence-related mental models that relate to what type of activities the hybrid should do)
3. Lack of inter-organizational sanction systems (motivational mental models)
4. Lack of interest in other member organizations (motivational mental models)
5. Powerlessness of member organizations (power)
6. Member organizations as information system silos (information system)
7. Inability to aggregate member organization data (information)
8. Clashes of organizational cultures (organizational culture)
9. Prevailing culture among member organizations (organizational culture)
10. Inter-organizational territorialism (organizational structure)
11. Representative lineup rotation in boundary-crossing performance dialogues (organizational structure)
12. Incompatible tasks of member organizations (organizational structure)

These 12 inter-organizational factors were harmful to performance information use because they hampered open and respectful communication of views on measured performance results. Lack of open and respectful communication inhibited innovation, limited learning, and hindered knowledge-sharing potential, which was embedded in boundary-crossing performance dialogues. Overall, the interorganizational factors causing challenges in boundary-crossing performance dialogues provided reasons for the non-use and limited use of performance information in the examined hybrid. The developed content analysis method can be used in future studies investigating non-use in hybrids. Thus, the study provided methodological contributions.

7 Blame avoidance strategies in governmental performance measurement

The article explored blame avoidance strategies in performance measurement systems to understand how performance information can be used to avoid blame. Using measures to avoid blame can make blame avoidance a purpose that drives performance information use. By nature, this purpose is a competence-related mental model. This adds blame avoidance to the purposes of performance information use. By doing so, the article expands the theoretical framework created in the literature review into new dimensions. The article also showed how blame avoidance as a purpose that drives performance information use leads to the non-use of performance information. As a methodological contribution, the study demonstrated new ways to investigate how blame-avoidance strategies utilized in performance measures affected performance information use.

The following research question was used in this article: How can blame avoidance strategies be embedded in performance measurement? The results of the article showed the use of four blame avoidance strategies in performance measurement. Such results indicate that produced performance information can be used for blame-avoidance purposes if it is designed in the right way.

According to the first finding, the studied performance measure applied the strategy called *herding* when the central government calculated one productivity figure for the whole government by aggregating measures of multiple agencies operating under the government while not reporting the productivity numbers of these agencies separately. This strategy reduces the use of agency-level performance information among citizens and the media. The second result demonstrated that the measurement design was based on reorganization, in which the productivity measures' content was constantly reorganized by the government. The constant reorganization can contribute to low use, as performance information users, such as citizens and the media, cannot track development in time. They must constantly update their knowledge on the used performance measures.

As a third result, the study found that the government gave different performance information to external citizens and public sector actors while hiding evaluation techniques and calculations from the public. Information not given to citizens or the media is quarantined to remain un-used. Thus, this provides reasons for the non-use of performance information. Last, the article found that criticized performance measures revealing poor results were terminated by the government. The termination of measures made it difficult to use of already collected performance information from the central government agencies because there is no-one actively using the collected data in Statistic Finland and the data can be obtained only by paying a service fee. This finding demonstrates that blame avoidance can be one reason driving the non-use of performance information. Cutting human resources that are used to disseminate the collected data is one way to limit the performance information use.

The general assumption is that performance measures are utilized for holding someone accountable (Behn, 2003). However, in this study, performance measures enabled blame avoidance although productivity measurement in Finland was used for accountability purposes. The public pressure for government productivity also related to performance measures incorporating blame avoidance strategies. Productivity was a very appealing notion in Finnish politics. Its importance was emphasized in the eyes of politicians and the public. Because the Finnish central government terminated productivity statistics while many other performance measures remained intact, the study's results are in line with research showing that blame avoidance strategies may be more prevalent in some policy areas compared with others (Johnsen, 2012; Nicholson-Crotty et al., 2006; 2017). This indicates that social pressures can significantly influence performance information use. This is in line with the theoretical framework created in the literature review of this thesis.

Blame avoidance as a mental model driving performance information use and non-use embedded ideas concerning stakeholder pressures, political environment, social rules of society, organizational structure and culture, organizational performance, and demographic variables as a form of individual attributes (see appendixes 1–7). Mental models on blame avoidance influenced use and non-use and were based on the following ideas:

1. Stakeholders with a negativity bias are looking for blameworthy actions (negativity bias is considered an attitudinal mental model, and it promotes better accountability and stakeholder pressure is considered a contextual factor placed under social rules of society)
2. Political conflicts get fuel from poor performance (political environment and performance results as contextual factors)
3. Performance information can be reorganized (reforming organizational structures as contextual factors)
4. Information and information systems can be made to change constantly (information and information channels as contextual factors)
5. Performance culture uses performance information for accountability purposes (organizational culture as a contextual factor).

These mental models describing how blame avoidance is incorporated into governmental performance measures provided fruitful examples on how mental models interacted with other factors, such as political environment and reforming organizational structures, affecting performance information use, and how these other factors were seen in mental models. This thesis analyzed the results of this article in a hermeneutic manner and demonstrated the interaction between the parts of theory and the whole theoretical framework presented in the literature review section of this thesis.

8 Numerical performance information in presidential rhetoric—comparing Estonia and Lithuania

The article studied the numerical performance information use in the rhetoric of two presidents, Dalia Grybauskaitė (Lithuania) and Kersti Kaljulaid (Estonia). The last article focused on age, education, gender, work experience, power, political context, speech context, speech length, and economic environment. The previous three articles did not focus on these factors driving use, which made the fourth article essential to the testing of the theoretical framework describing reasons for performance information use and non-use. From the tested factors associated with performance information use, speech length was a new factor driving performance information use which past studies have not investigated. As the objective was to assess the use of performance information by two sitting presidents, the study results cannot be generalized. Therefore, the results were preliminary and described the study objects and their use, not presidents in general. The results of the study offer hypotheses for future studies. The hypotheses are presented after the overview in which the primary results are explained.

After qualitative and quantitative content analysis, the study found frequent and extensive use of performance information while outcome information was the most used information type. For presidents conducting economic leadership, outcome information may be useful because it describes the state of the nation with indicators capturing socioeconomic development and well-being. Middle-aged female presidents with backgrounds in financial management and economics were active performance information users according to the results of the study. The examined presidents also operated in a similar political context with the same type of constitutional powers. These associations between individual attributes of the presidents, political context, and performance information use both support and contradict past findings. For example, it has been argued that political conflict increases performance information use (Askim, 2009). This argument was supported by the results of this article. However, research has linked higher age and education to low use (Askim,

2009), while this study provided contradictive results by associating frequent and extensive performance information use with high age and education levels.

According to the article, the examined presidents of Lithuania and Estonia also differed concerning performance information use despite their similar individual and contextual characteristics. The indication of this is that the level of similarity between the presidents was not adequate, and distinctions in individual characteristics relate to the dissimilarities found in performance information use. Longer political career, different professional career, older age, and higher education as demographic attributes were associated with lower use in this study. Of course, it is also possible that distinctions in contextual factors, such as economic development, political conditions, and socio-demographic aspects, can explain why the two presidents used performance information differently.

The article also demonstrated how there were significant disparities within performance information use categorized as extensive. This highlights that it might be important to use sensitive measurement techniques that can spot these kinds of differences. When the speech length was controlled, some variations vanished, while others remained. In general, performance information use correlated with speech length in both presidents' speeches. By sampling the most common words used in the English language and calculating this ratio between numbers and other words, the article demonstrated that the ratio between numbers and other words in the English language is fairly constant, which would indicate the following: More numbers in speeches would mean more words in public talks.

The research evidence in the article supported the already-established idea that performance information use can vary depending on the use context and users (e.g., Johansson & Siverbo, 2009; Giacomini et al., 2016). According to the article, performance information is being used more in conference contexts than ceremonies when the two presidents were observed separately. This study proposes the addition of the speech context, a new contextual factor, to the theoretical framework used in this thesis. Under the category of the speech context, divergent types of context incentivize use or non-use of performance information. However, perceived dissimilarities in performance information use between the presidents could not be explained by the speech context as these dissimilarities existed even after the speech context was controlled. Speech length explained some of these differences. The findings of the article suggest that performance information use models should consider the inclusion of speech length as important variables capable of influencing use. Overall, when combining the results of article three and those of past studies, the following hypotheses were formulated:

- H1: Presidents focus on outcome information more than other types of performance information;
- H2: Presidents are active users of performance information;
- H3: Higher education is associated with high performance information use when education relates to economics or business studies;
- H4: Longer political career is associated with lower performance information use;
- H5: Higher education relates to lower performance information use;

- H6: Higher age relates to lower performance information use;
- H7: Females are more active performance information users than males;
- H8: Performance information is used less in ceremonies than lectures and conferences;
and
- H9: Speech length correlates with performance information use.

The fourth article of the dissertation suggests useful novel research methods useful to future research that evaluates factors that drive performance information use. For example, the study suggested the use of conceptual frameworks that would enable an evaluation of performance information use intensity in the written and spoken language. Moreover, the study developed three methods to study the intensity of use. With the first proposed method, it was possible to examine how many of the speeches included performance information to determine intensity. According to the second method, intensity of use can be identified by observing the extent to which the speeches' content comprised performance information. The third method was used to examine whether certain types of performance information were utilized more frequently than others.

The way intensity is measured affects an individual's perception of the factors that drive high performance information use, for example. Thus, evaluating the intensity of performance information use should be considered when examining the factors that elucidate performance information use. It seems that most of the studies in the field that used surveys left it to the respondents to define the intensity or frequent use of performance information. As this article showed, it is not easy to define intensity of use. This may have affected the results of past surveys and therefore our current understanding of factors that explain performance information use. However, the interpretative frameworks for performance information use intensity suggested in this article were not intended to be normative as the definitions in the frameworks were debatable. Instead, the frameworks can be used to initiate conversations on the topic and address the definition of the intensity of performance information use. This type of conversation is much needed in the field of study on factors that incentivize performance information use, especially since the current research provided very few concrete definitions of high- or low-performance information use (see, e.g., Sterck & Scheers, 2006; Raudla, 2012; Buylens & Christiaens, 2016).

As a second methodological contribution, this article developed measurement methods to perform a comparison of performance information use by two actors. With these methods, future research can compare two actors in terms of the performance information used, how information is used in terms of intensity, and how its use is explained. The article also suggested an analytical model developed to utilize Student's *t*-test, Fisher's exact test, the χ^2 test, and the Kruskal–Wallis H test to compare the performance information use of two speakers. As the study only evaluated two presidents, further testing is warranted in future studies to validate the findings.

9 Conclusion and discussion

This thesis attempted to answer the following research question: What reasons stimulate the use and non-use of public sector performance information? To provide an answer to this research question, a scoping literature review on the reasons for information use in the fields of accounting, economics, psychology, and political science was conducted. Moreover, a systematic literature review (e.g., Mauro et al., 2017) focusing on performance information use in the field of public administration and management was performed. The goal of these reviews was to categorize past findings under several new umbrella concepts that could be tested in the research articles. As a result, the thesis created a novel theoretical framework describing the factors motivating the use and non-use of performance information (see appendixes 1–7 or the list in pages 29–30). The umbrella concepts describing factors driving use and non-use were the following: demographic attributes (e.g., Askim, 2007), mental models (e.g., Kroll, 2015a), power (e.g., Cavalluzzo & Ittner, 2004), social pressures (Raudla, 2015), information provider (e.g., Pollitt, 2006), information channel (e.g., Rajala et al., 2018), and information (e.g., Lu et al., 2015).

The theoretical framework resulted from an interdisciplinary approach that utilized research on accounting (e.g., Chewning & Harrell, 1999), economics (e.g., Akerlof, 1995), psychology (e.g., Taber & Lodge, 2006), political science (e.g., Weaver, 1986), and public administration and management (e.g., van Dooren, 2006). It also provided an answer to the first sub-question, as follows: How does an interdisciplinary perspective help one to understand the reasons for performance information use in the field of public administration and performance management? The thesis found that reasons seen in other disciplines (e.g., Jones & Dewing, 1997; Epley & Gilovich, 2016; Watkins & Terrell, 1988; Calvert, 2013) covered well the findings in the field of public administration and management and operated as umbrella concepts presented above. However, the public administration and management studies provided more in-depth insights to performance information use and non-use because they supplied several subcategories to the umbrella concepts found, such as mental models.

The presented theoretical framework proposes changes to concepts used in past research and to their content. For example, research has used various concepts when individual attributes have been addressed (e.g., Saliterer & Korac, 2014; Charbonneau & Nayer, 2012; Andrews, 2006). This study suggests that these concepts used in the literature fall under three umbrella concepts: demographic attributes, mental models, and power. Similarly, past studies talk about various contextual factors (e.g., Bourdeaux, 2006; Moynihan & Hawes, 2012; Abdel-Maksoud et al., 2015). However, here, it is argued that all these contextual factors can be placed under four umbrella concepts: social pressures (e.g., external environment, organizational factors) information provider, information channel, and information. All four articles of this article dissertation provided support for the theoretical framework generated in this thesis and demonstrated that it is applicable to three empirical settings and one theoretical study. Thus, it can be used in theoretical developments. Moreover, the performance information use of politicians, oversight bodies, public managers, private-sector and third-sector actors, service producers, donors of funds, citizens, and the media could be studied using this theoretical framework as a starting point, providing options and possibilities for what to study. The theoretical framework was also tested in different policy sectors (i.e., healthcare, social care, education and technical services). This synthesis also tried to understand what more can be known from the theoretical framework describing reasons to use performance information based on the articles presented in the synthesis.

The main implication of the study and its results can be summarized in one argument that also answers the second research sub-question asking what the characteristics would be for a more holistic theory of performance information use and non-use in the field of public administration and management. According to this argument, it is possible to create a more holistic theory from the factors driving performance information use, as the current theory is underdeveloped in the field of public administration and management because:

1. The coverage of important factors explaining use is lacking in the research
2. The ambiguity relating to factors driving use and non-use is mostly ignored
3. How factors are conceptually embedded into each other or relate to one another is largely neglected
4. A theory of non-use is generally missing

9.1 The coverage of important factors

The theoretical frameworks in past studies have been selective in terms of what possible drivers of performance information have been included and excluded (e.g., Askim, 2009; Lu & Willoughby, 2015; Rajala et al., 2018). Comparing the results of this study (see appendix 1–7) to the past theories (e.g., Moynihan & Lavertu, 2012; Buylen & Christiaens, 2016; Jorge et al., 2016) reveals that some individual attributes (i.e., power, different mental models, and demographic attributes) and/or some of the aspects from the context use (i.e.,

social pressures, organizational performance, information provider, information channel, and information) have been neglected. Theories have been lacking (e.g., Bourdeaux, 2006; Ho, 2011) because they have not been based on such an extensive literature review as the framework introduced here. Thus, the theoretical framework presented in this thesis is more comprehensive than its predecessors (e.g., Dull, 2009; Raudla, 2012), and it challenges the validity of the previous theoretical models (e.g., Askim, 2009; Bjørnholt et al., 2016) that have more caveats in their coverage of important factors affecting use. The exclusion of even a single important factor from the study design substantially increases the risk of omitted variable bias; similarly, limiting the focus to a few key factors, for the sake of precision, is associated with the danger of describing a false picture in detail. Indeed, the introduction of omitted factors could have altered the results in past studies (e.g., Hammerschmid et al., 2013; Kroll, 2015b). Uncertainty over the meaning of past results is unfortunate as it is possible to measure the effects of demographic attributes, mental models, power, social pressure, performance information providers, channels, and quality of performance information in one study. The measurement does not have to capture all aspects of, for example, the mental models, but it would be good to consider at least one relevant mental model in the study design.

The omission of variables is difficult to avoid, and this frequently occurs in research design. Nonetheless, acknowledging this does not mean that omitted variables should be ignored. Instead, existence of omitted variables emphasizes the logic of research design (Clarke, 2005). In research design, natural experiments could be preferred, and the omitted variables could be taken into consideration during the sampling phase (Hanushek and Jackson, 1977). When applying data collection methods, dynamic survey forms could be designed that enable the respondents to name, for example, the type of social pressure that affected their use of performance information; the analysis could achieve depth similar to that seen in current studies.

Other advice, given in relation to past findings, is to test broad theories in narrow, focused, and controlled circumstances (Clarke, 2005). According to Rosenbaum (1999), broad theories are useful as they are able to make predictions across different contexts. In line with this thinking, the current thesis argued that studies, even small or mid-range studies, should attempt to measure or at least consider in their research design how performance information use is affected by the following factors: demographic attributes, mental models, power, information providers, information channels, features of information, and social pressure. If consensus is lacking among scholars that the factors presented in this dissertation are conceptually and theoretically important, then at least serious discussions can be held regarding the theory of performance information use. The creation of a theory based on the current theoretical hypotheses would provide a common reference point in the field so that the results in different contexts could be reflected.

The present studies are theoretically interesting, but at the same time, they are often very disconnected from the theoretical perspective. This thesis was one attempt to start

the discussions about the broad theory of performance information use that will provide testable predictions in different parts of the world. Here, the theoretical framework proposed in this thesis is a suggestion for a broad theory of performance information use. More debate on the topic is warranted in future so that the ideas presented here can be validated.

To move away from overly simplistic theories that are lacking in terms of coverage, the theoretical framework presented in this study is a step forward because it covers the findings from past studies (see appendixes 1–7) while expanding the current theoretical understanding to new research areas. The articles attempted to further develop and enrich the understanding of the theoretical framework by providing more nuanced information to already identified factors (e.g., motivational mental models) or expanding the framework to new dimensions by introducing new factors, such as mental models on blame avoidance.

The proposed theoretical framework provides a point of reference for performance information use studies because it helps scholars to locate key concepts (i.e., factors that drive use) and assists critical studies to address relevant topics. To illustrate the implications of common reference point, let us consider academics who conduct small-scale studies and wish to focus on the role of organizational factors in performance information use. The theoretical framework proposed in this thesis could reveal to scholars operating in the micro context that other studies have investigated variations in performance information use related to task structures (e.g., Rabovsky, 2014), management structures (e.g., Charbonneau & Bellavance, 2015), the integration of structures (Jääskeläinen & Roitto, 2014), reformed structures (Kloot, 1999), and the legitimacy of structures (Poister & Streit, 1999). The framework also demonstrates that it would be useful to consider other factors when studying the role of organizational factors in performance information use, such as mental model and power. Through familiarization with the framework and its ideas, scholars performing micro research can aggregate the research results and accumulate their knowledge after taking into consideration seven categories (i.e., demographic factors, mental models, power, social pressure, information providers, information channels, and features of information). Specifically, the framework could help to identify the research traditions to which scholars conducting small-scale research on the topic of performance information use are contributing.

With respect to word limits in books and academic journals and problems in theoretical coverage caused by the limits, the establishment of a theory of performance information use as a common reference point could free up space for other issues to be addressed in studies because this would lessen the need to justify each theory used in a particular study. Other scholars could conduct extensive literature reviews on the topic to provide counterarguments to the proposed theoretical framework. Literature reviews are useful for mitigating the challenges of word limits as they gather together what is already known, and scholars can easily refer to the literature reviews. They also enlighten scholars operating in

the field by summarizing a vast amount of literature and creating a common understanding of results in the field.

For practitioners, the theoretical framework is an archive of possible drivers of performance information use and non-use. This archive of drivers can be used to find ideas concerning how to increase performance information use in the organization, or it can be used to stimulate non-use if this is desired for some reason. However, the conceptual work remains far from ready. Future studies could further develop the concept system by adding levels of hierarchy below the current ones and by examining the conceptual relationships between factors.

9.2 Ambiguity in factors driving performance information use

The research process in this thesis showed how complex and ambiguous the umbrella concepts, such as mental models, social pressures, and information, are in terms of content (see appendixes 1–7). Indeed, academic scholars have addressed many types of things that can be placed under the umbrella concepts (e.g., Charbonneau & Nayer, 2012; Saliterer & Korac, 2014). This demonstrates ambiguity. The ambiguity itself has consequences for research results. For example, even if scholars have defined their concepts clearly in studies (e.g., Saliterer & Korac, 2013), it is still questionable whether the interviewed persons or survey responders understand the concepts similarly given the complexity and ambiguity of the concepts. As an example, it is far from certain that scholars and practitioners are talking about the same aspects of organizational structures when they address how such structures influence performance information use (e.g., Kroll, 2015a). Survey respondents can understand, for instance, organizational structures in multiple ways, as scholars have also understood the concept in various ways (see appendix 3–5 to see how scholars have approached organizational structures in their research). This raises the difficult question of what has been actually said when a scholar or practitioner is reporting that organizational structures do or do not affect performance information use. Are all the different notions of organization structure affecting the use or not according the respondent or scholar reporting the results? Again, ambiguity is very much present in factors driving performance information use and should be considered if the attempt is to create more holistic theory on performance information use.

The conceptual complexity and ambiguity are something that past research has neglected to a large extent. Therefore, the results of this study challenge the use of overly simplified concepts used in past studies, which are designed to narrow down the focus rather than see the richness in these concepts (e.g., Andrews, 2004; Amirkhanyan, 2009). The thesis attempts to create a change in the way concepts are seen because these seemingly clear concepts create an illusion of mutual understanding among practitioners and scholars in situations where the concepts are ultimately ambiguous. With the conceptual work conducted in this thesis, the author aims to highlight the ambiguity in all the basic concepts

used to describe the reasons for performance information use and non-use in the field of public administration and management (see appendixes 1–7 to confirm the ambiguity). Mental models, demographic attributes, power, social pressures, information providers, information channels, and information as drivers of performance information use seem to be very ambiguous concepts, as there are, for instance, many types of mental models.

Ignoring conceptual ambiguity makes scientific communication disconnected and unorganized, making it difficult to understand the current knowledge on the topic. As interpretation depends on concepts that ensure their communicability (Novak, 1966), this is a problem in current research. For practitioners, ambiguity creates challenges for developing performance information use in public administration because vague concepts provide opportunities for misunderstanding and incorrect reforms resulting from these misunderstandings (e.g., Vakkuri, 2010). For example, concentrating only on management structures while neglecting the integration of organizational structures can blind public managers to seeing how the excellent or poor integration of structures affects performance information use. This can create misunderstandings and adverse reforms based on these misunderstandings.

9.3 Conceptual relationships between reasons for performance information use and non-use

All four articles also provided examples of how the diverse theoretical factors interplay with the whole theoretical framework and vice versa. For example, the first article addressing why output information might be preferred over outcome information in some situations demonstrated how mental models interacted and were intertwined with the whole theoretical framework consisting of demographic factors, power, social pressures, information provider, information channel, and information. This type of investigation enriches our understanding of the interplay between the factors driving the use and non-use of performance information, as past studies have not examined these in depth (e.g., Grossi et al., 2016; Giacomini et al., 2016; Van Helden, 2016).

The articles also provided examples of how intertwined these concepts are in some empirical and theoretical cases. For example, mental models embedding other reasons for performance information use show how performance information use co-occurs and is intertwined. Table 4 demonstrates how the factors driving performance information use are often embedded into each other and, therefore, co-occur in real life. The fact that reasons for performance information use do co-occur and are embedded into each other raises questions whether it is possible to separate these factors from each other in practice and, if so, what are the circumstances in which the factors are not intertwined? Another important but unaddressed question arising from the co-occurrence is whether practitioners can report the intertwined nature of the reasons when researchers ask them

about the rationales driving performance information use (e.g., Askim, 2007; Johansson & Siverbo, 2009). It is possible that responders only report what individual factors have promoted their performance information use, and they have not considered how factors are actually embedded into each other. If respondents are not aware of how factors are intertwined, they report their perceptions on how these things are not embedded into each other according to their view.

It is also unclear how well academics are aware of how embedded into each other the factors driving performance information use are. Currently, quantitative studies concentrate on reporting how organizational culture or some other factor all by itself drives performance information use when other separate factors are being controlled in statistical analysis and covariances are being checked (e.g., Melkers & Willoughby, 2005; Taylor, 2011). The question here is, how does one, for instance, control the effects of political power on organizations' structures and performance information use if organizational structures are manifestations of political power? The statistical methods used in current studies (e.g., Moynihan & Landuyt, 2009; Folz et al., 2009) cannot detect how factors are embedded in each other unless respondents understand how things are intertwined and report that every factor embedded in, for example, social pressures is also affecting use in addition to social pressures. Researchers using qualitative or quantitative approaches also may not ask anything about the co-occurrence and how the respondents can recognize factors embedded into each other (e.g., Berman & Wang, 2000; Rajala et al., 2018). Generally, the problem with not identifying the intertwined nature of reasons driving use is the following: using simplified concepts that are artificially separated from each other can inhibit one from seeing relationships between reasons of performance information use.

Understanding how incentives are intertwined makes the hunt for particular drivers of performance information use questionable processes in some instances. The question here is what we learn from the statements, noting that organizational culture is associated with performance information use but mental models are not (e.g., Saliterer & Korac, 2013). These types of statements are important if one is interested in how respondents understand drivers of performance information use in a particular context. However, such statements are misleading and oversimplifying if one is seeking to understand how the factors are actually intertwined and operate as packages that form a system where some factors linked with performance information use might encourage or discourage use and non-use while other factors are just allowing use or non-use but not really encouraging or discouraging it. For example, an organizational structure may not encourage or discourage performance information use, but it allows other factors, such as mental models, to drive performance information use. Future changes in the allowing factors could change performance information use habits in the organization, although these are not encouraging or discouraging use currently. This makes allowing factors important for performance information use.

Focusing only on factors driving performance information use neglects the relationships between allowing factors and performance information use and non-use (e.g., Sinervo & Haapala, 2019). These allowing factors can turn into discouraging or encouraging factors in varying situations. Thus, this thesis attempts to challenge the view that the focus of academics and practitioners should only be on discouraging and encouraging factors incentivizing performance information use (e.g., Moynihan et al., 2012; Kroll & Vogel, 2014). Recognizing allowing factors would alter how scholars and practitioners see relationships between performance information use and factors influencing such use. Currently, the allowing factors associated with use and non-use are rather latent elements in the literature that must be inferred by the reader. Studies do not currently ask how performance information use might change if allowing factors were changed (e.g., Ho, 2006).

In the factor packages, the discouraging or encouraging factors form more or less from different mental models, demographic attributes, power, and factors forming the performance information use context, be they laws, stakeholder pressures, matters relating to information quality, organizational structure, or culture (e.g., Kroll, 2015a; Laihonen & Mäntylä, 2018). In a similar manner, allowing factors can be mental models, demographic attributes, power, and factors forming the performance information use context. Therefore, separating, for example, mental models from the quality of information in factor packages is an artificial maneuver that can create the illusion that mental models describing what information is do not relate to performance information use at all (e.g., Saliterer & Korac, 2013). This impression is false, as information is essentially a mental model passed from the information sender to the information receiver. This thesis has been a though experiment aimed to help readers see how factors incentivizing performance information are layered on top of each other and how these factors conceptually relate one another. Indeed, seeing how stakeholder pressures or organizational culture is present in mental models enables one to see relationships between these elements.

Following from the above, this thesis argues that more holistic theory describing reasons for performance information use would consider how factors are intertwined and would not try to artificially separate these factors from each other in situations where factors are clearly embedded into each other. Table 4 is an attempt to move into more holistic theory about the reasons for performance information use. It identifies the intertwined nature of the reasons driving use. Moreover, more holistic theory would accept the idea that there are also allowing factors, not just encouraging or discouraging factors driving use.

Table 4. Intertwined reasons for performance information use (i.e., the chicken-and-egg dilemma)

<i>Co-occurring reason</i>	<i>Education as a demographic factor</i>	<i>Mental model</i>	<i>Power</i>	<i>Social pressure</i>	<i>Performance information producer</i>	<i>Performance information channel</i>	<i>Performance information</i>
Education as a demographic factor	-	Education creates mental models in students' minds	Education provides capabilities and therefore provides power to do things	Education teaches socially acceptable behavior	Education teaches how to be information producer	Education teaches how to find and use information channels	Education teaches how to read information
Mental model	-	-	Mental models describe how power works in real life	Mental models describe what the social pressure is and what actions are socially desirable	Mental models describe what is expected from information producers	Mental models describe how information channels work or should work	Mental models guide how to interpret information
Power	-	Use of power changes ideas about how world works	-	Power is used to create social pressure toward individuals	Power is used to determine who can act as information producer	Power determines who has access to certain channels and what channels can operate	Power defines the content of the information by guiding information production
Social pressure	-	Social pressure conveys ideas about what is socially accepted conduct	Social pressure attempts to use power over those who deviate from social norms	-	Social pressures can determine who can be information producer	Social pressures push people to use specific information channels	Social pressures affect what type of information is socially acceptable

Continues on the next page

Table 4. continues

Performance information producer	Producers determining what is to be taught in education	Information producer shapes and creates mental models by producing information	Information producer uses power to determine what or how information is being produced	Producer can create social pressure by publishing blameworthy information	-	Information producer is part of the information channel as the sender of the information	Information producer defines the content of the information by choosing what is being produced
Performance information channel	Education system as information channel	Information channels determine what mental models can be accessed and what cannot	Information channel provides power to access information	Information channels need to work so social pressures can be transmitted	Information channels determine producers that are connected to information users	-	Information channels set restrictions on the content of the information
Performance information	Information determining what is being taught in schools	Information conveys thoughts on how world works or has worked	Information has power to change behavior	Information delivers knowledge about social pressures	Information about the producers affects the producer's ability to provide information	The type of information affects what type of information channel works	-

9.4 Theory of non-use

As a conceptual contribution, this study has argued that performance information non-use as a concept needs to be more broadly recognized in the field. Non-use is currently less important than use (e.g., high or low use), if the past research settings reflect importance (e.g., Behn, 2003; ter Bogt, 2004; Van Dooren & Van de Walle, 2011; Moynihan & Lavertu, 2012; Angiola & Bianchi, 2015). Conversely, non-use as a research topic has been standing in the shadows far too long. To understand performance information use, it is helpful to comprehend what factors stimulate non-use or the purposes of performance information non-use. As most studies ask what drives performance information use (e.g., Moynihan & Ingraham, 2004; Askim, 2011; Kroll, 2015b), the theory of performance information non-use or models of non-use are currently underdeveloped. This is a shortage in the current research that needed fixing. To make the theory of reasons driving performance information use more comprehensive, non-use must be included in it and studied more. Therefore, the theory of non-use was connected in this thesis to the generated theoretical framework describing the drivers of performance information used to expand the theory.

This thesis tested the theoretical framework in empirical and theoretical settings designed to understand non-use. Indeed, the framework provided a fertile starting point in understanding non-use because it enabled one to see how individual attributes (i.e., mental models, power, and demographic attributes) and factors in performance information use context (social pressures, information provider, information channel, and information) related to non-use. However, the articles in this study expanded the theoretical framework created in the literature review by reporting many new reasons explaining non-use or limited use compared to previous studies (e.g., Johnsen, 2011).

The article addressing why output information may be used instead of outcome information displayed the idea that there exists a performance information hierarchy that can explain non-use in certain situations. This performance information hierarchy reveals why certain information is considered worth using while some other information about performance is not. The hierarchy also offers insights into why non-use can be a rational act from the performance information user. Indeed, the non-use of some types of performance information often occurs because it makes more sense to use other types of performance information in, for instance, decision making. Concerning the stability of the performance information hierarchy, this thesis argued that the hierarchy is often context-specific, meaning that different types of performance information switch places in the hierarchy depending on the situation where the information use is considered (see also Ho & Ni, 2005; Johansson & Siverbo, 2009). Changes in the performance information hierarchy indicate that scholars should focus on windows of opportunities showing situations where a specific performance information becomes relevant.

Besides the performance information hierarchy, the first article and third article addressing blame avoidance revealed that non-use also relates to factor packages driving non-use. Therefore, understanding encouraging and discouraging, as well as allowing

factors and the relationships between these factors is important for non-use as well. To support the use of performance information, one needs to be able to solve problems related to factors bundled into packages incentivizing non-use. Partial solutions considering some factors on the packages while neglecting others might not inhibit non-use, as factors are intertwined, as shown in Table 4. This makes studying the ideas presented in Table 4 highly relevant in non-use situations. However, solving non-use is not necessarily valuable and mandatory if non-use is actually beneficial for the activities of the public sector and to the well-being of the citizens. This thesis showed that performance information non-use can be valuable if one wants to avoid blame from poor performance. Hood (2011) has noted that blame avoidance can be useful in the public sector.

In situations where non-use is rational, knowing the factor packages driving non-use helps to see why non-use is considered valuable. To think that non-use would be somehow valuable is a novel and even radical idea in the field of public administration and management because the current research seems to have this underlying assumption that only use is valuable and non-use is a waste of resources. Indeed, this assumption declaring non-use as a waste is vividly presented in the quotation presented at the introduction of this thesis. Let us recap this quotation taken from Talbot (2005):

If much performance reporting has not, in practice, been meant to change anything why has it been done? If the audience is inattentive, the directors uninterested and the actors apathetic, what's the point? In one sense this can be seen as a classic "implementation failure."

To make a counter-argument to Talbot's (2005) notion and the general assumption made in the field, this thesis claims that non-use can be valuable, as using information has opportunity costs that can be too high on some occasions. Blame avoidance is an example of how the non-use of poor performance results can save politicians re-election (e.g., Weaver, 1986). This is a reflection of the "dark" side of non-use performance information as a means of serving political self-interest. Nevertheless, the public can also gain value from non-use as every piece of information is not useful in every situation (e.g., Montesinos et al., 2016).

To understand the value of non-use, it is useful to recognize that it can be either temporary or permanent. In addition, non-use relates to actors, which indicates that the use and non-use of the same performance information could occur simultaneously because one user might use the information, and another might not do so. If all users were committed to the permanent non-use of particular performance information, then information that is not used would reflect a wasteful use of resources in performance measurement. The removal of measures that produce information that no-one ever uses would be valuable as it would result in future savings as a consequence of the termination of measurement. In this way, non-use performance information can reveal which measures are redundant and wasting future resources. Thus, non-use of performance information could become a valuable resource allocation tool, similar to the way that not buying certain products guides

market allocation to achieve a more efficient solution. Hence, the value of non-use is its ability to produce future savings in information production. However, care must be taken with halting the production of performance information as terminating measures would inhibit the option of obtaining information from the shelf when it is needed (Johnsen, 2011). Hence, the removal of the option of being able to use information could have severe consequences for public sector activities.

In situations where the use and non-use of a particular performance information occurs simultaneously, both the use and non-use can be a sign of the efficient division of labor. The central idea behind labor of division is that everybody does not do the same thing, which enables specialization and efficiency within organizations. If public managers everywhere do not need to read the same performance information, this would save time and money, and the organization could still utilize all the performance information efficiently and rationally. Well-designed non-use saves organizational resources, thus becoming valuable. In designing who needs to use specific performance information, it is equally important to consider who does not need to use it. The value of non-use of performance information is easy to overlook as many people automatically identify information that is and is not relevant to their task. However, it can become costly if public managers cannot identify which information is irrelevant to the extent that non-use of performance information is justified. In psychological and behavioral economics, there are numerous examples of the harmful effects of information overloading in decision-making. To avoid information overloading, mastering non-use performance information is a valuable skill in working life. Lastly, non-use of performance information can be invaluable when misinformation or disinformation is involved in decision-making because non-use protects from decision-making errors caused by use of misinformation of performance information. Using accurate performance information is more likely to lead to better decision-making, thereby resulting in future value, compared to a situation where the decision is based on misinformation about performance, for example. Thus, the value of non-use is an aspect that is missing from current theories on performance information use. Therefore, the suggestion is that, to ensure that the performance information theory is comprehensive, it should cover the value of non-use of performance information.

Ignorance of the importance of non-use is understandable if the rational and technocratic ideas that motivate a large part of the research on performance management are adopted. According to technocratic and rational ideals seen in the literature, an increase in performance information use by government makes the public sector more efficient, and effective in service production and more rational in decision-making because decisions are more informed. By proposing that temporary non-use and even permanent non-use by certain actors is rational when exercised in a highly skilled manner, this study suggests that the most rational government is not the one where every actor uses every possible piece of performance information the maximal amount. A government setting where everyone uses the performance information to the maximal amount is likely to reflect a

wasteful government characterized by poor decisions caused by an information overload. Indeed, what is proposed is that marginal utility in performance information use should be considered on a case by case basis. The efficient labor of division, not maximal use, should be sought regarding performance information use. Of course, the assumption is that people do not invent a decision-making “machine” that considers all performance information and understands everyone’s preferences and values. Provided that people perform decision-making, the rationality of temporary non-use and even permanent non-use by certain actors should be taken seriously.

The study addressing performance dialogues in a hybrid organization highlighted inter-organizational factors associated with non-use, whereas the article focusing on presidents as performance information users displayed non-use in specific speech contexts. In the hybrid setting, performance information users were public and private sector managers. Presidents, though, represent politicians. In attempts to understand non-use, there is a growing need to study the topic in different contexts and among dissimilar user groups, be they citizens, oversight bodies, or funders. This study has contributed to the investigations addressing the non-use of various user groups in different contexts and argued that a more holistic theory of performance information use would consider non-use in a wide variety of contexts and among a diverse set of performance information users. It is also likely that ambiguity is tied to concepts describing non-use in a similar manner ambiguity was found in the concepts describing factors driving use. An improved theory of performance information use would consider this ambiguity. In conclusion, this study has contributed to the field of public administration and management by providing models of performance information non-use that can be further tested in future studies.

9.5 Limitations of the thesis

Because the articles conducted as part of this thesis included explanations of their limitations, this section focuses on the limitations of the synthesis and the whole thesis. The scoping review conducted in the study fields of accounting, economics, psychology, and political science did not examine grey literature, unpublished works, or studies written in other languages than English. For this reason, some viewpoints may have been left out. As the search method is not systematic (e.g., Levati, Campbell, & Frost, 2016), the studies included in the scoping review do not represent the population of studies from which the chosen studies were taken. Thus, the results of scoping reviews cannot be generalized in the same manner as the conclusions drawn from systematic reviews can be generalized. Scoping reviews are often used to examine the main sources of evidence available (e.g., Kidholm & Kristensen, 2018). Therefore, this study applied a scoping review.

When considering the systematic literature review conducted in the theoretical section, there were some limitations. First, the literature review was designed to examine studies published only in academic books and international academic journals, which limits some

perspectives (e.g., Mauro et al., 2017). For example, there are many relevant studies to the theoretical framework created in this research that fell outside the scope of the systematic review as they were either published in non-academic journals (e.g., McPhee, 2005) or official documents written by practitioners. Second, language limitations and restrictions in the time frame and databases might have also left out some valuable insights on the topic. Because inferring the categories of reasons driving performance information was based on an inductive approach, the validity of the logic behind the categories can be confirmed from appendixes 1–7. As a limitation, the categorization of verbal descriptions is rarely straightforward, as words can have many meanings to people (e.g., Empson, 2004). It is, therefore, possible that some other researcher would have placed some factor driving performance information use under different category. This makes the categorization debatable to some extent, but the author of this thesis can defend the categorization, as every factor categorized has its justification.

Incompatible results of qualitative and quantitative approaches are typical problems for mixed methods (e.g., Johnson & Onwuegbuzie, 2004), but this thesis did not involve any incompatible results. However, all the articles were based on sampling methods that do not permit the generalization of results (e.g., Patton, 2014). The thesis attempted to fix this problem related to generalization by adopting a comprehensive literature review approach so the reasons for performance information use and non-use could be generalized to apply to the whole study field of public administration and management. As a lack of coverage, ambiguity in reasons for use and non-use, intertwined reasons, and an omitted theory of non-use were confirmed from both the systematic literature review and the articles, the main conclusions of this thesis can be generalized.

9.6 Where to go from here? Topics for future researchers

This study opened several new research paths for future studies by exploring new research areas in public sector performance information use and developing novel research methods. The proposed theoretical framework and the hierarchy of performance information suggested in the first article of this thesis need further validation. This would reveal windows of opportunity for types of performance information able to influence public sector performance management. Studies could examine how a conflict-oriented context affects incentives to utilize outcome information and identify in which contexts output measures are adequate proxies of outcome measures.

Aside from the hierarchy of performance information, the non-use and use of performance information could be further explored in different contexts and among assorted user groups while considering the factor packages and ambiguities related to each factor driving use. In particular, comparative studies could be used to fulfill the research gap relating to the question that addresses differences in performance information use in different hybrid settings (e.g., Hodges, 2012). Non-use should be further investigated

within purely public sector service production, not just in hybrid organizations (e.g., Rajala et al., 2018). Large-scale surveys could be conducted to identify factors that contribute to the non-use of performance information and whether these same factors explain high performance information use in hybrids and the public sector.

Performance information use in rhetoric is another topic on which more research is required. There are many unexplored questions and potential for machine learning applications in the rhetorical analysis of performance information use (e.g., Buylen & Christiaens, 2016). Machine learning could be especially useful when conducting extensive studies on local, regional, and central governments. The type of performance information used in political argumentation and taking place at different governmental levels could be evaluated, along with an exploration of any differences between the levels. It might also be possible to conduct classical argumentation analysis at all levels of government to determine whether performance information is used in different roles in argumentation. To expand current knowledge of performance information use by presidents, studies could compare the operations of presidents in different continents. The coronavirus disease 2019 pandemic also offers opportunities to study the role of performance information use in crisis rhetoric. The methodological contributions provided in this thesis offer a starting point from which scholars can progress in rhetoric analysis using the suggested research proposals.

Blame avoidance, as a driver of performance information use, should also be further assessed. Non-use of performance information arising from blame avoidance is a largely unexamined topic, as is performance information use that has the objective of avoiding blame. For example, the identification of the type of blame avoidance strategies used to increase non-use of performance information and use could be investigated by scholars. Another unexamined question is whether public managers and politicians are risk averse, loss averse, risk neutral, or risk seeking when it comes to deciding whether or not to enable outcome information use in public sector.

The author acknowledges that the proposed theoretical framework and methods could be improved. Researchers could expand upon the current knowledge of unexplored demographic attributes, mental models, power, social pressure, information providers, features of information, information channels, and performance information. Reasons for the ambiguity associated with drivers of performance information use could be evaluated further, along with the intertwined reasons that drive use of performance information use in the public sector. For example, interviews could be used to ask whether the interviewees recognized the ambiguity relating to the reasons for use and non-use of performance information use. Compared to surveys, the use of interviews would better facilitate understanding of the ambiguity because they make it possible to react to interviewee responses using different questions to reveal how and whether the ambiguity related to factors driving use is understood in practice. An action research approach could be used to study how the theoretical framework could be applied to problem-solving to address the

issue of limited information use. This could operate as a test to measure the practical value of the framework. Lastly, to improve performance information use in the public sector, there is considerable need to gain deeper insight into the value of information non-use. Do practitioners see value in information non-use, and, if so, what type of value is produced as a consequence of information non-use? Currently, these two questions remain unanswered, along with several others that were alluded to in this thesis. Therefore, I conclude this study by inviting you to consider the questions that I have proposed. Honorable scientific community, there is work to be done.

References

- Abdel-Maksoud, A., Elbanna, S., Mahama, H., & Pollanen, R. (2015). The use of performance information in strategic decision making in public organizations. *International Journal of Public Sector Management*, 28(7), 528–549. <https://doi.org/10.1108/IJPSM-06-2015-0114>
- Agostino, D., & Arnaboldi, M. (2015). The new public management in hybrid settings: New challenges for performance measures. *International Review of Public Administration*, 20(4), 353–369.
- Agostino, D., & Arnaboldi, M. (2018). Performance measurement systems in public service networks. The What, Who and How of Control. *Financial Accountability & Management*, 1–14.
- Akerlof, G. (1995). The market for lemons: Quality uncertainty and the market mechanism. In S. Estrin & A. Marin (Eds), *Essential readings in economics* (pp. 175–188). New York: St. Martin's Press.
- Akerlof, G., & Kranton, R. (2010). *Identity Economics*. Princeton, NJ: Princeton University Press.
- Alexa (2019). *Top sites*. Retrieved from <https://www.alexa.com/topsites>
- Amirkhanyan, A. A. (2011). What is the effect of performance measurement on perceived accountability effectiveness in state and local government contracts? *Public Performance and Management Review*, 35(2), 303–339. <https://doi.org/10.2753/PMR1530-9576350204>
- Amirkhanyan, A. A., Kim, H. J., & Lambright, K. T. (2014). The performance puzzle: Understanding the factors influencing alternative dimensions and views of performance. *Journal of Public Administration Research and Theory*, 24(1). <https://doi.org/10.1093/jopart/muto21>
- Ammons, D. N., & Rivenbark, W. C. (2008). Factors influencing the use of performance data to improve municipal services: Evidence from the North Carolina benchmarking project. *Public Administration Review*, 68(2), 304–318. <https://doi.org/10.1111/j.1540-6210.2007.00864.x>
- Ammons, D. N., & Roenigk, D. J. (2015). Performance management in local government: Is practice influenced by doctrine? *Public Performance and Management Review*, 38(3), 514–541. <https://doi.org/10.1080/15309576.2015.1006461>
- Andersen, S. C., & Moynihan, D. P. (2016). How leaders respond to diversity: The moderating role of organizational culture on performance information use. *Journal of Public Administration Research and Theory*, 26(3), 448–460. <https://doi.org/10.1093/jopart/muvo38>
- Andrews, M. (2004). Authority, acceptance, ability and performance-based budgeting reforms. *International Journal of Public Sector Management*, 17(4), 332–344. <https://doi.org/10.1108/09513550410539811>

- Andrews, M. (2006). Beyond “best practice” and “basics first” in adopting performance budgeting reform. *Public Administration and Development*, 26(2), 147–161. <https://doi.org/10.1002/pad.401>
- Andriani, Y., Kober, R., & Ng, J. (2010). Decision usefulness of cash and accrual information: Public sector managers’ perceptions. *Australian Accounting Review*, 20(2), 144–153. <https://doi.org/10.1111/j.1835-2561.2010.00087.x>
- Angiola, N., & Bianchi, P. (2015). Public managers’ skills development for effective performance management: Empirical evidence from Italian local governments. *Public Management Review*, 17(4), 517–542. <https://doi.org/10.1080/14719037.2013.798029>
- Ansari, S., & Euske, K. J. (1987). Rational, rationalizing, and reifying uses of accounting data in organizations. *Accounting, Organizations and Society*, 12(6), 549–570.
- Ariely, D. (2008). *Predictably Irrational*. New York: Harper Collins.
- Aristigueta, M. P., & Zarook, F. N. (2011). Managing for results in six states: A decade of progress demonstrates that leadership matters. *Public Performance & Management Review*, 35(1), 177–201.
- Arnaboldi, M., Lapsley, I., & Steccolini, I. (2015). Performance management in the public sector: The ultimate challenge. *Financial Accountability & Management*, 31(1), 1–22.
- Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, 70, 1–70.
- Askim, J. (2007). How do politicians use performance information? An analysis of the Norwegian local government experience. *International Review of Administrative Sciences*, 73(3), 453–472. <https://doi.org/10.1177/0020852307081152>
- Askim, J. (2009). The demand side of performance measurement: Explaining councillors’ utilization of performance information in policymaking. *International Public Management Journal*, 12(1), 24–47. <https://doi.org/10.1080/10967490802649395>
- Askim, J. (2011). Determinants of performance information utilization in political decision making. W. van Dooren, S.T. van de Walle (Eds), *Performance information in the public sector. How it is used*. Palgrave Macmillan, Basingstoke (2008), 125–139.
- Ayers, R. S. (2013). Building goal alignment in federal agencies’ performance appraisal programs. *Public Personnel Management*, 42(4), 495–520. <https://doi.org/10.1177/0091026013496077>
- Baddeley, M. (2017). *Behavioural economics: a very short introduction* (Vol. 505). Oxford University Press.
- Baekgaard, M. (2015). Performance information and citizen service attitudes: Do cost information and service use affect the relationship? *International Public Management Journal*, 18(2), 228–245.
- Baekgaard, M. (2017). Prospect theory and public service outcomes: Examining risk preferences in relation to public sector reforms. *Public Administration*, 95(4), 927–942.
- Baekgaard, M., & Serritzlew, S. (2016). Interpreting performance information: Motivated reasoning or unbiased comprehension. *Public Administration Review*, 76(1), 73–82. <https://doi.org/10.1111/puar.12406>
- Bandy, G. (2011). *Financial management and accounting in the public sector*. Abingdon, OX: Routledge.
- Barbato, G., & Turri, M. (2017). Understanding public performance measurement through theoretical pluralism. *International Journal of Public Sector Management*, 30(1), 15–30. <https://doi.org/10.1108/IJPSM-11-2015-0202>

- Barnow, B. S & Heinrich, C. J. (2017). One standard fits all? The pros and cons of performance standard adjustments. *American Society for Public Administration Stable*, 70(1), 60–71, <http://www.jstor.org/stable/40469111>
- Barton, A. D. (2004). How to profit from defence: A study in the misapplication of business accounting to the public sector in Australia. *Financial Accountability & Management*, 20(3), 281–304.
- Bartoli, A. (2008). The study of public management in France. La spécificité du modèle français d'administration. In *The Study of Public Management in Europe and the US* (pp. 28–55). Routledge.
- Baumgartner, F. R., & Jones, B. D. (2015). *The politics of information: Problem definition and the course of public policy in America*. Chicago: University of Chicago Press.
- Behn, R. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 588–606. <https://doi.org/10.1111/1540-6210.00322>
- Bellé, N., Cantarelli, P., & Belardinelli, P. (2018). Prospect theory goes public: Experimental evidence on cognitive biases in public policy and management decisions. *Public Administration Review*, 78(6), 828–840. <https://doi.org/10.1111/puar.12960>
- Bergmann, A. (2012). The influence of the nature of government accounting and reporting in decision making: Evidence from Switzerland. *Public Money & Management*, 32(1), 15–20.
- Berman, E., & Wang, X. (2014). Performance measurement in U.S. counties: Capacity for reform. Performance Measurement: Building Theory, Improving Practice. *Public Administration Review*, 60(), 128–145.
- Becker, G. S. (1976). *The economic approach to human behavior*. Chicago: The University of Chicago Press.
- Bianchi, C., & Rivenbark, W. C. (2014). Performance management in local government: The application of system dynamics to promote data use. *International Journal of Public Administration*, 37(13), 945–954. <https://doi.org/10.1080/01900692.2014.944990>
- Bimber, B. (1991). Information as a factor in congressional politics. *Legislative Studies Quarterly*, 16(4), 585–605.
- Bingman, C. F. (2006). Proposals for improving GPRA annual performance plans. *Public Budgeting & Finance*, 26(2), 143–154.
- Bjørnholt, B., Bækgaard, M., & Houlberg, K. (2016). Does fiscal austerity affect political decision-makers' use and perception of performance information? *Public Performance and Management Review*, 39(3), 560–580. <https://doi.org/10.1080/15309576.2015.1137766>
- Bol, J. C., & Smith, S. D. (2011). Spillover effects in subjective performance evaluation: Bias and the asymmetric influence of controllability. *The Accounting Review*, 86(4), 1213–1230.
- Borgida, E., & Nisbett, R. E. (1977). The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology*, 7, 258.
- Boston, J. Martin, J. Pallot, J. & Walsh, P (1996). *Public Management: The New Zealand Model*. Auckland: Oxford University Press.
- Bouckaert, G., & Halligan, J. (2008). Comparing performance across public sectors. *Performance Information in the Public Sector*, August, 72–93. https://doi.org/10.1007/978-1-137-10541-7_6
- Bouckaert, G., & Halligan, J. (2008). *Managing performance*. London: Routledge.
- Bourdeaux, C. (2006). Do legislatures matter in budgetary reform? *Public Budgeting & Finance*, 26(1), 120–142.

- Bourdeaux, C. (2008). Integrating performance information into legislative budget processes. *Public Performance & Management Review*, 31(4), 547–569. <https://doi.org/10.2753/pmr1530-9576310403>
- Bourgon, J. (1998). Fifth Annual Report to the Prime Minister on the Public Service of Canada, Ottawa, Privy Council Office.
- Bourgon, J. (2008). Performance management: It's the results that count. *Asia Pacific Journal of Public Administration*, 30(1), 41–58. <https://doi.org/10.1080/23276665.2008.10779342>
- Bourmistrov, A. (2017). Mental models and cognitive discomfort: Why do users reject even a small change in a financial accounting report? *Pacific Accounting Review*, 29(4), 490–511. <https://doi.org/10.1108/PAR-11-2016-0109>
- Boyne, G., Gould-Williams, J., Law, J., & Walker, R. (2002). Plans, performance information and accountability: The case of best value. *Public Administration*, 80(4), 691–710. <https://doi.org/10.1111/1467-9299.00324>
- Bozeman, B. (1987). *All organizations are public: Bridging public and private organizational theories*. San Francisco: Jossey-Bass.
- British Library of Political and Economic Science & Charles Booth Collection. (1988). Life and labour of the people of London: The Charles Booth Collection, 1885–1905, British Library of Political and Economic Science, London: A listing and guide. Brighton, Sussex: Harvester Press Microform Publications.
- Brusca, I. (1997). The usefulness of financial reporting in Spanish local governments. *Financial Accountability and Management*, 13(1), 17–34. <https://doi.org/10.1111/1468-0408.00024>
- Brusca, I., & Montesinos, V. (2016). Implementing performance reporting in local government: A cross-countries comparison. *Public Performance and Management Review*, 39(3), 506–534. <https://doi.org/10.1080/15309576.2015.1137768>
- Brusca. (1997). The usefulness of financial reporting in Spanish local governments. *Financial Accountability & Management*, 13(1), 17–34.
- Bulmer, M., Bales, K., & Sklar, K. K. (Eds). (1991). *The social survey in historical perspective, 1880—1940*. Cambridge: Cambridge University Press.
- Burchell, S., Clubb, C., Hopwood, A., Hughes, J., & Nahapiet, J. (1980). The roles of accounting in organizations and society. *Accounting, organizations and society*, 5(1), 5–27.
- Bursztyn, L., & Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9, 131–153.
- Buylen, B., & Christiaens, J. (2016). Talking numbers? Analyzing the presence of financial information in councilors' speech during the budget debate in Flemish municipal councils. *International Public Management Journal*, 19(4), 453–475. <https://doi.org/10.1080/10967494.2015.1064502>
- Bäckstrand, K. (2008). Accountability of networked climate governance: The rise of transnational climate partnerships. *Global environmental politics*, 8(3), 74–102.
- Cairney, P. (2016). *The politics of evidence-based policy making*. London: Springer.
- Calderini, M., Chiodo, V., & Michelucci, F. V. (2018). The social impact investment race: Toward an interpretative framework. *European Business Review*, 30(1), 66–81. <https://doi.org/10.1108/EBR-10-2016-0134>
- Calvert, R. (2013). *Models of imperfect information in politics*. Oxfordshire: Taylor & Francis.
- Carlin, T., & Guthrie, J. (2003). Accrual output-based budgeting systems in Australia | The rhetoric-reality gap. *Public Management Review*, 5(2), 145–162.
- Carnegie, G. D., & West, B. P. (2005). Making accounting accountable in the public sector. *Critical Perspectives on Accounting*, 16(7), 905–928. <https://doi.org/10.1016/j.cpa.2004.01.002>

- Carpenter, D. P., Esterling, K. M., & Lazer, D. M. (2004). Friends, brokers, and transitivity: Who informs whom in Washington politics? *The Journal of Politics*, *66*(1), 224–246.
- Carvalho, J., Jorge, S., & José Fernandes, M. (2006). New local government accounting in Portugal. *Public Money and Management*, *26*(4), 211–216. <https://doi.org/10.1111/j.1467-9302.2006.00527.x>
- Cavalluzzo, K. S., & Irtner, C. D. (2004). Implementing performance measurement innovations: Evidence from government. *Accounting, Organizations and Society*, *29*(3–4), 243–267. [https://doi.org/10.1016/S0361-3682\(03\)00013-8](https://doi.org/10.1016/S0361-3682(03)00013-8)
- Cepiku, D., Hinna, A., Scarozza, D., & Savignon, A. B. (2017). Performance information use in public administration: An exploratory study of determinants and effects. *Journal of Management and Governance*, *21*(4), 963–991. <https://doi.org/10.1007/s10997-016-9371-3>
- Charbonneau, E., & Bellavance, F. (2012). Blame avoidance in public reporting: Evidence from a provincially mandated municipal performance measurement regime. *Public Performance & Management Review*, *35*(3), 399–421.
- Charbonneau, É., & Bellavance, F. (2015). Performance management in a benchmarking regime: Quebec's municipal management indicators. *Canadian Public Administration*, *58*(1), 110–137. <https://doi.org/10.1111/capa.12098>
- Charbonneau, É., & Nayer, G. (2012). Barriers to the use of benchmarking information: Narratives from local government managers. *Journal of Public Management & Social Policy*, *18*(2), 25–47.
- Charbonneau, É., & van Ryzin, G. G. (2015). Benchmarks and citizen judgments of local government performance: Findings from a survey experiment. *Public Management Review*, *17*(2), 288–304. <https://doi.org/10.1080/14719037.2013.798027>
- Chatman, J. A. (1989). Matching people and organizations: Selection and socialization in public accounting firms. *Academy of Management Proceedings*, *1989*(1), 199–203. <https://doi.org/10.5465/ambpp.1989.4980837>
- Chen, S.-H. (2017). Who cares: Stakeholder participation and the use of performance information in strategic planning. *Chinese Public Administration Review*, *8*(2), 86. <https://doi.org/10.22140/cpar.v8i2.125>
- Chewing Jr., E. G., & Harrell, A. M. (1990). The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task. *Accounting, Organizations and Society*, *15*(6), 527–542.
- Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict management and peace science*, *22*(4), 341–352.
- Cohen, S. (2008). Identifying the moderator factors of financial performance in Greek municipalities. *Financial Accountability & Management*, *24*(3), 265–294. <https://doi.org/10.1111/j.1468-0408.2008.00453.x>
- Colarelli, S. M., Hechanova-Alampay, R., & Canali, K. G. (2002). Letters of recommendation: An evolutionary psychological perspective. *Human Relations*, *55*, 315–344.
- Collins, K. M. T. (2000). Implementing mathematics curricula standards: Effective instruction for “all” students? Paper presented at the annual meeting of the Midwestern Educational Research Association, Chicago, IL (2000, October).
- Coulter, K. S., & Coulter, R. A. (2005). Size does matter: The effects of magnitude representation congruency on price perceptions and purchase likelihood. *Journal of Consumer Psychology*, *15*(1), 64–76.

- Courty, P., Heinrich, C., & Marschke, G. (2005). Setting the standard in performance measurement systems. *International Public Management Journal*, 8(3), 321–347. <https://doi.org/10.1080/10967490500439628>
- Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Los Angeles: SAGE Publications.
- Cuéllar, I., & González, G. (2000). Cultural identity description and cultural formulation for Hispanics. In R. H. Dana (Ed.), *Personality and clinical psychology series. Handbook of cross-cultural and multicultural personality assessment* (p. 605–621). Lawrence Erlbaum Associates Publishers.
- Cunningham, G. M., & Harris, J. E. (2001). A heuristic framework for accountability of governmental subunits. *Public Management Review*, 3(2), 145–165.
- Cunningham, G. M., & Harris, J. E. (2005). Toward a theory of performance reporting to achieve public sector accountability: A field study. *Public Budgeting and Finance*, 25(2), 15–42. <https://doi.org/10.1111/j.0275-1100.2005.00359.x>
- Currstine, T. (2006). Government performance. *OECD Journal on Budgeting*, 5(1), 127–151. <https://doi.org/10.1787/budget-v5-art6-en>
- da Silva Nogueira, S. P., & Jorge, S. M. F. (2017). The perceived usefulness of financial information for decision making in Portuguese municipalities. *Journal of Applied Accounting Research*, 18(1), 116–136
- Dal Molin, M., Turri, M., & Agasisti, T. (2017). New public management reforms in the Italian universities: Managerial tools, accountability mechanisms or simply compliance? *International Journal of Public Administration*, 40(3), 256–269. <https://doi.org/10.1080/01900692.2015.1107737>
- Dandeneau, S. D., & Baldwin, M. W. (2004). The inhibition of socially rejecting information among people with high versus low self-esteem: The role of attentional bias and the effects of bias reduction training. *Journal of Social and Clinical Psychology*, 23(4), 584–603.
- Daniels, J. D., & Daniels, C. E. (1991). Municipal financial reports: What users want. *Journal of Accounting and Public Policy*, 10(1), 15–38.
- Davidson, J., & Sternberg, R. (Eds). (2003). *The psychology of problem solving*. Cambridge, England: Cambridge University Press.
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2–21.
- de Boer, N., Eshuis, J., & Klijn, E. (2018). Does disclosure of performance information influence street-level bureaucrats' enforcement style? *Public Administration Review*, 78(5), 694–704.
- de Castro e Silva Neto, D., Cruz, C. O., Rodrigues, F., & Silva, P. (2016). Bibliometric analysis of PPP and PFI literature: Overview of 25 years of research. *Journal of Construction Engineering and Management*, 142(10), 06016002.
- de Figueiredo Jr, R. J. (2002). Electoral competition, political uncertainty, and policy insulation. *American Political Science Review*, 321–333.
- de Lancer Julnes, P., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Performance Measurement: Building Theory, Improving Practice: Building Theory, Improving Practice*, 166–189. <https://doi.org/10.1111/0033-3352.00140>
- de Lancer Julnes, P. (2008). Performance measurement beyond instrumental use. In W. van Dooren & S. van de Walle (Eds), *Performance Information in the Public Sector: How It Is Used* (pp. 58–71). Houndmills, UK: Palgrave Macmillan.

- Del Giudice, A., & Migliavacca, M. (2019). Social impact bonds and institutional investors: An empirical analysis of a complicated relationship. *Nonprofit and Voluntary Sector Quarterly*, 48(1), 50–70.
- Demaj, L. (2017). What can performance information do to legislators? A budget-decision experiment with legislators. *Public Administration Review*, 77(3), 366–379. <https://doi.org/10.1111/puar.12713>
- Dewey, J. (1948, 1920). *Reconstruction in philosophy*. Boston, MA: Beacon Press
- Dewey, J., 1999. The Essential Dewey (two volumes edited by Hickman, L. and Alexander, T.), Bloomington: Indiana University Press, 1999.
- Diclemente, C. C., Marinilli, A. S., Singh, M., & Bellino, L. E., (2001). The role of feedback in the process of health behavior change. *American Journal of Health Behavior*, 25, 217–227.
- di Francesco, M. (1999). Measuring performance in policy advice output: Australian developments. *International Journal of Public Sector Management*, 12(5), 420–431. <https://doi.org/10.1108/09513559910300181>
- di Francesco, M. (2012). Grand designs? The “managerial” role of ministers within Westminster-Based public management policy. *Australian Journal of Public Administration*, 71(3), 257–268. <https://doi.org/10.1111/j.1467-8500.2012.00783.x>
- Dolan, P., Hallsworth, M., Halpern, D., King, D., & Vlaev, I. (2010). *MINDSPACE: Influencing behaviour through public policy*. London, UK: Cabinet Office.
- Dowrick, S. and Quiggin, J., (1998). Measures of economic activity and welfare: the uses and abuses of GDP, in R. Eckersley (ed.), *Measuring Progress: is life getting better?* Collingwood, Australia: CSIRO Publishing, 93–121.
- Drechsler, W. and Kattel, R. (2008) ‘Towards the Neo-Weberian State? Perhaps, but Certainly, Adieu, NPM!’. *NISPAce Journal of Public Administration and Policy, Special issue: ‘A Distinctive European Model? The Neo-Weberian State*, 1:2, pp. 95–9. <https://www.nispa.org/journal.php?sid=569>
- Dretske, Fred. (1985). Precis of knowledge and the flow of information. In Hilary Kornblith (Ed.), *Naturalizing epistemology* (pp. 169–187). Cambridge, MA: MIT Press.
- Duhaime, I. M., & Schwenk, C. R. (1985). Conjectures on cognitive simplification in acquisition and divestment decision making. *Academy of Management Review*, 10(2), 287–295.
- Duisenberg, P. (2016). New development: How MPs in the Dutch parliament strengthened their budgetary and accounting powers. *Public Money and Management*, 36(7), 521–526. <https://doi.org/10.1080/09540962.2016.1237153>
- Dunwoody, S., & Griffin, R. J. (2014). The role of channel beliefs in risk information seeking. In J. Arvai & L. Rivers III (Eds), *Effective risk communication* (pp. 220–233). London: Taylor & Francis.
- Dull, M. (2009). Results-model reform leadership: Questions of credible commitment. *Journal of public administration research and theory*, 19(2), 255–284.
- Dunleavy, P., H. Margetts, S. Bastow and J. Tinkler (2006). *Digital Era Governance*. Oxford: Oxford University Press.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Döring, H., Downe, J., & Martin, S. (2015). Regulating public services: How public managers respond to external performance assessment. *Public Administration Review*, 75(6), 867–877. <https://doi.org/10.1111/puar.12400>
- Edwards, G. (2003). *On deaf ears: The limits of the bully pulpit*. Yale University Press: New Haven, CT.

- Elliott, B., & Elliott, J. (2007). *Financial accounting and reporting*. Essex: Pearson Education.
- Ellul, L., & Hodges, R. (2019). Reforming the government budgeting system in Malta. *Journal of Public Budgeting, Accounting & Financial Management*, 31(4), pp. 518–538.
- Empson, W. (2004). *Seven Types of Ambiguity*. London: Pimlico.
- Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(3), 133–40.
- Ezzamel, M., Hyndman, N. S., Johnsen, Å., Lapsley, I., & Pallot, J. (2004). Has devolution increased democratic accountability? *Public Money and Management*, 24(3), 145–152. <https://doi.org/10.1111/j.1467-9302.2004.00411.x>
- Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14, 159–181.
- Ferlie, E., Pollitt, C., & Lynn, L. E. (2005). *The Oxford handbook of public management*. Oxford: Oxford University Press.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston: Row, Peterson.
- Flamholtz, E. (1996). *Effective management control: Theory and practice*. Boston: Springer.
- Folz, D. H., Abdelrazek, R., & Chung, Y. (2009). The adoption, use, and impacts of performance measures in medium-size cities. *Public Performance & Management Review*, 33(1), 63–87. <https://doi.org/10.2753/pmr1530-9576330103>
- Franklin, A. L., & Carberry-George, B. (1999). Analyzing how local governments establish service priorities. *Public Budgeting & Finance*, 19(3), 31–46.
- Frey, B. S., & Osterloh, M. (Eds.). (2001). *Successful management by motivation: Balancing intrinsic and extrinsic incentives*. Springer Science & Business Media.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40, 351–401.
- Furnham, A. (1986). Economic locus of control. *Human relations*, 39(1), 29–43.
- Furnham, A. (1986). Response bias, social desirability, and dissimulation. *Personality and Individual Differences*, 7(3), 385–400. [https://doi.org/10.1016/0191-8869\(86\)90014-0](https://doi.org/10.1016/0191-8869(86)90014-0)
- Gadamer, H. G. (2004). *Truth and Method*. London: Continuum.
- George, B., Baekgaard, M., Decramer, A., Audenaert, M., & Goeminne, S. (2018). Institutional isomorphism, negativity bias and performance information use by politicians: A survey experiment. *Public Administration*, 1–15. <https://doi.org/10.1111/padm.12390>
- Giacomini, D., Sicilia, M., & Steccolini, I. (2016). Contextualizing politicians' uses of accounting information: Reassurance and ammunition. *Public Money and Management*, 36(7), 483–490. <https://doi.org/10.1080/09540962.2016.1237128>
- Goh, S. C. (2002). Managing effective knowledge transfer: An integrative framework and some practice implications. *Journal of Knowledge Management*, 6(1), 23–30.
- Greenwald, B. C., & Stiglitz, J. E. (1986). Externalities in economies with imperfect information and incomplete markets. *The Quarterly Journal of Economics*, 101, 229–264.
- Grizzle, G. A. (1981). A manager's guide to the meaning and uses of performance measurement. *The American Review of Public Administration*, 15(1), 16–28. <https://doi.org/10.1177/027507408101500104>
- Grossi, G., Reichard, C., & Ruggiero, P. (2016). Appropriateness and use of performance information in the budgeting process: Some experiences from German and Italian municipalities. *Public Performance and Management Review*, 39(3), 581–606. <https://doi.org/10.1080/15309576.2015.1137770>
- Gruening, G. (2001). Origin and theoretical basis of New Public Management. *International Public Management Journal*, 4(1), 1–25.

- Guba, E. G., & Lincoln, Y. S. (1998). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds), *The Sage handbook of qualitative research* (3rd edn., pp. 191–215). Thousand Oaks: Sage Publications.
- Guthrie, J., & English, L. (1997). Performance information and programme evaluation in the Australian public sector. *International Journal of Public Sector Management*, 10(3), 154–164. <https://doi.org/10.1108/09513559710166039>
- Haas, N., & Speckbacher, G. (2017). Everything under my control: CEO characteristics and the evaluation of middle manager performance in small and medium-sized firms. *Schmalenbach Business Review*, 18(2), 109–128. <https://doi.org/10.1007/s41464-017-0027-x>
- Hahn, M., Lawson, R., & Lee, Y. G. (1992). The effects of time pressure and information load on decision quality. *Psychology & Marketing*, 9, 365–378.
- Hammerschmid, G., van de Walle, S., & Stimac, V. (2013). Internal and external use of performance information in public organizations: Results from an international survey. *Public Money and Management*, 33(4), 261–268. <https://doi.org/10.1080/09540962.2013.799803>
- Hampton, W. (2018). "Things for Those Who Wait: Predictive Modeling Highlights Importance of Delay Discounting for Income Attainment". *Frontiers in Psychology*, 9 (1545). <https://doi.org/10.3389/fpsyg.2018.01545>
- Harries, U., Elliott, H., & Higgins, A. (1999). Evidence-based policy-making in the NHS: exploring the interface between research and the commissioning process. *Journal of Public Health*, 21(1), 29–36.
- Hasan, H., & Hasan, S. (1997). Computer-based performance information for executives in local government. *Australian Journal of Public Administration*, 56(3), 24–29. <https://doi.org/10.1111/j.1467-8500.1997.tb01261.x>
- Hatry, H. P. (2006). *Performance measurement: Getting results*. Washington, DC: The Urban Institute.
- Hatry, H. P. (2008). The many faces of use. In W. Van Dooren & S. Van de Walle (Eds), *Performance information in the public sector: How it is used* (pp. 125–140). Basingstoke, UK: Palgrave Macmillan.
- Hatry, H. (2011). Epilogue: The many faces of use. In W. Van Dooren & S. Van de Walle (Eds), *Performance information in the public sector: How it is used* (pp. 227–240). Basingstoke, UK: Palgrave Macmillan.
- Hanushek, E. A., and J. E. Jackson. 1977. *Statistical methods for social scientists*. New York: Academic Press.
- Henderson, A. C., & Bromberg, D. E. (2016). Performance information use in local Government: Monitoring relationships with emergency medical services agencies. *Public Performance and Management Review*, 39(1), 58–82. <https://doi.org/10.1080/15309576.2016.1071162>
- Hibbard, J. H., Slovic, P., & Jewett, J. J. (1997). Informing consumer decisions in health care: Implications from decision-making research. *Milbank Quarterly*, 75(3), 395–414. <https://doi.org/10.1111/1468-0009.00061>
- Hildebrand, R., & McDavid, J. C. (2011). Joining public accountability and performance management: A case study of Lethbridge, Alberta. *Canadian Public Administration*, 54(1), 41–72. <https://doi.org/10.1111/j.1754-7121.2011.00160.x>
- Hill, H., & Andrews, M. (2005). Reforming budget ritual and budget practice: The case of performance management implementation in Virginia. *International Journal of Public Administration*, 28(3–4), 255–272. <https://doi.org/10.1081/PAD-200047319>
- Hinterleitner, M., & Sager, F. (2015). Avoiding blame: A comprehensive framework and the Australian home insulation program fiasco. *Policy Studies Journal*, 43(1), 139–161.

- Hirschey, M. (2008). *Fundamentals of managerial economics*. Mason: Cengage Learning.
- Ho, A. T. K., & Im, T. (2015). Challenges in building effective and competitive government in developing countries: An institutional logics perspective. *The American Review of Public Administration*, 45(3), 263–280.
- Ho, A. T. K., & Ni, A. Y. (2005). Have cities shifted to outcome-oriented performance reporting?—a content analysis of city budgets. *Public Budgeting and Finance*, 25(2), 61–83. <https://doi.org/10.1111/j.0275-1100.2005.00361.x>
- Ho, A. T.-K. (2006). Accounting for the value of performance measurement from the perspective of Midwestern mayors. *Journal of Public Administration Research and Theory*, 16(2), 217–237.
- Hodges, R. (2012). Joined-up government and the challenges to accounting and accountability researchers. *Financial Accountability & Management*, 28(1), 26–51.
- Holloway, D., Alam, M., Griffiths, A., & Holloway, D. (2012). Performance management in Australia's public mental health service: A state-based perspective. *Australian Journal of Public Administration*, 71(1), 20–32. <https://doi.org/10.1111/j.1467-8500.2012.00749.x>
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3–19.
- Hood, C. (2011). *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton, NJ: Princeton University Press
- Hood, C. (2014). Accountability and blame-avoidance. In M. Bovens, R. E. Goodin, & T. Schillemans (Eds), *The Oxford handbook of public accountability* (pp. 603–616). Oxford: Oxford University Press.
- Hoontis, P., & Kim, T. (2012). Antecedents to municipal performance measurement implementation. *Public Performance and Management Review*, 36(1), 158–173. <https://doi.org/10.2753/PMR1530-9576360108>
- Hou, Y., Lunsford, R. S., Sides, K. C., & Jones, K. A. (2011). State performance-based budgeting in boom and bust years: An analytical framework and survey of the states. *Public Administration Review*, 71(3), 370–388. <https://doi.org/10.1111/j.1540-6210.2011.02357.x>
- Hughes, O. E. (1998). New public management. In *Public Management and Administration* (pp. 52–80). UK: Macmillan Education.
- Hyndman, N. (1990). Charity accounting—an empirical study of the information needs of contributors to UK fundraising charities. *Financial Accountability & Management*, 6, 295–307. <https://doi.org/10.1111/j.1468-0408.1990.tb00335.x>
- Innes, J. E. (1990). *Knowledge and public policy: The search for meaningful indicators*. New Brunswick: Transaction Publishers.
- Ionciă, M., Petrescu, E. C., Ionciă, D., & Constantinescu, M. (2012). The role of education on consumer behavior on the insurance market. *Procedia-Social and Behavioral Sciences*, 46, 4154–4158.
- Jääskeläinen, A., & Roitto, J. M. (2014). Drivers of personnel satisfaction toward performance information usage: Findings from Finnish municipalities. *International Journal of Public Sector Management*, 27(6), 530–547. <https://doi.org/10.1108/IJPSM-03-2014-0044>
- James, O. (2011). Managing citizens' expectations of public service performance: evidence from observation and experimentation in local government. *Public Administration*, 89(4), 1419–1435.
- James, W. (1995, 1907). *Pragmatism*. New York: Dover.
- Jansen, E. P. (2008). New public management: Perspectives on performance and the use of performance information. *Financial Accountability & Management*, 24(2), 169–191.

- Jantz, B., Christensen, T. & Læg Reid, P. (2015) Performance Management and Accountability: The Welfare Administration Reform in Norway and Germany. *International Journal of Public Administration*, 38(13), 947–959. <https://doi.org/10.1080/01900692.2015.1069838>
- Johansen, M., & Zhu, L. (2014). Market competition, political constraint, and managerial practice in public, nonprofit, and private American hospitals. *Journal of Public Administration Research and Theory*, 24(1), 159–184. <https://doi.org/10.1093/jopart/muto29>
- Johansson, T., & Siverbo, S. (2009). Explaining the utilization of relative performance evaluation in local government: A multi-theoretical study using data from Sweden. *Financial Accountability & Management*, 25(2), 197–224. <https://doi.org/10.1111/j.1468-0408.2009.00474.x>
- Johnsen, Å (2012). Why does poor performance get so much attention in public policy? *Financial Accountability and Management*, 28(2), 121–142.
- Johnsen, Å. (2005). What does 25 years of experience tell us about the state of performance measurement in public policy and management? *Public Money and Management*, 25(1), 9–17. <https://doi.org/10.1111/j.1467-9302.2005.00445.x>
- Johnsen, Å. (2011). Performance information and educational policy making. In W. Van Dooren & S. Van de Walle (Eds), *Performance information in the public sector: How it is used* (pp. 161–77). Basingstoke, UK: Palgrave Macmillan.
- Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, 21(1), 93–110. [https://doi.org/10.1016/S0149-7189\(97\)00048-7](https://doi.org/10.1016/S0149-7189(97)00048-7)
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26. <https://doi.org/10.3102/0013189X033007014>
- Johnson-Laird, P. (1998). Imagery, visualization, and thinking. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 441–467). New York: Academic Press.
- Johnston, M. P. (2017). Secondary data analysis: A method of which the time has come. *Qualitative and quantitative methods in libraries*, 3(3), 619–626.
- Jones, C. S., & Dewing, I. P. (1997). The attitudes of NHS clinicians and medical managers towards changes in accounting controls. *Financial Accountability and Management*, 13(3), 261–280. <https://doi.org/10.1111/1468-0408.00037>
- Jones, L. R., & McCaffery, J. L. (2010). Performance budgeting in the US federal government: History, status and future implications. *Public Finance and Management*, 10(3), 482.
- Jorge, S., Jesus, M. A. J. de, & Nogueira, S. P. (2019). The use of budgetary and financial information by politicians in parliament: A case study. *Journal of Public Budgeting, Accounting and Financial Management*, 31(4), 539–557. <https://doi.org/10.1108/JPB AFM-11-2018-0135>
- Jorge, S., Jorge de Jesus, M. A., & Nogueira, S. (2016). Information brokers and the use of budgetary and financial information by politicians: The case of Portugal. *Public Money and Management*, 36(7), 515–520. <https://doi.org/10.1080/09540962.2016.1237152>
- Joyce, P. G. (1996). Jesse Burkhead and the multiple uses of federal budgets: A contemporary perspective. *Public Budgeting and Finance*, 16(2), 59–78. <https://doi.org/10.1111/1540-5850.01068>
- Joyce, P. G. (2011). The Obama Administration and PBB: Building on the legacy of federal performance-informed budgeting? *Public Administration Review*, 71(3), 356–367.
- Joyce, P. G., & Levy, A. (2008). The program assessment rating tool and its role in evaluating the effectiveness of U. S. government programs. *Journal of the Washington Institute of China Studies, [S.L.]*, 3(1), 25–36.

- Jung Wook Lee & Yoon Jik Cho (2011). In Search of a Better Fit: Exploring Conditions for Successful Performance-Based Human Resources Management in U.S. Federal Agencies. *International Review of Public Administration*, 16(3), 89–112. <https://doi.org/10.1080/12294659.2011.10805209>
- Kahneman, D & Tversky, A (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 2, 263–291.
- Kahneman, D., & Tversky, A. (1999). Evaluation by moments: Past and future. In D. Kahneman & A. Tversky (Eds), *Choices, values and frames* (pp. 2–23). New York: Cambridge University Press.
- Kahneman, D (2011). *Thinking fast and slow*. London: Allen Lane.
- Kahneman, D. & Tversky, A. (2013). Choices, values, and frames. *Handbook of the Fundamentals of Financial Decision Making*, July 2013, 269–278.
- Kang, H. Y., Kim, S. J., Cho, W., & Lee, S. (2009). Consumer use of publicly released hospital performance information: Assessment of the National Hospital Evaluation Program in Korea. *Health Policy*, 89(2), 174–183. <https://doi.org/10.1016/j.healthpol.2008.05.009>
- Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty*, 38, 95–115.
- Kasdin, S. (2010). Reinventing reforms: How to improve program management using performance measures. Really. *Public Budgeting and Finance*, 30(3), 51–78. <https://doi.org/10.1111/j.1540-5850.2010.00962.x>
- Kaufman, R. (1992). The challenge of Total Quality Management in education. *International Journal of Educational Reform*, 1(2), 149–165.
- Kennedy, M. M. (2007). Defining a literature. *Educational Researcher*, 36(3), 139–147.
- Kidholm, K., & Kristensen, M. B. D. (2018). A scoping review of economic evaluations alongside randomised controlled trials of home monitoring in chronic disease management. *Applied Health Economics and Health Policy*, 16(2), 167–176. <https://doi.org/10.1007/s40258-017-0351-9>
- Kim, Y., & Kang, M. (2016). The Performance Management System of the Korean healthcare sector: Development, challenges, and future tasks. *Public Performance and Management Review*, 39(2), 297–315. <https://doi.org/10.1080/15309576.2015.1108771>
- Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V. Caracelli & H. Preskill (Eds), *The expanding scope of evaluation use*. New Directions for Evaluation, No. 88 (pp. 5–24). San Francisco, CA: Jossey-Bass.
- Kloot, L. (1999). Performance measurement and accountability in Victorian local government. *International Journal of Public Sector Management*, 12(7), 565–584. <https://doi.org/10.1108/09513559910308039>
- Kober, R., Lee, J., & Ng, J. (2010). Mind your accruals: Perceived usefulness of financial information in the Australian public sector under different accounting systems. *Financial Accountability & Management*, 26(3), 267–298. <https://doi.org/10.1111/j.1468-0408.2010.00502.x>
- Kock, N. (2010). *Evolutionary psychology and information systems research*. US: Springer.
- Kool, D., & Bekkers, V. (2016). The perceived value-relevance of open data in the parents' choice of Dutch primary schools. *International Journal of Public Sector Management*, 29(3), pp. 271–287. <https://doi.org/10.1108/IJPSM-02-2016-0022>
- Korir, S., Rotich, J., & Bengat, J. (2015). Performance management and public service delivery in Kenya. *European Journal of Research and Reflection in Management Science*, 3(4).
- Kristensen, J. K., Groszyk, W., & Bühler, B. (2002). Outcome-focused management and budgeting. *OECD Journal on Budgeting*, 1(4), 7–34.

- Kroll, A. (2013). The other type of performance information: Nonroutine feedback, its relevance and use. *Public Administration Review*, 73(2), 265–276. <https://doi.org/10.1111/j.1540-6210.2012.02648.x>
- Kroll, A. (2014). Why performance information use varies among public managers: Testing manager-related explanations. *International Public Management Journal*, 17(2), 174–201.
- Kroll, A. (2015a). Drivers of performance information use: Systematic literature review and directions for future research. *Public Performance and Management Review*, 38(3), 459–486. <https://doi.org/10.1080/15309576.2015.1006469>
- Kroll, A. (2015b). Explaining the use of performance information by public managers: A planned-behavior approach. *The American Review of Public Administration*, 45(2), 201–215.
- Kroll, A. (2016). Exploring the link between performance information use and organizational performance: A contingency approach. *Public Performance and Management Review*, 39(1), 7–32. <https://doi.org/10.1080/15309576.2016.1071159>
- Kroll, A., & Proeller, I. (2013). Controlling the control system: Performance information in the German childcare administration. *International Journal of Public Sector Management*, 26(1), 74–85. <https://doi.org/10.1108/09513551311293435>
- Kroll, A., & Vogel, D. (2014). The PSM–leadership fit: A model of performance information use. *Public Administration*, 92(4), 974–991. <https://doi.org/10.1111/padm.12014>
- Kroll, A., & Moynihan, D. (2015). Creating public value using performance information. In J. M. Bryson, B. C. Crosby, & L. Bloomberg (Eds), *Valuing public value* (pp. 243–275). Washington, DC: Georgetown University Press.
- Larrick, Richard P. 2004. Debiasing. In Derek J. Koehler and Nigel Harvey (Eds), *Blackwell Handbook of Judgment and Decision Making* (pp. 316–38). Malden, MA: Blackwell.
- Laihonen, H., & Mäntylä, S. (2018). Strategic knowledge management and evolving local government. *Journal of Knowledge Management*, 22(1), pp. 219–234
- Lam, W. F., & Wang, X. (2014). The cognitive foundation of a co-production approach to performance measurement: How do officials and citizens understand government performance in China? *Public Administration and Development*, 34(1), 32–47.
- Lapsley, I. (1999). Accounting and the new public management: Instruments of substantive efficiency or a rationalising modernity? *Financial Accountability & Management*, 15(3–4), 201–207.
- Lee, J. (2008). Preparing performance information in the public sector: An Australian perspective. *Financial Accountability & Management*, 24(2), 117–149. <https://doi.org/10.1111/j.1468-0408.2008.00449.x>
- Lee, K., & Lee, S. (2015). A study of performance information use and credibility in performance management systems: Focusing on the self-evaluation system. *Korean Journal of Policy Studies*, 30(2), 23–45.
- Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality and Quantity*, 43(2), 265–275. <https://doi.org/10.1007/s11135-007-9105-3>
- Lepistö, L. (2014). Label in context: On the enterprise resource planning system in a medium-sized enterprise. *The Finnish Journal of Business Economics*, 63(1), 48–72.
- Levati, S., Campbell, P., Frost, R., et al. (2016). Optimisation of complex health interventions before a randomised controlled trial: A scoping review of strategies used. *Pilot Feasibility Stud.* 2(17). <https://doi.org/10.1186/s40814-016-0058-y>
- Levine, J., Chan, K. M. A., & Satterfield, T. (2015). From rational actor to efficient complexity manager: Exorcising the ghost of Homo economicus with a unified synthesis of cognition research. *Ecological Economics*, 114, 22–32. <https://doi.org/10.1016/j.ecolecon.2015.03.010>

- Levine, L. J., & Bluck, S. (1997). Experienced and remembered emotional intensity in older adults. *Psychology and Aging, 12*, 514–523.
- Lewis, J. K. (2016). Corporate social responsibility/sustainability reporting among the fortune global 250: Greenwashing or green supply chain? *Entrepreneurship, Business and Economics, 1* (pp. 347–362). Springer, Cham.
- Liguori, M., Sicilia, M., & Steccolini, I. (2012). Some like it non-financial... Politicians' and managers' views on the importance of performance information. *Public Management Review, 14*(7), 903–922.
- Lindblom, C. E., & Cohen, D. K. (1979). *Usable knowledge: Social science and social problem solving*. New Haven: Yale University Press.
- Liner, B., Hatry, H., Vinson, E., Allen, R., Dusenbury, P., Bryant, S., et al. (2001). *Making results-based state government work*. Washington, DC: Urban Institute.
- Liu, J., Love, P., Smith, J., Regan, M., & Sutrisna, M. (2014). Public-private partnerships: A review of theory and practice of performance measurement. *International Journal of Productivity and Performance Management, 63*(4), 499–512.
- Lönqvist, A., Sillanpää, V., & Carlucci, D. (2009). Intellectual capital management in practice: Assessment of implementation and outcomes. *Knowledge Management Research & Practice, 7*(4), 308–316.
- Lu, Y. (2007). Performance budgeting: The perspective of state agencies. *Public Budgeting and Finance, 27*(4), 1–17. <https://doi.org/10.1111/j.1540-5850.2007.00890.x>
- Lu, Y., Willoughby, K., & Arnett, S. (2009). Legislating results: Examining the legal foundations of PBB systems in the states. *Public Performance & Management Review, 33*(2), 266–287.
- Lu, Y. (2011). Individual engagement to collective participation: The dynamics of participation pattern in performance budgeting. *Public Budgeting and Finance, 31*(2), 79–98. <https://doi.org/10.1111/j.1540-5850.2011.00980.x>
- Lu, E. Y. (2013). Beginning to unlock the black box of the budgetary performance evaluation practices in China: A case study of evaluation reports from Zhejiang province. *Public Money and Management, 33*(4), 253–260. <https://doi.org/10.1080/09540962.2013.799802>
- Lu, E. Y., & Willoughby, K. (2015). Performance budgeting in American states: A framework of integrating performance with budgeting. *International Journal of Public Administration, 38*(8), 562–572. <https://doi.org/10.1080/01900692.2014.949751>
- Lu, E. Y., Mohr, Z., & Ho, A. T. K. (2015). Taking stock: Assessing and improving performance budgeting theory and practice. *Public Performance and Management Review, 38*(3), 426–458. <https://doi.org/10.1080/15309576.2015.1006470>
- Luft, J., Shields, M. D., & Thomas, T. F. (2016). Additional information in accounting reports: Effects on management decisions and subjective performance evaluations under causal ambiguity. *Contemporary Accounting Research, 33*(2), 526–550. <https://doi.org/10.1111/1911-3846.12156>
- Lynn Jr., L. E. (1998). The new public management: How to transform a theme into a legacy. *Public Administration Review, 58*(3), 231–237.
- Mack, J. (2004). *An investigation of the information requirements of users of Australian public sector financial reports*. PhD Thesis, QUT.
- MacKuen, M. (1983). Political drama, economic conditions, and the dynamics of presidential popularity. *American Journal of Political Science, 27*(2), 165–192.
- MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology, 8*(1), 189–217.

- Malhotra, N. K. (1984). Information and sensory overload. *Information and sensory overload in psychology and marketing. Psychology & Marketing, 1*(3-4), 9–21.
- Malmi, T., & Brown, D. (2008). Management control systems as a package—opportunities, challenges, and research directions. *Management Accounting Research, 19*(4), 287–300.
- Mann, L., Harmoni, R., & Power, C. (1991). The GOFER course in decision making. In Baron, Jonathan, Brown, Rex V. (Eds), *Teaching decision making to adolescents*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 61–78. ISBN 978-0805804973.
- Masal, D., & Vogel, R. (2016). Leadership, use of performance information, and job satisfaction: Evidence from police services. *International Public Management Journal, 19*(2), 208–234. <https://doi.org/10.1080/10967494.2016.1143422>
- Mauro, G. (2016). *Performance-based budgeting: Understanding practice variation through the lens of institutional theories*. Pisa: S'ant Anna School of Advanced Studies.
- Mauro, S. G., Cinquini, L., & Grossi, G. (2017). Insights into performance-based budgeting in the public sector: A literature review and a research agenda. *Public Management Review, 19*(7), 911–931.
- Mayston, D. J. (1985). Non-profit performance indicators in the public sector. *Financial Accountability & Management, 1*(1), 51–74.
- Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy & Marketing, 25*, 1–21.
- McAdam, R., & Walker, T. (2004). Evaluating the best value framework in UK local government services. *Public Administration and Development, 24*(3), 183–196. <https://doi.org/10.1002/pad.275>
- McBeath, B., Jolles, M. P., Carnochan, S., & Austin, M. J. (2015). Organizational and individual determinants of evidence use by managers in public human service organizations. *Human Service Organizations Management, Leadership and Governance, 39*(4), 267–289. <https://doi.org/10.1080/23303131.2015.1044588>
- McGeough, F. (2014). Performance information in the Irish public sector: A comparison with the UK. *Public Money and Management, 34*(3), 197–204. <https://doi.org/10.1080/09540962.2014.908010>
- McGrath, R. G. (1999). Falling forward: Real options reasoning and entrepreneurial failure. *Academy of Management Review, 24*(1), 13–30.
- McKinsey Global Institute (2012). *The social economy: Unlocking value and productivity through social technologies*. New York: McKinsey and Company.
- McPhee, I. (2005). Outcomes and outputs: Are we managing better as a result? In *CPA National Public Sector Convention, 20*.
- Melkers, J., & Willoughby, K. (2005). Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review, 65*(2), 180–190.
- Melville, A. (2008). *International financial reporting: A practical guide*. Essex: Pearson Education.
- Miller, P., & Rose, N. (2008). *Governing the present: Administering economic, social and personal life*. Cambridge: Polity Press.
- Modell, S. (2004). Performance measurement myths in the public sector: A research note. *Financial Accountability & Management, 20*(1), 39–55.
- Montesinos, V., Brusca, I., Rossi, F. & Aversano, N. (2013). The usefulness of performance reporting in local government: Comparing Italy and Spain. *Public Money & Management, 33*(3), 171–176. <https://doi.org/10.1080/09540962.2013.785701>

- Morewedge, C. K., Gilbert, D. T., & Wilson, T. D. (2005). The least likely of times: How remembering the past biases forecasts of the future. *Psychological Science* 16(8), 626–630.
- Mortimer, J. T., & Lorence, J. (1979). Work experience and occupational value socialization: A longitudinal study. *American Journal of Sociology*, 84, 1361–1385.
- Mosher, F. (1968). *Democracy and the Public Service*, 2d ed. New York: Oxford University Press.
- Moynihan, D. P., & Ingraham, P. W. (2004). Integrative leadership in the public sector: A model of performance-information use. *Administration and Society*, 36(4), 427–453. <https://doi.org/10.1177/0095399704266748>
- Moynihan, D. P. (2005a). Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), 203–216. <https://doi.org/10.1111/j.1540-6210.2005.00445.x>
- Moynihan, D. P. (2005b). Why and how do state governments adopt and implement “managing for results” reforms? *Journal of Public Administration Research and Theory*, 15(2), 219–243. <https://doi.org/10.1093/jopart/mui012>
- Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Washington D.C.: Georgetown University Press.
- Moynihan, D., & Landuyt, N. (2009). How do public organizations learn? Bridging cultural and structural perspectives. *Public Administration Review*, 69(6), 1097–1105.
- Moynihan, D. P. (2009). Public performance & management review through a glass, darkly | Understanding the effects of performance regimes. *Public Performance & Management Review*, 32(4), 592–603. <https://doi.org/10.2753/PMR1530-9576320409>
- Moynihan, D. P., & Pandey, S. K. (2010). The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory*, 20(4), 849–866. <https://doi.org/10.1093/jopart/muq004>
- Moynihan, D. P., & Hawes, D. P. (2012). Responsiveness to reform values: The influence of the environment on performance information use. *Public Administration Review*, 72(SUPPL.1), 95–105. <https://doi.org/10.1111/j.1540-6210.2012.02653.x>
- Moynihan, D. P., & Lavertu, S. (2012). Does involvement in performance management routines encourage performance information use? Evaluating GPRA and PART. *Public Administration Review*, 72(4), 592–602.
- Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2012). Prosocial values and performance management theory: Linking perceived social impact and performance information use. *Governance*, 25(3), 463–483. <https://doi.org/10.1111/j.1468-0491.2012.01583.x>
- Murphy, J. P. (1990). *Pragmatism: From Peirce to Davidson*. Boulder, CO: Westview.
- Murray, R. (2002). Citizens’ control of evaluations: Formulating and assessing alternatives, *Evaluation*, 8(1), 81–100.
- Nagtegaal, R., Tummers, L., Noordegraaf, M., & Bekkers, V. (2020). Designing to Debias: Measuring and Reducing Public Managers’ Anchoring Bias. *Public Administration Review*, 80(4), 565–576.
- Nakrošis, V. 2008. “Reforming Performance Management in Lithuania: Towards Results-Based Government.” In G. B. Peters (Ed.), *Mixes, Matches and Mistakes: New Public Management in Russia and the Former Soviet Republics*, 53–116. Budapest: Open Society Institute.
- Newcomer, K. (2007). How does program performance assessment affect program management in the federal government? *Public Performance and Management Review*, 30(3), 332–350.
- Nicholson-Crotty, S., Nicholson-Crotty, J., & Fernandez, S. (2017). Performance and management in the public sector: Testing a model of relative risk aversion. *Public Administration Review*, 77(4), 603–614. <https://doi.org/10.1111/puar.12619>

- Nicholson-Crotty, S., Theobald, N., & Nicholson-Crotty, J. (2006). Disparate measures: Public managers and performance-measurement strategies. *Public Administration Review*, 66(1), 101–113. <https://doi.org/10.1111/j.1540-6210.2006.00559.x>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.
- Nielsen, P. A., & Baekgaard, M. (2015). Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment. *Journal of Public Administration Research and Theory*, 25(2), 545–569. <https://doi.org/10.1093/jopart/muto51>
- Nielsen, P. A., & Moynihan, D. P. (2017). How do politicians attribute bureaucratic responsibility for performance? Negativity bias and interest group advocacy. *Journal of Public Administration Research and Theory*, 27(2), 269–283. <https://doi.org/10.1093/jopart/muw060>
- Nõmm, K., & Randma-Liiv, T. (2012). Performance measurement and performance information in new democracies: A study of the Estonian central government. *Public Management Review*, 14(7), 859–879.
- Norman, R. (2002). Managing through measurement or meaning? Lessons from experience with New Zealand's public sector performance management systems. *International Review of Administrative Sciences*, 68(4), 619–628.
- Novak, J. D. (1966). *The role of concepts in science teaching*. In *Analyses of Concept Learning* (pp. 239–254). Amsterdam: Elsevier.
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32, 303–330.
- O'Reilly, C. A. (1982). Variations in decision makers' use of information sources: The impact of quality and accessibility of information. *Academy of Management Journal*, 25, 756.
- Olsen, A. L. (2016). The numerical psychology of performance information: Implications for citizens, managers, and policymakers. *Public Performance and Management Review*, 39(1), 100–115. <https://doi.org/10.1080/15309576.2016.1071167>
- Ongaro, E. (2009). *Public Management Reform and Modernization: Trajectories of Administrative Change in Italy, France, Greece, Portugal and Spain*. Cheltenham and Northampton: Edward Elgar.
- Onwuegbuzie, A. J., & Frels, R. K. (in press). *Seven steps to a comprehensive literature review: A multimodal and cultural approach*. London, England: Sage.
- Onwuegbuzie, A. J., DaRos-Voseles, D. A. (2001). The role of cooperative learning in research methodology courses: A mixed-methods analysis. *Research in Schools*, 8, 61–75.
- Osborne, D., & Gaebler, T. (1992). *Reinventing government: How the entrepreneurial spirit is transforming government*. Reading, MA: Addison Wesley Public Comp.
- Osbourne, S. (2010). *The New Public Governance: Emerging Perspectives on the Theory and Practice of Public Governance*. London and New York, Routledge/Taylor and Francis.
- Ospina, S., Grau, N. C., & Zaltsman, A. (2004). Performance evaluation, public management improvement and democratic accountability: Some lessons from Latin America. *Public Management Review*, 6(2), 229–251. <https://doi.org/10.1080/1471903042000189119>
- Ostrom, C., & Smith, R. (1993). Error correction, attitude persistence, and executive rewards and punishments: A behavioral theory of presidential approval. *Political Analysis*, 4, 127–184.
- Oulasvirta, L. (2014). The reluctance of a developed country to choose International Public Sector Accounting Standards of the IFAC. A critical case study. *Critical Perspectives on Accounting*, 25(3), 272–285.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Los Angeles: Sage Publications.

- Paulsson, G. (2006). Accrual accounting in the public sector: Experiences from the central government in Sweden. *Financial Accountability and Management*, 22(1), 47–62. <https://doi.org/10.1111/j.0267-4424.2006.00392.x>
- Peirce, C. S. (1998). *The essential Peirce selected philosophical writings*. Bloomington, IN: Indiana University Press.
- Pervin, L. (2003). *The science of personality* (2nd ed.). New York: Oxford University Press.
- Phinney, J. S. (1996). Understanding ethnic diversity: The role of ethnic identity. *American Behavioral Scientist*, 40(2), 143–152.
- Pierson, P. (1996). The new politics of the welfare state. *World Politics*, 48(2), 143–179.
- Plous, S. (1993). *McGraw-Hill series in social psychology. The Psychology of Judgment and Decision Making*. Pennsylvania: McGraw-Hill Book Company.
- Poister, T. H., & Streib, G. (2016). Performance measurement in municipal government: Assessing the state of the practice author(s). *Public Administration Review*, 59(4), 325–335.
- Pollitt, C. (1986). Performance measurement in the public services: Some political implications. *Parliamentary Affairs*, 39(3), 315–329.
- Pollitt, C., and Bouckaert, G. 2011. *Public management reform: A comparative analysis*. Oxford, UK: Oxford University Press.
- Pollitt, C. (2006). Performance information for democracy: The missing link? *Evaluation*, 12(1), 38–55. <https://doi.org/10.1177/1356389006064191>
- Poocharoen, O. O., & Wong, N. H. L. (2016). Performance management of collaborative projects: The stronger the collaboration, the less is measured. *Public Performance and Management Review*, 39(3), 607–629. <https://doi.org/10.1080/15309576.2015.1137767>
- Popper, K. R. (1959). *The logic of scientific discovery*. London, UK: Routledge.
- Posner, P. L., & Fantone, D. M. (2007). Assessing federal program performance: Observations on the US Office of Management and Budget's Program Assessment Rating Tool and its use in the budget process. *Public Performance & Management Review*, 30(3), 351–368.
- Previts, G and Merino, B. (1979). *A History of Accounting in America: An historical interpretation of the cultural significance of accounting*. New York: Wiley.
- Provan, K. G., & Milward, H. B. (2001). Do networks really work? A framework for evaluating public sector organizational networks. *Public Administration Review*, 61(4), 414–423.
- Quinlivan, D., Nowak, M., & Klass, D. (2014). From accountability to assurance—stakeholder perspectives in local government. *Australian Journal of Public Administration*, 73(2), 206–217. <https://doi.org/10.1111/1467-8500.12080>
- Rabin, M. & Schrag, J. (1999). First impressions matter: A model of confirmatory bias. *Quarterly Journal of Economics*, 114(1), 37–82.
- Rabovsky, T. (2014). Support for performance-based funding: The role of political ideology, performance, and dysfunctional information environments. *Public Administration Review*, 74(6), 761–774. <https://doi.org/10.1111/puar.12274>
- Rajala, T. (2019a). Leadership that drives social change with performance measures. In Kangas, A., Kujala J., Heikkinen, A., Lönnqvist, A., Laihonon H., and Bethwaite J. (Eds), *Leading Change in a Complex World: Transdisciplinary Perspectives* (pp. 201–222). Tampere: Tampere University Press
- Rajala, T. (2019b). Mind the information expectation gap. *Journal of the Knowledge Economy*, 10(1), 104–125. <https://doi.org/10.1007/s13132-016-0445-x>
- Rajala, T. and Laihonon, H. (2019). Managerial choices in orchestrating dialogic performance management. *Baltic Journal of Management*, Vol. 14 No. 1, pp. 141–157. <https://doi.org/10.1108/BJM-12-2017-0416>

- Rajala, T., Laihonon, H. and Haapala, P. (2018). Why is dialogue on performance challenging in the public sector? *Measuring Business Excellence*, 22(2), 117–129. <https://doi.org/10.1108/MBE-06-2017-0032>
- Raudla, R. (2012). The use of performance information in budgetary decision making by legislators: Is Estonia any different? *Public Administration*, 90(4), 1000–1015. <https://doi.org/10.1111/j.1467-9299.2012.02041.x>
- Raudla, R., & Savi, R. (2015). The use of performance information in cutback budgeting. *Public Money and Management*, 35(6), 409–416. <https://doi.org/10.1080/09540962.2015.1083685>
- Raudla, R., Taro, K., Agu, C., & Douglas, J. W. (2016). The impact of performance audit on public sector organizations: The case of Estonia. *Public Organization Review*, 16(2), 217–233.
- Reddy, T. R. (2012). Total quality management and knowledge management integrations in library and information centers: a study. *Journal of Research in International Business and Management*, 2(11), 292–298.
- Reichard, C. (2016). Can training help to make politicians more active users of performance information? *Public Money and Management*, 36(7), 481–482. <https://doi.org/10.1080/09540962.2016.1237119>
- Rhee, D. Y. (2014). The impact of performance information on congressional appropriations. *Public Performance and Management Review*, 38(1), 100–124. <https://doi.org/10.2753/PMR1530-9576380105>
- Rich, R. F., & Oh, C. H. (2000). Rationality and use of information in policy decisions: A search for alternatives. *Science Communication*, 22(2), 173–211.
- Robinson, M., & Last, M. D. (2009). *A basic model of performance-based budgeting*. Washington D.C.: International Monetary Fund.
- Roehrich, J. K., Lewis, M. A., & George, G. (2014). Are public–private partnerships a healthy option? A systematic literature review. *Social Science & Medicine*, 113, 110–119.
- Rosenbaum, P. R. 1999. Choice as an alternative to control in observational studies. *Statistical Science*, 14, 259–304.
- Rossi, P. H., & Freeman, H. E. (1985). *Evaluation: A systematic approach*. (3rd ed.) Beverly Hills, CA: Sage.
- Saliterer, I., & Korac, S. (2013). Performance information use by politicians and public managers for internal control and external accountability purposes. *Critical Perspectives on Accounting*, 24(7–8), 502–517. <https://doi.org/10.1016/j.cpa.2013.08.001>
- Saliterer, I., & Korac, S. (2014). The discretionary use of performance information by different local government actors—Analysing and comparing the predictive power of three factor sets. *International Review of Administrative Sciences*, 80(3), 637–658. <https://doi.org/10.1177/0020852313518170>
- Samuelson, W., & Zeckhauser, R. J. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1, 7–59.
- Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, 80(1), 1–22.
- Schaie, K. W. (2001). Cognitive aging. In N. J. Smelser, & P. B. Baltes (Eds), *International Encyclopedia of the Social and Behavioral Sciences*. Oxford, UK: Pergamon, 2072–2075.
- Shiller, R. J. (2015). *Irrational exuberance*. Princeton, NJ: Princeton University Press.
- Schmidle, T. (2011). Performance management. *Public Performance & Management Review*, 35(2), 370–389. <https://doi.org/10.2753/PMR1530-9576350206>

- Schul, Y., Mayo, R., & Burnstein, E. (2004). Encoding under trust and distrust: The spontaneous activation of incongruent cognitions. *Journal of Personality and Social Psychology*, 86, 668–679.
- Senne, T. A., & Rikard, G. L. (2002). Experiencing the portfolio process during the internship: A comparative analysis of two PETE portfolio models. *J. Teach. Phys. Educ.* 21, 309–336.
- Seppänen, A., Laitinen, M., Rajala, T., & Laihonon, H. (2019). Public value and performance of libraries. *Qualitative and Quantitative Methods in Libraries (QQML)* 6, 271–281.
- Shiratori, R. (1995). The politics of electoral reform in Japan. *International Political Science Review*, 16(1), 79–94.
- Simon, H. A. (1957). *Models of man: Social and rational | Mathematical essays on rational human behavior in society setting*. New York: Wiley.
- Simon, H. A. (1978). Rational decision-making in business organizations. Nobel Memorial Lecture, 8 December, 1978
- Simon, H. A. (1979). *Models of Thought*. New Haven, CT: Yale University Press.
- Sinervo, L., & Haapala, P. (2019). Presence of financial information in local politicians' speech. *Journal of Public Budgeting, Accounting & Financial Management*, 31(4), 558–577. <https://doi.org/10.1108/JPBAFM-11-2018-0133>
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 18(2–3), 277–310.
- Smith, P. C. (1996). *Measuring outcome in the public sector*. New York: Taylor and Francis.
- Sinclair, D., & Zairi, M. (2000). Performance measurement: a critical analysis of the literature with respect to total quality management. *International Journal of Management Reviews*, 2(2), 145–168.
- Solesbury, W. (2001). *Evidence-based policy: Whence it came and where it's going*. London: ESRC UK Centre for Evidence-Based Policy and Practice.
- Soper, J. C., & Walstad, W. B. (1983). On measuring economic attitudes. *The Journal of Economic Education*, 14(4), 4–17.
- Spano, A., & Asquer, A. (2011). Performance review and assessment of public managers: Some evidence from local governments in Italy. *International Journal of Public Administration*, 34(7), 461–472. <https://doi.org/10.1080/01900692.2011.576108>
- Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30, 337–360.
- Sterck, M., & Scheers, B. (2006). Trends in performance budgeting in seven OECD countries. *Public Performance & Management Review*, 30(1), 47–72. <https://doi.org/10.2753/pmr1530-9576300103>
- Steup, M. (2005). Epistemology. In Edward N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Spring 2014 ed.).
- Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy*, 69, 213–225.
- Stickers, K. W. (2003). Transaction, development, and capacity: Commentary on “Toward a transactional theory of decision making.” *Journal of Economic Methodology*, 10(2). <https://doi.org/10.1080/1350178032000071048>
- Struthers, C. W., Eaton, J., Santelli, A. G., Uchiyama, M., & Shirvani, N. (2008). The effects of attributions of intent and apology on forgiveness: When saying sorry may not help the story. *Journal of Experimental Social Psychology*, 44(4), 983–992.
- Svensson, J., & Yanagizawa, D. (2009). Getting prices right: The impact of the market information service in Uganda. *Journal of the European Economic Association*, 7(2–3), 435–445.

- Syed-Ikhsan, S. O. S., & Rowland, F. (2004). Knowledge management in a public organization: A study on the relationship between organizational elements and the performance of knowledge transfer. *Journal of Knowledge Management*, 8(2), 95–111. <https://doi.org/10.1108/13673270410529145>
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755–769.
- Talbot, C. (2000). Performing 'performance'—a comedy in five acts. *Public Money and Management*, 20(4), 63–68. <https://doi.org/10.1111/1467-9302.00238>
- Talbot, C. (2005). Performance management. In *The Oxford Handbook of Public Management*.
- Tan, H. T. R., & Harvey, G. (2016). Unpacking the black box: A realist evaluation of performance management for social services. *Public Management Review*, 18(10), 1456–1478. <https://doi.org/10.1080/14719037.2015.1112422>
- Tantardini, M., & Kroll, A. (2016). The role of organizational social capital in performance management. *Public Performance and Management Review*, 39(1), 83–99. <https://doi.org/10.1080/15309576.2016.1071163>
- Taylor, J. (2006). Statutory bodies and performance reporting: Hong Kong and Singapore experience. *Public Organization Review*, 6, 289–304. <https://doi.org/10.1007/s11115-006-0018-8>
- Taylor, J. (2009). Strengthening the link between performance measurement and decision making. *Public Administration*, 87(4), 853–871.
- Taylor, J. (2011). Factors influencing the use of performance information for decision making in Australian state agencies. *Public Administration*, 89(4), 1316–1334. <https://doi.org/10.1111/j.1467-9299.2011.02008.x>
- Taylor, J. (2015). Closing the rhetoric–reality gap? Employees' perspective of performance management in the Australian public service. *Australian Journal of Public Administration*, 74(3), 336–353.
- Taylor, J. (2020). Public Officials' Gaming of Performance Measures and Targets: The Nexus between Motivation and Opportunity. *Public Performance & Management Review*, preprint, 1–22.
- Taylor, R. N., & Dunnette, M. D. (1974). Influence of dogmatism, risk-taking propensity, and intelligence on decision making strategies for a sample of industrial managers. *Journal of Applied Psychology*, 59, 420–423.
- ter Bogt, H. (2001). Politicians and output-oriented performance evaluation in municipalities. *European Accounting Review*, 10(3), 621–643. <https://doi.org/10.1080/713764629>
- ter Bogt, H. J. (2004). Politicians in search of performance information? Survey Research on Dutch aldermen's use of performance information. *Financial Accountability and Management*, 20(3), 221–252. <https://doi.org/10.1111/j.0267-4424.2004.00387.x>
- Terhune, K. W., & Kennedy, J. L. (1963). *Exploratory analysis of a research and development game*. Princeton: Princeton University Press.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, 4(3), 199–214.
- Thaler, R. H., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
- Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. Allen Lane.
- Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative Inquiry*, 17(6), 511–521.

- Thorndike, A. N., Sonnenberg, L., Riis, J., Barraclough, S., & Levy, D. E. (2012). A 2-phase labeling and choice architecture intervention to improve healthy food and beverage choices. *American Journal of Public Health, 102*(3), 527–533.
- Tillema, S., Mimba, N. P. S., & Van Helden, G. J. (2010). Understanding the changing role of public sector performance measurement in less developed countries. *Public Administration and Development, 30*(3), 203–214.
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty, 5*(4), 297–323.
- Vakkuri, J. & Meklin, P. (2006). Ambiguity in Performance Measurement—A Theoretical Approach to Organisational Uses of Performance Measurement. *Financial Accountability and Management, 22*(3), 235–250.
- Vakkuri, J. (2010). Struggling with ambiguity: Public managers as users of NPM-oriented instruments. *Public Administration, 88*(4), 999–1024.
- Van Dooren, W & Van de Walle, S. (Eds). (2011). *Performance Information in the Public Sector: How It Is Used*. Houndmills, UK: Palgrave Macmillan.
- Van der Kolk, B., & Kaufmann, W. (2018). Performance measurement, cognitive dissonance and coping strategies: Exploring individual responses to NPM-inspired output control. *Journal of Management Control, 29*(2), 93–113.
- Van Dooren, W. (2006). Performance measurement in the Flemish public sector: A supply and demand approach. *Faculteit Sociale Wetenschappen, 96*, 1–366. <https://lirias.kuleuven.be/handle/1979/205>
- Van Dooren, W. (2008). Nothing new under the sun? Change and continuity in the twentieth-century performance movements. In *Performance Information in the Public Sector* (pp. 11–23). Palgrave Macmillan UK.
- Van Dooren, W. (2011). Better performance management: Some single- and double-loop strategies. *Public Performance and Management Review, 34*(3), 420–433. <https://doi.org/10.2753/PMR1530-9576340305>
- Van Dooren, W., Bouckaert, G., & Halligan, J. (2010). *Performance management in the public sector*. London: Routledge.
- Van Helden, J. (2016). Literature review and challenging research agenda on politicians' use of accounting information. *Public Money and Management, 36*(7), 531–538. <https://doi.org/10.1080/09540962.2016.1237162>
- Van Helden, J., & Reichard, C. (2019). Making sense of the users of public sector accounting information and their needs. *Journal of Public Budgeting, Accounting and Financial Management, 31*(4), 478–495. <https://doi.org/10.1108/JPBAFM-10-2018-0124>
- Van Helden, J., Johnsen, Å., & Vakkuri, J. (2008). Distinctive research patterns on public sector performance measurement of public administration and accounting disciplines. *Public Management Review, 10*(5), 641–651.
- Van Maanen, J., & Schein, E. (1979). *Towards a theory of organizational socialization*. Cambridge: MIT Press.
- Van Thiel, S., & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public performance & management review, 25*(3), 267–281.

- Verhoest, K., & Wynen, J. (2018). Why do autonomous public agencies use performance management techniques? Revisiting the role of basic organizational characteristics. *International Public Management Journal*, 21(4), 619–649. <https://doi.org/10.1080/10967494.2016.1199448>
- Vedung, E. (2010). Four waves of evaluation diffusion. *Evaluation*, 16(3), 263–277.
- Veggeland, N. (2007). *Paths of Public Innovation in the Global Age: Lessons from Scandinavia*. Cheltenham and Northampton, Mass., Edward Elgar.
- Vinnari, E. M., & Näsi, S. (2008). Creative accrual accounting in the public sector: ‘Milking’ water utilities to balance municipal budgets and accounts. *Financial Accountability & Management*, 24(2), 97–116.
- Vohs, K. D., Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *Science*, 314(5802), 1154–1156.
- Wagner, R. K. (1991). Managerial problem solving. In R. J. Sternberg & P. A. Frensch (Eds), *Complex problem solving: Principles and mechanisms* (pp. 159–184). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Walsh, D., & Downe, S. (2005). Meta-synthesis method for qualitative research: A literature review. *Journal of Advanced Nursing*, 50(2), 204–211.
- Wang, X. H. (2008). Convincing legislators with performance measures. *International Journal of Public Administration*, 31(6), 654–667. <https://doi.org/10.1080/01900690701641232>
- Ward, T. (2015). The Irish parliament and the scrutiny of departmental performance reports. *Public Money and Management*, 35(2), 153–160. <https://doi.org/10.1080/09540962.2015.1007713>
- Watkins, C. E., & Terrell, F. (1988). Mistrust level and its effects on counseling expectations in Black client–White counselor relationships: An analogue study. *Journal of Counseling Psychology*, 35(2), 194.
- Weaver, R. K. (1986). The politics of blame avoidance. *Journal of Public Policy*, 6(4), 371–398.
- Weigel R. H. (1983). Environmental attitudes and the prediction of behavior. In N. R. Feimer & E.S. Geller (Eds), *Environmental Psychology: Directions and Perspectives*. New York: Praeger.
- Weiss, C. (1983) Ideology, interests, and information: The basis of policy positions. In D. Callahan, B. Jennings (Eds), *Ethics, the social science, and policy analysis*. New York: Plenum Press.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public administration review*, 39(5), 426–431.
- Wholey, J., Newcomer, K., & Associates (1989). *Improving government performance: Evaluation strategies for strengthening public agencies and programs*. San Francisco, CA: Jossey–Bass.
- Wilensky, H. L. (1967). *Organizational intelligence*. New York: Basic Books
- Williamson, A., & Snow, D. (2014). From accountability to decision-making? Budgeting with mandated performance measures. *International Journal of Public Administration*, 37(4), 202–214. <https://doi.org/10.1080/01900692.2013.809593>
- Willoughby, K. G. (2004). Performance measurement and budget balancing: State government perspective. *Public Budgeting and Finance*, 24(2), 21–39. <https://doi.org/10.1111/j.0275-1100.2004.02402002.x>
- Wilson, W. (1887). The study of administration. *Political Science Quarterly*, 2(2), 197–222.
- Wimbush, E. (2011). Implementing an outcomes approach to public management and accountability in the UK—are we learning the lessons? *Public Money and Management*, 31(3), 211–218. <https://doi.org/10.1080/09540962.2011.573237>
- Wook Lee, J., & Jik Cho, Y. (2011). In search of a better fit: Exploring conditions for successful performance-based human resources management in US federal agencies. *International Review of Public Administration*, 16(3), 89–112.

- Wollmann, H. (2003). Evaluation in Public-Sector Reform: Towards a “Third Wave” of Evaluation? In Hellmut Wollmann (Ed.), *Evaluation in Public-Sector Reform: Concepts and Practice in International Perspective* (pp. 1–11). Cheltenham: Edward Elgar.
- Wright, B. E., & Pandey, S. K. (2010). Transformational leadership in the public sector: Does structure matter? *Journal of Public Administration Research and Theory*, 20(1), 75–89. <https://doi.org/10.1093/jopart/mup003>
- Wynen, J., & Verhoest, K. (2016). Internal performance-based steering in public sector organizations: Examining the effect of organizational autonomy and external result control. *Public Performance and Management Review*, 39(3), 535–559. <https://doi.org/10.1080/15309576.2015.1137769>
- Yamamoto, T. (2008). Asset impairment accounting and appraisers: Evidence from Japan. *Appraisal Journal*, 76(2), 179–188.
- Yang, K., & Hsieh, J. (2007). Managerial effectiveness of government performance measurement: Testing a middle-range model. *Public Administration Review*, 67(5), 861–879.
- Ye, L., & Ni, X. (2016). Assessing agency-level performance evaluation reform in China: Can it truly serve as a management innovation? *Chinese Public Administration Review*, 7(1), 7–34. <https://doi.org/10.22140/cpar.v7i1.96>
- Yetano, A. (2013). What drives the institutionalization of performance measurement and management in local government? *Public Performance and Management Review*, 37(1), 59–86. <https://doi.org/10.2753/PMR1530-9576370103>
- Zak, P. J. (2010-12-16). *Moral markets: The critical role of values in the economy*. Princeton University Press. p. 158. ISBN 9781400837366.
- Zhang, H., van de Walle, S., & Zhuo, Y. (2016). Does trust in the performance measurement organization influence how public managers use performance information? *Public Performance and Management Review*, 40(2), 409–430. <https://doi.org/10.1080/15309576.2016.1239583>
- Zinyama, T., & Nhema, A. G. (2016). Zimbabwe performance-based budgeting: Concepts and success factors. *Public Policy and Administration Review*, 4(1), 33–60. <https://doi.org/10.15640/ppar.v4n1a3>

Appendices

Appendix 1.

Table 5. Individual attributes associated with performance information use

<i>Umbrella concept for the set of individual attributes associated with use or non-use</i>	<i>Specific individual attributes influencing performance information use</i>
Demographic attributes	Age (Saliterer & Korac, 2014)
	Education (Askim, 2009)
	Race (McBeath et al., 2015)
	Gender (Kroll, 2015a)
	Salary as an income (Bourdeaux, 2006)
	Experience (work or political) (Kroll, 2015a; Askim, 2009)
	Number of years employed (McBeath et al., 2015)
	Private sector experience (Lu, 2007)
	Public administration background (Kroll, 2013)
	Political party (Lu & Willoughby, 2015)
	Generalist leader (Moynihan & Pandey, 2010)
Attitudinal mental models	Acceptance (Andrews, 2006)
	Responsiveness to organizational change (McBeath et al., 2015)
	Unwillingness to use (Charbonneau & Bellavance, 2015)
	Willingness to learn from other townships (Hoontis & Kim, 2012)
	Innovation mindedness (McBeath et al., 2015)
	Cynicism (Kroll, 2015a)
	Negativity bias (Nielsen & Moynihan, 2016)
	Internal focus (Kroll, 2013)
Pro-NPM attitude (Kroll, 2015a)	
Mental models of identity	Identity of council chair (Buylen & Christiaens, 2016)
	Manager identity (Kroll, 2014)
	Partisanship (or political ideology) (Demaj, 2017)

Competence-related mental models	Understanding how program activities affect performance (Barnow & Heinrich, 2010)
	Incomprehension (Charbonneau & Nayer, 2012)
	Inability to understand what to do with performance information or how to interpret it (Charbonneau & Nayer, 2012)
	Evaluation style called operation-conscious style that focused on the activities of the manager and the organizations (ter Bogt, 2001)
	Competence (Angiola & Bianchi, 2015)
	Lack of expertise (Charbonneau & Nayer, 2012)
	Ability or inability (e.g., to understand performance information and influence decisions) (Andrews, 2006)
	Capabilities (Ayers et al., 2013)
	Emotional intelligence (Kroll, 2014)
	Managerial competence (Molin et al., 2017)
	Networking behavior (i.e., ability to network) (Kroll, 2015a)
	Clear expectations (Hoontis & Kim, 2012)
	Information needs of the public manager (Cepiku et al., 2017)
	Competence with measurement (Cunningham & Harris, 2001)
Motivational mental models	Motivation to embrace comparisons (Ammons & Rivenbark, 2008)
	Prosocial motivation (Kroll, 2015a)
	Intentions (Kroll, 2015b)
	Public service motivation (Kroll, 2014)
Power	Authority (Andrews, 2006)
	Leadership (Cavalluzzo & Ittner, 2004)
	Political position (opposition or government) (Reichard, 2016)
	Dominance of party whip (Ward, 2015)
	Managerial power (Rajala et al., 2018)

Appendix 2.

Table 6. External environment factors as contextual factors associated with performance information use

<i>Umbrella concept for the set of environmental factors associated with use or non-use</i>	<i>First subcategory of specific environmental factors influencing performance information use</i>
Environmental factors	<p>Service users include the following:</p> <ul style="list-style-type: none"> – Intensity of use among stakeholders (Amirkhanyan et al., 2014) – Service user diversity (Moynihan & Hawes, 2012) – Heterogeneity of preferences in the population (Bourdeaux, 2006) – Diversity in population (McAdam & Walker, 2004) – Population size (Saliterer & Korac, 2014) – Population density (Charbonneau & Bellavance, 2015) – Average unemployment rate (Lu & Willoughby, 2015) – Demographic variables (Amirkhanyan et al., 2014) – Dissatisfaction with service among citizens (Quinlivan et al., 2017) – Perception of the utility of local government services (Quinlivan et al., 2017) – Lack of competing demands for public services (Moynihan, 2009) <hr/> <p>Service production environment including the following:</p> <ul style="list-style-type: none"> – Size of the city (i.e., geographical size) (Kroll, 2014) – Environmental complexity (Moynihan & Hawes, 2012) – Dysfunctions in the external information environment (Rabovsky, 2014) – Performance measurement environment (McBeath et al., 2015) – Benchmark organizations (public, private, or third sector) Angiola & Bianchi (2015) – Stakeholder attributes (age, race, education level) (McBeath et al., 2015) – Stakeholders' experience with performance measurement (Abdel-Maksoud et al., 2015) <hr/> <p>Political environment, including the following:</p> <ul style="list-style-type: none"> – Political pressure for improvements (Raudla, 2015) – Political context (Pollitt, 2006) – Political influences (Lee & Cho, 2011). – Political competition or devilsation (Charbonneau & Bellavance, 2015) – Political conflict Bourdeaux (2006) – Pressure toward politicians (Raudla, 2015) – Political fragmentation (Buylen & Christiaens, 2016) – Constituent positions formed from particular performance information (Bourdeaux, 2008) – Governor–legislator confrontations (Lu & Willoughby, 2015) – Stakeholder conflicts (Moynihan & Hawes, 2012) – Political (mis)use risk (Kroll, 2015a) – Political constraints (Johansen & Zhu, 2013) – Political support (Kroll, 2015a) – Political bargaining (Joyce, 1996)

Environmental factors	<p>Economic environment including the following:</p> <ul style="list-style-type: none"> – Fiscal stress/pressure (Paulsson, 2006) – Tax rate (Buylen & Christiaens, 2016) – Economic downturns and upturns (Hou et al., 2011) – Market competition (Johansen & Zhu, 2013) – Income per capita (Moynihan & Pandey, 2010)
	<p>Social rules of the society including the following:</p> <ul style="list-style-type: none"> – Legislation supporting use (Bourdeaux, 2006) – Lack of external requirements for reporting (Hildebrand & McDavid, 2011) – Legislation (i.e., laws and regulations) (Lu & Willoughby, 2015). – External results control (Verhoest & Wynen, 2016) – Coercive isomorphism (Tan & Harvey, 2016) – Raised interest on accountability in society (Kloot, 1999) – Professional influence and professionalism (Lee & Cho, 2011) – Objections by professional associations or unions (Liner et al., 2001, p. 15) – Workers' unionization (Folz et al., 2009) – Influence from citizens' associations (Depiku et al., 2017) – Stakeholder pressure (Tillema et al., 2010) – Community values (Quinlivan et al., 2017) – Concerns about community issues (Quinlivan et al., 2017) – Greater demand for performance information among stakeholders (Moynihan, 2009) – Media interest (Raudla et al., 2015)

Appendix 3.

Table 7. Organizational structures as contextual factors associated with performance information use

<i>Umbrella concept for the set of organizational factors associated with use or non-use</i>	<i>Specific organizational factor influencing performance information use</i>
Current organizational structure: Labor of division in public administration (includes service production tasks and arrangements, financial duties, administrative jobs, and features of the tasks, authorities related to the responsibilities, roles and personnel required to do the above tasks, administrative organs,	Government structure (Pollitt, 2006).
	Administrative levels (Hammerschmid et al., 2013)
	Governing board structure (Rabovsky, 2014)
	Governance arrangements (Quinlivan et al., 2017).
	Organizational design (Spano & Asquer, 2011)
	Policy fields (Askim, 2007)
	Public sector type (Abdel-Maksoud et al., 2015)
	Organization or agency size (Abdel-Maksoud et al., 2015)
	Organizational boundaries (Döring, Downe, & Martin, 2015)
	Complexity (McAdam & Walker, 2004)
	Complex programs (Moynihan, 2009)
	Diversity (McAdam & Walker, 2004)
	Centralization (Moynihan, Pandey, & Wright, 2012)
	Central oversight (Ammons & Roenigk, 2015)
	Decentralization (Van Dooren, 2011)
	Autonomy or lack of autonomy (Moynihan, 2009)
	Decision flexibility and level of discretion (Depiku et al., 2017)
	Decision-making authority (Abdel-Maksoud et al., 2015)
	Devolved decision authority (Ammons & Roenigk, 2015)
	Decisions with difficult trade-offs (Demaj, 2017)
	Governor's institutional power (Lu & Willoughby, 2015)
	Legal distance from government (Verhoest & Wynen, 2016)
	Task structure (Pollitt, 2006)
	Primary tasks (Verhoest & Wynen, 2016)
	Red tape (Moynihan, Pandey, & Wright, 2012)
	Political regime (Johansson & Siverbo, 2009)
	Term limits (Lu & Willoughby, 2015)
	Professional tasks (Kroll, 2013)
	Service specifications (Amirkhayan, 2011)
	Participative process (Lu, 2008)
	Independent agencies participation (Lu & Willoughby, 2015)
	Stakeholder involvement (e.g., politicians, citizens, companies) Chen (2017)
	Process in which stakeholder involvement is arranged (Lu, 2011)
	Employee involvement (Depiku et al., 2017)
	Reliance on partnerships (Moynihan & Hawes, 2012)
	Co-production (Lam & Wang, 2014)
	Co-operative arrangements (Saliterer & Korac, 2013)
	Marketization (Saliterer & Korac, 2013)
	Contract relationship (Amirkhayan et al., 2014)
	Collaboration type (Poocharoen & Wong, 2016)
Use of professional help (Lu et al., 2015)	
Complicated service delivery system (Holloway et al., 2012)	
Fragmented and networked service system (Holloway et al., 2012)	
Top-down or bottom-up approach to performance budgeting (Lu et al., 2015)	

Current organizational structure: Labor of division in public administration (includes service production tasks and arrangements, financial duties, administrative jobs, and features of the tasks, authorities related to the responsibilities, roles and personnel required to do the above tasks, administrative organs,	Shared responsibilities for performance-based budgeting among stakeholders (Lu et al., 2015).
	Adequate personnel resources (Depiku et al., 2017)
	Income sources (Verhoest & Wynen, 2016)
	Agencies financed with fees (Paulsson, 2006)
	Formal role Askim (2009)
	Executive formal participation (Lu & Willoughby, 2015)
	Legislative formal participation (Lu & Willoughby, 2015)
	Political composure of political organs (Rabovsky, 2014)
	Split legislative control (Lu & Willoughby, 2015)
	Senior performance councils (Ayers et al., 2014)
	Adversarial role of budget staff (Moynihan & Pandey, 2010).
	Number of core businesses (Lu, 2007)
	Staff (Lu et al., 2015).
	Non-permanent employees (Lee & Cho, 2011)
	Job-person fit (Saliterer & Korac, 2014).
	Role diversity (Andersen & Moynihan, 2016)
Organizational role (planner, decision maker) (Kroll, 2014)	
Education diversity (Andersen & Moynihan, 2016)	

Appendix 4.

Table 8. Management structures and integration of organizational structures as contextual factors associated with performance information use

<i>Umbrella concept for the set of organizational factors associated with use or non-use</i>	<i>Specific organizational factor influencing performance information use</i>
Management structures (includes planning, budgeting, measuring, analyzing, controlling, and managing, as well as the roles and authority required to do these tasks)	Managerial level (Abdel-Maksoud et al., 2015)
	Appointed agency management (Lu, 2007)
	Performance management routines (Askim, 2009)
	Control of government (Bourdeaux, 2006)
	Agenda control and control tightness (Moynihan & Ingraham, 2004)
	Mission orientation (Moynihan, 2009)
	Goal orientation (de Lancer Julnes & Holzer, 2001)
	Strategic planning (Charbonneau & Bellavance, 2015)
	Poor strategic planning (Nomm & Randma-Liiv, 2012)
	Strategy (Joyce, 1996)
	Balanced scorecard (McAdams & Walker, 2004)
	Vague and ambiguous goals (Barbato & Turri, 2017)
	Goal clarity (Moynihan, 2009).
	Coherence between goals and measures (Molin et al., 2017)
	Sophistication of performance measurement systems (Kroll & Proeller, 2013)
	Effectiveness of performance indicators (Hatry, 2008)
	Measures concentrate on issues that can be influenced (Grizzle, 1981)
	Coherent conceptual level (Kim & Kang, 2016)
	Meaningful performance measures (Cavalluzzo & Ittner, 2004)
	Performance measurement density Melkers & Willoughby, 2005)
	Nature of the measured activity (Barbato & Turri, 2017)
	Systematic review processes (Duisenberg, 2016)
	Collecting adequate data (de Lancer Julnes & Holzer, 2001).
	Lack of meaningful measures (Liner et al., 2001, p. 15)
	Difficulties in choosing performance measures (Cavalluzzo & Ittner, 2004)
	Measuring costs (Amirkhayan, 2011)
	Accounting problems (Charbonneau & Nayer, 2012)
	Benchmarking (Angiola & Bianchi, 2015)
	More than token measurement (Ammons & Roenigk, 2015)
	Numerous and complicated measures (Kasdin, 2010)
	Variety and amount of performance information (Hibbard et al., 1997)
	Time spent on performance measurement (Lu, 2007)
	Ambiguity in performance measures (Poister & Streib, 1999)
	Different measures tracking the same outcomes provide different results (Nicholson-Crotty et al., 2006)
	Complexity of performance being measured (Di Francesco, 1999)
	Performance measurement transparency (Melkers & Willoughby, 2005)
	Performance information ownership (Kroll, 2014)
	Maturity of measurement system (Kroll, 2015a)
	Complications arising from indicator's churning (Charbonneau & Nayer, 2012)
	Degree to which measures are embedded into key management systems (Ammons & Rivenbark, 2008)
Political performance management (Van Dooren, 2011)	
Legislative mandates (Cavalluzzo & Ittner, 2004)	
Legislative authority over budget (Bourdeaux, 2006)	
Legislative engagement in oversight of performance information (Bourdeaux, 2006)	
Professional status of the legislature (Bourdeaux, 2006)	
Budget process is too time constrained (Raudla, 2012)	
Complicated funding system (Holloway et al., 2012)	
Appropriations and policy tools grants (Lee & Cho, 2011; Rabovsky, 2014)	

Management structures (includes planning, budgeting, measuring, analyzing, controlling, and managing, as well as the roles and authority required to do these tasks)	Budget format (Raudla, 2012)
	Budget session length (Bourdeaux, 2006)
	High-quality budget analysis (Wang, 2008)
	Size of the budget (Charbonneau & Bellavance, 2015)
	Performance-based funding (Rabovsky, 2014)
	Incentives and sanctions (Ammons & Roenigk, 2015)
	Reward expectation (Moynihan & Pandey, 2010)
	Lack of rules for incentive systems (Liner et al., 2001, p. 15)
	State employee pay (Moynihan & Ingraham, 2004)
	System dynamic modelling in the organization (Bianchi & Rivenbark, 2014)
	Learning forums (Moynihan, 2005)
	Appropriate data analysis (Hatry, 2008)
	Poor policy analysis (Nomm & Randma-Liiv, 2012)
	Difficulties in interpreting measures (Cavalluzzo & Ittner, 2004)
	Difficulties in connecting staff actions to results (Liner et al., 2001, p. 15)
	Audit (Lu et al., 2015)
	Online reporting (Lu et al., 2015)
Supportive or controlling use (Masal & Vogel, 2016)	
Integration of organizational structures	Including performance information in budget proposals (Demaj, 2017).
	Reporting processes are not in line with budget structures (Joyce & Levy, 2008)
	Objectives linked to indicators (Brusca & Montesinos, 2016)
	Aligning performance management practices (Ayers et al., 2014)
	Integrating goals into performance appraisal plans (Ayers et al., 2014)
	Enhancing collaboration among internal stakeholders (Ayers et al., 2014)
	Inter-administrative management (Jääskeläinen & Roitto, 2014)
	Measurement should be integrated into goal setting, strategic planning, and communication processes (Jääskeläinen & Roitto, 2014)
	Inadequate links between databases (Willoughby, 2004)
	Integration of the performance measures into different departments and processes (Depiku et al., 2017)
	Ability to connect measures to actions (Moynihan & Lavertu, 2012)
	Coordination of statewide performance measures (Lu & Willoughby, 2015)
	Social integration (structural capital) (Tantardini & Kroll, 2016)
	Managerial networking (Moynihan & Hawes, 2012)
	Structural insulation (Lee & Cho, 2011)
	Integrating results-oriented activities into all facets of the financial management process (Hill & Andrews, 2005)
	Integration of a performance management system into the management system (Depiku et al., 2017)
Broad involvement in all government levels and integration and communication into strategic planning and agency management (Cunningham & Harris, 2005)	
Ease of negotiations (Henderson & Bromberg, 2015)	
State actors' influence (Rabovsky, 2014)	

Appendix 5.

Table 9. Reforms and legitimacy of organizational structures as contextual factors associated with performance information use

Reforming/changing organizational structures (here, <i>leadership</i> is understood as leading change)	Capacity building (Lu & Willoughby, 2015)	
	Change routines (Johansson & Siverbo, 2009)	
	Reform time (Lu et al., 2015)	
	Support of reform (Bourdeaux, 2006)	
	Changes imposed by central government (Kloot, 1999)	
	Uniform changes (Nomm & Randma-Liiv, 2012)	
	Incremental rather than radical changes (Yetano, 2013)	
	Early adopter (Moynihan & Hawes, 2012)	
	Transformational leadership (Wright & Pandey, 2010)	
	Parliament has only limited role in making changes (Raudla, 2012)	
	Gubernatorial leadership (Moynihan & Ingraham, 2004)	
	Legislative leadership (Moynihan & Ingraham, 2004)	
	Agency leadership (Moynihan & Ingraham, 2004)	
	Central agency leadership (Moynihan & Ingraham, 2004)	
	Managerial implementation style (Saliterer & Korac, 2013)	
	Legitimizing implementation style (Saliterer & Korac, 2013)	
	Managerial–contractual interventions (Kim & Kang, 2015)	
	Capability-enhancing interventions (Kim & Kang, 2015)	
	Market-type interventions (Kim & Kang, 2015)	
	Systematic development (Saliterer & Korac, 2014)	
	Random development (Saliterer & Korac, 2014)	
	Promotion of new language (Yetano, 2013)	
	Agile public organizations (Van Dooren, 2011)	
	Establishment of planning and reporting routines (Yetano, 2013)	
	Good implementation in terms of training and dialogue (Depiku et al., 2017)	
	Training opportunities (Amirkhanyan et al., 2014)	
	Need for training (Charbonneau & Nayer, 2012)	
	Training effects (Abdel-Maksoud et al., 2015)	
	Difficulty in selecting performance measures (Abdel-Maksoud et al., 2015)	
	Legitimacy of current organizational structures	Employees do not support performance information system (Poister & Streib, 1999)
		Managers do not support performance information system (Poister & Streib, 1999)
		City council does not support performance information system (Poister & Streib, 1999)
		Legislative support for performance information use (Lu et al., 2015)
Executive support for performance information use (Lu et al., 2015)		
Management support for performance information use (Lu et al., 2015)		
Leadership support for performance information use (Kroll, 2015a)		
Leadership commitment (Lee & Lee, 2015)		
Staff buy-in (Lu et al., 2015)		
Lack of champions (Aristiqueta & Zarook, 2011)		
Agreement from the audit criteria (Raudla, 2015)		
Dissonant demands for performance (Ho, 2007)		
Advocacy (Moynihan, 2009)		
Advocates of performance information (Pollitt, 2006)		
Measurement is guided by theory (Grizzle, 1981)		
Support capacity (Kroll, 2015a)		
Expenditures on support services (Bourdeaux, 2006)		
Resource decency (Tan & Harvey, 2016)		

Appendix 6.

Table 10. Organizational culture and performance capacity and results as contextual factors associated with performance information use

Organization's culture (includes norms, group habits, and traits seen in organization)	Performance-oriented organizational culture (Lu et al., 2015)
	Hierarchical culture (Kroll, 2015a)
	Rational culture (Saliterer & Korac, 2013)
	Group culture (Kroll, 2015a)
	Developmental culture (Kroll, 2014)
	Shared culture (Moynihan & Hawes, 2012)
	Internal performance culture (Lu, 2007)
	Internal orientation (Kroll, 2013)
	Social norms (Kroll, 2015b)
	Peer exchange (Kroll, 2013)
	Organizational traits (Amirkhanyan et al., 2014)
	Performance management is collectively viewed as an instrument of innovation rather than an instrument of control (Bourgon, 2008)
	Trust (Tandardini & Kroll, 2016)
	Political-administrative relations (Depiku et al., 2017)
	Staff devoted to the evaluation of performance measures (de Lancer, Julnes & Holzer, 2001)
	Limited use encourages limited use (Charbonneau & Nayer, 2012)
	Municipal experience in performance measurement (Foltz et al., 2009)
	Performance capacity and results (input, process, output, outcome, productivity, and cost-effectiveness results)
Past performance (Smith, 1995)	
Total revenues, total enrollment, research, change in graduation rates (Rabovsky, 2014)	
Perverse behavioral effects (Spano & Asquer, 2011)	
Current administrative capabilities and capacities (Johansson & Siverbo, 2009)	
Management capacity (Amirkhanyan et al., 2014)	
Central administrative capacity (Moynihan & Hawes, 2012)	
Personnel capacity (Lu et al., 2015)	
Social capacity (Andrews, 2006)	
Organizational social capital (Tandardini & Kroll, 2016)	
Relational capital (Tandardini & Kroll, 2016)	
Cognitive capital (Tandardini & Kroll, 2016)	
Slack (Johansson & Siverbo, 2009)	
Bills produced (Bourdeaux, 2006)	
Service timeliness and disruptions (Amirkhayan, 2011)	
Client impact (Amirkhayan, 2011)	
Internal requirements (Depiku et al., 2017)	
Financial health (Saliterer & Korac, 2013)	
Chaos (Charbonneau & Nayer, 2012)	
Individual effects (Spano & Asquer, 2011)	
Instability (Nomm & Randma-Liiv, 2012)	
Expectancy disconfirmation (Moynihan, 2015)	
Effective adjustment of performance expectations for external influences (Courty et al., 2005)	

Appendix 7.

Table 11. Information producer, information channel, and information as contextual factors associated with performance information use

<i>Umbrella concept for the set of performance information use context factors associated with use or non-use</i>	<i>What is being expected?</i>
Information producer	Credible information source (Pollitt, 2006) Auditor expertise (Raudla et al., 2015) Audit team open to dialogue (Raudla et al., 2015) Performance review involvement (Spano & Asquer, 2011) Independent information producer (Zhang et al., 2016)
Information channel	Quality of the information channel (Abdel-Maksoud et al., 2015) "The information/data systems are hard to use, the challenge is to locate the correct and essential information" (Rajala et al., 2018) Different information systems do not communicate with each other (Rajala et al., 2018) Access to information (Rajala et al., 2018) Effective communication (Wang, 2008) Information system capacity (Lu et al., 2015) Existence of computer-based information systems (Hasan & Hasan, 1997) Communicating performance information (Jääskeläinen & Roiito, 2014) Executive verbal explanations (Buylen & Christiaens, 2016) Information brokers (Jorge et al., 2016) Citizen outreach (Charbonneau & Bellavance, 2015) Information systems (Cavalluzzo & Ittner, 2004) Simple reporting (Angiola & Bianchi, 2015)
Information	Information or measure type (Abdel-Maksoud et al., 2015) Quality of performance information is adequate (Lu et al., 2015) Information is useful and beneficial (Kroll, 2015a) Performance information needs to be provided together with cost information (Baekgaard, 2015) Responds to user needs (Bogt, 2004) Timeliness (Pollitt, 2006) Concrete recommendations (Raudla et al., 2015) Explanations concerning the shortcomings (Raudla et al., 2015) Too long and cumbersome performance documents (Raudla, 2012) Performance information availability (Saliterer & Korac, 2014) Broad comparative performance measures are more convincing than reflexive benchmarks (Charbonneau & Van Ryzin, 2015) Performance information is too simplistic, not realistic or valid (Charbonneau & Nayer, 2012) Performance information is useless or uninteresting (Charbonneau & Nayer, 2012) Lack of comparability between municipalities and their performance (Charbonneau & Nayer, 2012) Other indicators are more important (Charbonneau & Nayer, 2012) Content of the information (Depiku et al., 2017) Information overload (Ezzamel et al., 2004) Trustworthiness of performance reports (Kang et al., 2009) Presentation of information (Hatry, 2008) Range of information (Moynihan & Ingraham, 2004)

Original publications

Post-print

Rajala, T., Laihonon, H., & Vakkuri, J. (2018). Shifting from output to outcome measurement in public administration arguments revisited. In E. Borgonovi, E. Anessi-Pessina, & C. Bianchi (Eds.), *Outcome-based performance management in the public sector* (pp. 3–23). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-57018-1_1

Shifting from Output to Outcome Measurement in Public Administration - Arguments Revisited

Tomi Rajala, University teacher, Ph.D. student

University of Tampere, School of Management, FI-33014 University of Tampere, Finland

email tomi.rajala@uta.fi

Harri Laihonen, Research director

University of Tampere, School of Management, FI-33014 University of Tampere, Finland

email harri.laihonen@uta.fi

Jarmo Vakkuri, Professor

University of Tampere, School of Management, FI-33014 University of Tampere, Finland

email jarmo.vakkuri@uta.fi

Corresponding author: Tomi Rajala, University of Tampere, School of Management, FI-33014 University of Tampere, Finland, email tomi.rajala@uta.fi

Abstract

Moving to outcome-based measurement systems in the public sector has been difficult. In this article, we examine the contingent decision-making arguments stimulating output instead of outcome measurement in public management. Based on an argumentative literature review, we conclude that there exist several contingent arguments encouraging politicians and public managers to stick with outputs while ignoring outcomes in performance measurement. Mapping out these arguments contributes to understanding the difficulties in implementation of outcome-based measurement and management systems. This understanding is highly useful in performance management research and policy practice. We also suggest that these contingent arguments may be considered proposals for the future research in the area of public financial management and public sector performance measurement.

Keywords: Outcomes, Outcome-based performance measurement systems, Politicians, Public managers, Contingent arguments

1.1. Introduction

Outcome information is relevant to the public sector because it reports whether or not public services are producing desired outcomes to the society (Hatry 2005). This information is important to public managers seeking to improve performance as well as to other stakeholders such as voters and politicians aiming for a better societal welfare. However, it has remained extremely complicated to establish an

outcome-oriented measurement system for public sector purposes and practices. Governmental organizations continue to use output measures more often than outcome measures (Ferlie et al. 2005). This study aims to explore contingent arguments as to why output measures are sometimes preferred over outcome measures in the public sector.

There are many contingent arguments intrinsic to public sector behavior and performance measurement that have been acknowledged and addressed by the previous research literature (e.g., Smith 1996). However, the previous research has not been able to systematically and comprehensively understand contingent decision-making arguments for resisting the shift from an output-based to an outcome-based measurement system. Our study aims to fill this research gap by gathering these arguments together and presenting them under two topics: (1) pursue of value for money (second section) and (2) control of legitimacy (third section). In the value for money section, we are searching for arguments indicating that output information would provide more value for money than outcome information because the costs are bigger and/or benefits are not so evident in the latter. In the chapter, dealing with control of legitimacy, we are looking for arguments implicating that output information would provide more control over legitimacy than outcome information. Since legitimation (Bouckaert 1993) and value for money (Jackson 2012) are important parts of performance information use, this approach can be seen as justified.

The research follows constructivist epistemology (e.g., Guba and Lincoln 1998) and the logic of abductive reasoning (e.g., Peirce 1998). The contingent arguments are constructed from scientific arguments presented in performance management literature. As an example of our method, consider the following scenario: “scientist x has noted in her research that outcome measurement is not supported by the current entity-based information systems, and scientist y has stated that the current information systems support output measurement.” From these statements, we form a contingent argument stating that current information systems support output measurement and do not support outcome measurement. By forming this argument, we would create one possible answer to our research question. In the conclusion section, we place all these arguments under broader categories constructed in this study.

We conducted an argumentative literature review in order to construct these arguments. An argumentative literature review examines literature selectively in order to support an argument already established in the literature. The aim of this type of literature review is to develop a body of literature that establishes a contrarian viewpoint (Kennedy 2007). The contingent arguments presented in this article form a contrarian viewpoint to outcome measurement advocates listing the benefits of outcome measuring (see, e.g., Hatry 2005) compared to output measuring. These arguments describe mental models that argue against the use of outcome measures and favor output indicators. By “contingent” it is indicated that the truth value of every argument is contextual, not universal. Furthermore, these

arguments are not meant to be normative in any way, and their truth value may even be untrue. The point of this article is to raise discussion on whether or not outcome measurement can have negative effects in public sector.

As a main theoretical contribution, this research gathers together the dispersed arguments describing the possible reasons why output measuring is often more established than outcome measurement in the public sector. These reasons are described in the contingent arguments, and they can be understood as problems and limitations that incentivize public sector actors not to adopt outcome measurement. Understanding of these reasons is one of the first steps in better comprehending non-use of outcome measurement. The second significant theoretical contribution is the recognition of the future research questions proposed in this study. We are hoping that future research would examine empirically whether or not these arguments are capable of explaining why the implementation of the outcome measurement has been difficult.

Figure 1.1 depicts the structure of this article. Following the introduction, first, we examine whether or not outcome measurement provides less value for money than output measurement. Second, we investigate how legitimacy is affected by these two types of measurement. The final part includes conclusions and future research questions.

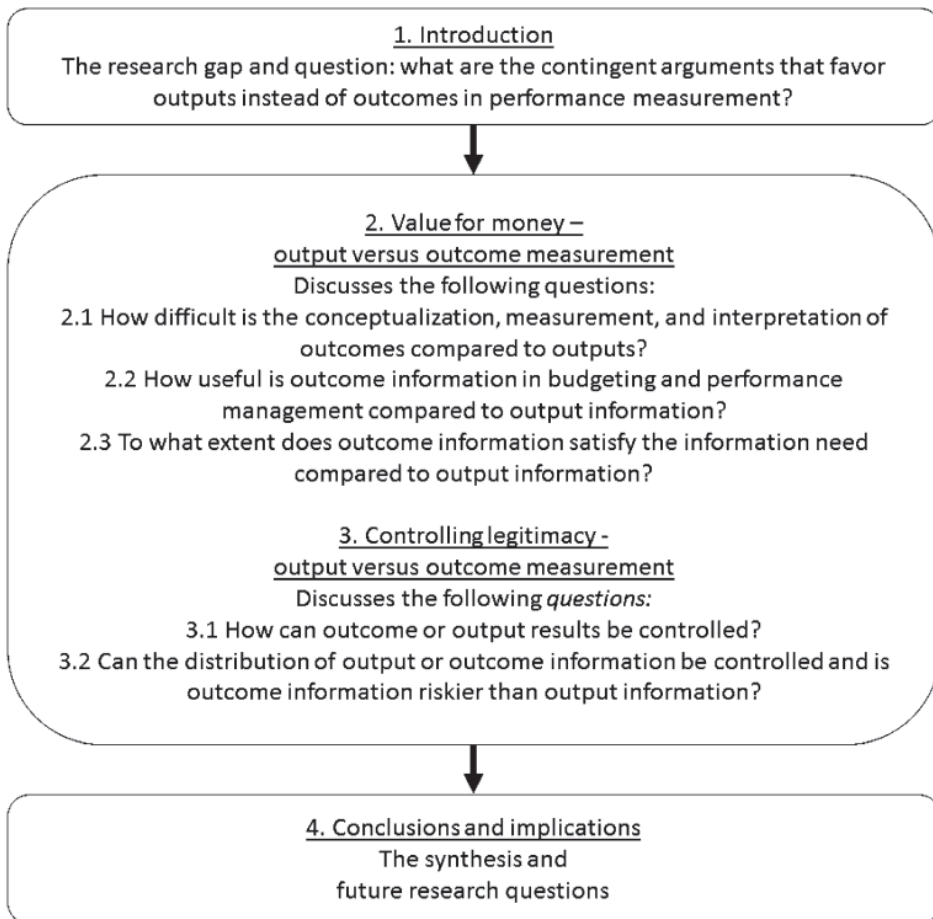


Fig. 1.1 The overall structure of the research

1.2 Value for Money—Output Versus Outcome Measurement

Value for money refers to the ideal combination of whole-life costs of public services and fitness for the purpose of meeting the user’s requirements (Jackson 2012). Both the outcome and output measurement have a purpose and cost. The important question in the context of this article is whether or not output measurement could provide more value for money than outcome measurement. More value for money would here indicate that output measures are cheaper to produce and/or they offer more fitness for purpose than outcome measures according to scientific arguments. Thus, in this section, we are displaying arguments presented in the literature which state that the costs of outcome measurement can be high. We are also representing previous academic statements asserting that the costs of output measuring are often low. In addition, we present ideas expressing that it might not be possible to determine the purpose of producing outcome information, or the fitness for purpose may be lacking

when outcomes are attached to the performance management system and budgeting system. Thus, the use of outputs can seem more favorable in these situations to public managers and politicians.

1.2.1 Outputs Versus Outcomes: The Conceptualization, Measurement, and Interpretation

As noted by Hatry (2006), two different types of outcome exist: intermediate and end outcomes. Intermediate outcomes lead to the ends desired, but they are not ends in and of themselves. The end outcomes are the desired results of the program according to the program customers and citizens (Hatry 2006). Vedung (1997) identifies an additional outcome type: immediate outcomes. These outcomes happen right after the actions are taken, whereas intermediate outcomes occur in the causal chain following immediate outcomes (Vedung 1997). However, in Hatry's (2006) typology, immediate outcomes can be placed under the concept of intermediate outcomes without breaking any theoretical assumptions of intermediate outcomes.

Vedung (1997) has also recognized more comprehensively the complex nature of outcomes, naming several different outcome types as follows:

1. Outcomes for customers and society.
2. Quantitative and qualitative outcomes.
3. Subjective and objective outcomes.
4. Short- and long-run outcomes.
5. External outcomes and internal outcomes.
6. Positive and negative outcomes.
7. Expected and unexpected outcomes.
8. Intended and unintended outcomes.

Yeung and Matheison (1998) identify six different areas where outcomes can appear: economic performance, competitiveness, education, health, environment, democracy, and freedom. Outcomes can also occur at the program/service level, the agency/organizational level, the state/community level, or any combination thereof (e.g., Martin 1997). Thus, some outcomes for public sector bodies may occur at a societal rather than at an organizational level (McGill 2001).

The public sector has multiple outcomes at different hierarchical levels; meanwhile, different units on the same hierarchical level can consider different outcomes to be important to them. Ideally, a goal congruency exists among these different goals set by units within the public sector organization, but this might not always be the case. For example, the General Accounting Office (1997, p. 6) has reported that "mission fragmentation" is common at the federal government level in the USA, and it is difficult to get stakeholders to think beyond their own program operations to the diversity of activities related to the common outcome.

The complex nature associated with conceptualizing the outcome becomes evident when all of these outcome categories and areas are combined in the public sector's hierarchical and horizontal dimensions. Bounded rationality and lack of know-how can magnify the complications associated with outcome definition. For example, researchers have documented problems in understanding the difference between outputs and outcomes (e.g., Dugan and Herson 2002). As Hogwood and Gunn (1992, p. 17) point out, the distinction between outcomes and outputs is often blurry in practice.

Outputs describe what the public sector does (Rosen 1993), whereas outcomes describe the effects that have been caused directly and indirectly by the outputs (e.g., Talbot 2010). The decision considering outputs of a program boils down to the following question: What are the goods and services the public sector wants to produce? No matter what the output is, all the different kinds of outcomes listed above in various areas can occur. By comparing one output to another output, only two things are compared. However, whenever the possible outcomes of the two outputs are compared, the comparison becomes far more complex. Identifying all the relevant outcomes can take more time and effort than the output identification. Value for money may not be achieved because the cost of conceptualization is too high in outcome measurement.

If outcomes cannot be defined, they cannot be measured. A key question is whether or not decision-makers agree on which types of outcomes are the most optimal and which ones can be ignored. In the public sector, a high level of subjectivity often relates to outcomes, and thus, even reaching consensus on outcomes can be difficult (e.g., Kurunmaki and Miller 2011) because outcomes can be multi-dimensional, qualitative by nature, and impossible to represent as a single quantitative measure (Carlin and Guthrie 2003). The difficulties in defining the outcomes are well known in different countries (e.g., Carlin and Guthrie 2003). For instance, Heinrich (2002) learned that federal managers in the USA considered the outcome-based performance management systems to be conceptually and practically one of their most difficult tasks to complete. By comparison, outputs often are easier to identify (e.g., Bandy 2011, p. 76). Moreover, usually outputs have to be defined because public sector produces outputs. However, it is not required to identify and name the important and unimportant outcomes in order to get the public production up and running.

It is common that different political parties strive for different societal outcomes (e.g., Spoon and Kluver 2014). Political outcome goals may inhibit outcome measuring if these outcome goals are contradictory. The inconsistency between the policy objectives set by politicians and the goals of executive agencies also creates problems in the public sector (e.g., Smith 1995). The policy objectives are contested both among politicians as well as between politicians and managers (Agranoff and McGuire 2001). There is often little consensus as to what constitutes outcome because the large number of diverse stakeholders in the public services holds different expectations toward these services (Smith 1996). The complex nature of outcomes can intensify this rivalry whenever outcomes are policy objectives instead of

outputs. The increased number of options in conceptualization simply offers more possibilities for disagreement. Investing resources in outcome measurement can lead to conflicts and inefficient resource use if outcomes cannot be defined or measured. If this scenario occurs, the purpose of outcome measurement cannot be identified properly.

Contingent argument: outputs are easier, cheaper, and less time-consuming to define and conceptualize than outcomes (nature of outcome, nature of output, and conflict orientation).

1.2.1.1 Technical Aspects of Measurement

Measuring outcomes can be astonishingly difficult (Smith 1996), and on the other hand, calculating outputs is usually fairly straightforward (Newcomer 2007). Obtaining information about the intervening variables affecting outcomes causes problems (Miller and Fox 2007), whereas variables that influence output production can be monitored and detected more easily in many cases. The problem with intervening variables in the context of outcomes is related to the problem of monitoring citizens and societal activities round-the-clock holistically. Anthony and Young (1988, p. 608) summarize this common problem that plagues many outcome performance measures (here “social indicator” means “outcome”):

A social indicator is a broad measure of output which significantly reflects the work of the organization. Unfortunately, few social indicators can be related to the work of a single organization because in almost all cases they are affected by exogenous forces; that is, forces other than those of the organization being measured. The crime rate in a city may reflect the activities of the police department and the court system, but it is also affected by unemployment, housing conditions, and other factors unrelated to the effectiveness of these organizations... Social indicators are so nebulous, so difficult to obtain on a current basis, so little affected by current program efforts, and so much affected by external influences that they are of limited usefulness in day-to-day management...

In public and private contexts, some outcomes cannot be measured directly, and some outcomes are not measurable at all (e.g., van der Valk and van Iwaarden 2011; Newcomer 2015). In such cases, output measurement has to suffice (Cunningham and Harris 2001). In addition, factors related to the reliability, validity, and accuracy of the measurement may favor output instead of outcome measurement. As McPhee (2005) points out, the reported information on output tends to be better than for outcomes because output indicators are often more appropriate and the method for output measurement is usually more robust and reliable than for outcomes. Outcomes are often encompassed by values of quality and satisfaction (e.g., Chalmers 2008). They are considered to be more difficult to measure than outputs (Currstine et al. 2008), which are often more quantitative by nature. As a consequence, outcomes are not utilized nearly as often as outputs in practice (e.g., OECD 2013).

One permanent problem with outcomes is that the impact of any governmental action requires information about what would have happened to citizens if those actions were not executed (e.g.,

Heinrich 2002). When assessing the effectiveness of government actions, it is difficult to isolate and measure the real difference between doing something and doing nothing. Again, unmeasured intervening variables and moderator variables can explain outcomes better than measured ones. On the contrary, doing something and doing nothing can be seen rather easily on production volumes (e.g., Rosen 1993).

Outputs are usually cheaper to measure (e.g., Marks 2005). In contrast, measuring all the relevant aspects of the outcomes would normally require rigorous quantitative and qualitative methodology with subjects over prolonged time periods (Schalock 2001). In such cases, outcome measurement can require extensive resources or tunnel vision focusing on some aspects while ignoring other critical aspects associated with outcomes by reducing the number of indicators used to track outcome development in the name of measurement efficiency (Lowe 2013). If extensive outcome measurement is chosen, frontline workers will often have to devote more time to reporting and less time to service production (e.g., Keevers et al. 2012). The question here might simply be whether we want to focus on reporting or on the actual service production.

Conflicts about the usefulness of different approaches to public sector performance measurement do exist (Harrison et al. 2012). Agreeing on appropriate performance measures has proven difficult in hybrid organizations (Kurunmaki et al. 2003). Performance measures are generally not neutral in the public sector context (e.g., Van de Walle and Van Dooren 2010), and there exist divergent opinions about the right performance indicators among politicians and between politicians and managers (Agranoff and McGuire 2001). The development of political debate dictates the assessment of public sector performance (Stewart and Walsh 1994). Because there can be a lack of consensus regarding the right indicators, performance measurement can cause dysfunctional effects (van Thiel and Leeuw 2002).

According to Chan (2004), outcome measures often are more difficult to define than output measures. Lack of consensus from the right outcome measures often occurs (Newcomer 2015). Again, the complexity of outcomes offers more possibilities to measure, meaning that there are more alternatives from which to choose the performance indicators. The diversity of preferences typical to public sector can utilize these alternatives to create conflicts. These conflicts can induce more costs and mean that the purpose of outcome measurement cannot be defined.

Contingent argument: outcome measurement causes more costs and conflicts about the right measures than output measurement. Meanwhile, outcomes cannot be measured comprehensively, whereas outputs can be (nature of outcome, nature of output, and conflict orientation).

1.2.1.2 Interpretation Problems in Outcome Results

The analysis of causes explaining the outcomes is often more complicated than the analysis of the activities producing the outputs (e.g., Pollitt and Bouckaert 2004). The linkage and interaction between

outcomes, outputs, intervening variables, and/or moderator variables makes the interpretation of outcome results considerably more difficult (e.g., Mascarenhas 1996). How different policies, programs, and agencies contribute to outcomes is often unclear (Newcomer 2015). The fact that a perception in a complex issue depends on when, where, and who is making the interpretation does not help in outcomes analysis (e.g., Kunda 1990; Van Maanen and Schein 1979). These complications in detecting the causes explaining the outcomes are called “the attribution problem” in the previous literature, and several researches have addressed this problem (e.g., Taro 2015).

Complex outcomes may cause information overload for politicians and public managers and therefore deteriorate the quality of decisions (c.f. Hahn et al. 1992). For this reason, simpler output information may seem a better choice (e.g., Chaston 2011). Kristensen et al. (2002) point out that politicians and public managers can devote focused attention to only limited areas at a time, and these actors have constraints on how much information they can utilize in their decision-making. If the outcomes form from complex processes, the decision-maker may not be able to utilize all of the information relating to the outcome achievement. Outcome measurement may not provide sufficient value for money if it deteriorates the quality of decisions or the information remains unused because we cannot interpret it properly or without conflicts and debates.

Contingent argument: interpretation of outcomes is more difficult and more prone to produce conflicts than interpretation of outputs (nature of outcome, nature of outputs, and conflict orientation).

1.2.1.3 Output as a Reflection of an Outcome

When outputs reflect outcomes reasonably well, the value of producing additional outcome measures can be very low. It is therefore important to analyze how well current outputs can approximate outcomes of a public organization (e.g., Smith 1996). For example, in the private sector, there often is no need to measure customer outcomes because the customers’ valuation of the products and services reflects their willingness to pay for them (Smith 1996). In a similar fashion, the willingness to use, for example, the public sport facilities or the public parks can tell us something about the customer valuation placed on these types of goods and services.

Contingent argument: outputs reflect outcomes adequately and accordingly; there is no need for outcome information (nature of outcome and nature of output).

1.2.2 How Outputs and Outcomes Connect to Budgeting and Performance Management

The budget demonstrates whether or not there is political and managerial demand for outcome measurement (c.f. Greenwood et al. 1977). The budget process reminds us that there are opportunity costs for measuring outcomes. The interesting question in the resource allocation context is what makes

output measurement more desirable than outcome measurement. The answer is threefold, relating to costs, current information systems, and the purpose of such systems.

From the perspective of budgetary allocations, the decision-making problem is about comparing uses of resources to the added value of measurements. Assuming that the added value is perceived to be similar between the two types of measurement, cost of measurement defines the choice. Performance information often focuses on output levels because these are easy and less costly to define, measure, and analyze. By comparison, program outcomes tend to be much more difficult to identify, measure (e.g., Robichau and Lynn 2009), and analyze (e.g., Mascarenhas 1996). For these reasons, outputs may be preferred. If the outcome information is more expensive than output information, the former would have to provide more value than the latter in order to be the first choice of the decision-maker when these two types of measurement are competing on the same resources.

*Contingent argument: From the budgetary perspective output measurement may provide more value for money because the nature of outcomes is problematic and more expensive to measure and analyze comprehensively (**opportunity costs/competition for resources**).*

The necessity of coupling the budget process to outcomes may be one reason explaining why output measuring is preferred over outcomes in the public sector. As Kristensen et al. (2002) stated government budgeting and financial systems may currently only be capable of generating rudimentary matches of resources and outcomes. It is difficult to put a price tag on outcomes (e.g., Midwinter 2009). In contrast, calculating the cost of achieving required output levels is a rather established procedure (e.g., Anderson et al. 2000). There is thus a twofold problem with connecting outcomes to budgets. The problem with cost calculation is more fundamental because it has to be solved before any information system can be built. However, the problem involving information systems should not be understated either. Performance measurement in the public context is often based around the traditional vertical hierarchies of government departments and developed within individual legal entities (e.g., Ryan and Walsh 2004). Outputs are often created in these entities, whereas end outcomes are not (e.g., Mayne 2007). Systems supplying information on the costs and benefits of working across accounting entities would be needed in order to do pooled budgets that assign resources to service outcomes, for instance (Hodges 2012). Thus, information systems may not support outcome measurement.

*Contingent argument: outcome measurement requires too many investments in information systems, while output measurement does not (**information system**).*

Difficulties in cost calculations mostly relate to the fact that outcomes may not be as accurately specified and measured as outputs. Also, the causal link between inputs and outcomes is often more difficult to perceive than the link between inputs and outputs. Thus, uncertainty may arise over how changes in resource levels may affect overall performance (Kristensen et al. 2002). This problem causes

difficulties, especially at the state and community level. Connecting resources and outcomes to the change in indicator values in the state and community level is difficult at best, and it raises validity issues that are not encountered at the other levels (Rossi 1997). As stated by Kristensen et al. (2002), outcome budgeting raises many difficult questions. For instance, who should estimate the resources needed for outcomes that are a result of cross-sectional government operations? And should outcome targets be set first and then resources after the targets or vice versa (Kristensen et al. 2002)? These quite practical questions demonstrate the challenging problems outcome budgeting can generate (e.g., Grizzle and Pettijohn 2002).

Timeliness is an important feature of the performance management system (Heinrich 2002). The ability to provide timely feedback to public managers creates opportunities for performance improvements and for adjustments in budget allocations, service contracts, management practices, and training strategies. The challenge here is to provide outcome information in a timely manner so that it can be connected to day-to-day performance management. If outcome information cannot be used in operational performance management, the purpose of providing such information becomes compromised.

According to Heinrich (2002), federal agencies in the USA have found it particularly difficult to transform their long-term missions or strategic goals into annual performance goals. These federal agencies have also found predicting the level of performance results attained over a shorter term to be particularly challenging. For this reason, short-run rather than long-run measures are normally used in performance standards systems (Heinrich 2002). The indication here seems to be that outcomes cannot be utilized if they require long-run measures, as they often do. A longer time frame is usually needed because evaluating how the programs have affected the outcome takes time (Bovaird 2014). Outputs, on the other hand, are more suitable for those performance management systems that aim to provide feedback with minimal lag from actual performance because outputs can usually be detected, measured, and reported more instantly and easily than outcome information. Outputs also can be used to control work more efficiently because, by defining outputs, the public managers and politicians actually define what is done at the operating level (e.g., Snell 1992). Outcomes, on the other hand, may provide more freedom to the frontline workers. For instance, it does not matter what precise actions are taken as long as customers are satisfied.

*Contingent argument: it is difficult to do budgets for outcomes and use outcome information in day-to-day performance management whereas outputs can be more easily connected to budgeting and performance management processes (**nature of outcomes and nature of outputs**).*

1.2.3 The Information Need

According to Dervin's (1983) sense-making approach, information needs arise from the gap that exists between the current situation and the desired situation, from the process that tries to make sense of the

current gap and from the efforts to bridge that gap. Put simply, information needs are conceived as individual attempts to answer questions and to make sense of a gap in order to move from the current situation to the desired situation (Dervin 1983).

If low information need causes problems for the adoption of outcome measurement, the politician or public manager fails to see outcome information as beneficial for four reasons. Firstly, outcome information perhaps cannot help the politicians and public managers to understand which outputs will produce certain outcomes. In this situation, the outcome information does not provide enough data on how to change the current system; therefore, this information does not lead to action. The lack of mutual congruence in the results analysis may also mean that the information would remain unused. Secondly, it might be that the political system and public managers cannot agree on what the desired situation, or outcome, should be. Thus, there exist multiple views on the situation, which leads to an inability to determine the kind of information is needed collectively in order to improve the quality of life in society. Thirdly, the lack of know-how in performance measurement can lower information needs if it is acknowledged that these limitations could deteriorate the quality of the information to a level where it is no longer useful. Finally, the information can be seen as a blame attractor and as uncontrolled risk if outcomes are not in control and transparent information is needed for the sake of legitimacy.

Contingent argument: outcome information is costly to produce and the value to the decision-maker is equal or less than output information, which is why output information is preferred (information need).

1.3 Controlling Legitimacy—Output Versus Outcome Measurement

It has been stated that the main motivation for the use of performance information is legitimacy-seeking rather than efficiency maximization (Modell 2001). For example, in symbolic use, the information can be used for legitimation purposes (Van de Walle and Van Dooren 2008). Thus, legitimacy can be considered an important aspect of the implementation of performance measurement, and therefore, it is justified to look how produced output and outcome information can be managed and controlled in order to gain legitimacy. Suchman (1995, p. 574) describes legitimacy as “a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions.” To Zimmerman and Zeitz (2002), this social system is a very broad concept that includes the operating environment in which the organization functions and which it needs to demonstrate consistency. Legitimacy can also mean the congruence between the organization’s activities and outcomes and society’s values, norms, and expectations (Ashforth and Gibbs 1990). In order to understand how legitimacy can be managed, it is important to take a more detailed look at the ability to control the output and outcome achievement, as well as the distribution of the performance information describing these achievements.

1.3.1 Controllability of Results

Gaining, maintaining, and repairing legitimacy may be more problematic when outcomes are reported instead of outputs. This difficulty rests in the fact that outputs often have higher controllability than outcomes (Irwin 1996).

*Contingent argument: due to higher controllability earning legitimacy with output rather than outcome information renders more control (**control of legitimacy**).*

Lack of control also leads to problems in accountability. It seems that using outputs in performance management could lead to situations where the government has fewer difficulties holding an agency accountable for delivering the agreed-upon outputs (Mayne 2007). Outcome measures, on the other hand, may be subject to multiple determinants, with the budget holder's activities representing just one. It is harder for the government to hold the agency accountable for outcome achievement if an agency has only partial control over outcomes (Mayne 2001). In a similar fashion, ministers cannot be held accountable if the outcomes are not within their control (e.g., Irwin 1996). In addition, the time frame related to outcomes is troublesome because there can be extensive time lags between resource use and performance outcomes (e.g., Bovaird 2014). Consequently, the attribution problems not only make it difficult to interpret outcome results but also produce problems to accountability (Mayne 2001).

Attribution of responsibility for outcomes becomes even more problematic when the services are supported by multiple funding sources or various providers, such as health service providers, measures are affected by so many determinants that change in outcomes cannot be attributed to the effectiveness of a specific program alone (U.S. Department of Health and Human Services 1997). The uncontrolled nature of outcomes can also raise another challenging question: whether or not it is politically or professionally wise to promise to deliver outcomes instead of outputs. It can be daunting to manage efforts to achieve outcomes that require actions across different agencies or will be achieved by more than one hierarchical level of government (Kristensen et al. 2002).

*Contingent argument: while outputs are under control, politicians or public managers have only partial control over outcomes and only partial accountability for the results. Thus, outcomes may offer less tools for principals to control performance (**control of accountability**).*

1.3.2 Controllability of Information

Van de Walle and Van Dooren (2010) note that information relates to power structures because any new information about the performance of organizational departments may have a significant effect on future budgets or staff allocations. Even the survival of the department within the wider organization can depend on performance information. It is therefore in organizational actors' interests to control information flows (Van de Walle and Van Dooren 2010).

If government agencies were to focus on outcomes, the stakeholders should understand that the agency is only one of many factors likely to affect outcomes (e.g., Schalock 2001). This recognition would explicitly indicate that public agencies have only partial control and, therefore, only partial accountability, according to the stakeholders (Hatry 1997, p. 2). Without this recognition from stakeholders, blame games and blame avoidance strategies will most likely play a role in performance management. However, the opposition versus government setting can prevent such recognition because political opposition can do little other than generate blame. They cannot hope to have an effective voice in the process of policy formulation so long as there is a majority government in the parliamentary system (e.g., Weaver 1986). Thus, whenever there exists opposition, there may also exist a need to control information.

It is an interesting question whether or not there is a larger need to control outcome information than information that describes outputs. To answer this question, we need to ask: What do the outcome and output information tell us, exactly? If the output goals are not met, it indicates that the government is not operating efficiently. However, it cannot be inferred from output levels that the government is doing the wrong things. Failing to achieve outcome goals more directly raises the question of whether or not the government is actually doing the wrong things. Thus, the comparison between output and outcome measurements relates to the comparison between efficiency and effectiveness, where efficiency is “doing things right” and effectiveness is “doing the right thing” (e.g., Gleason and Barnum 1982).

Doing the wrong thing is a more severe error than doing the right thing inefficiently. Taking the wrong actions not only wastes public resources but can also lower citizens’ well-being. Thus, outcome indicators have the potential to show more fundamental problems in government operations than output indicators. Moreover, doing the wrong things demonstrates problems in the political system and in the current government’s visions. For politicians and public managers, the rationale may be to think that nothing is worse than providing outcome information demonstrating that public sector is doing the wrong things. However, it is uneasy to demonstrate that the government is actually doing the wrong things because of the inherent ambiguity analysis of outcomes.

From the perspective of accountability, outputs provide no justification for failures. This observation could indicate that the agent accountable for the outputs would have a greater stake because no excuses for failure would be available when output levels are not achieved. If the existence of blame avoidance is assumed, there is a need to control output information. Outcomes, on the other hand, offer only partial control over results; however, this partial control also offers justification to fail. Understanding the nature of outcomes would therefore diffuse the blame, avoiding the need to use blame avoidance strategies. Thus, it is unclear which type of information can be a bigger threat to legitimacy. Ultimately, if people react to reality as they perceive it and not to reality itself (Lewin 1936), then the need to control different information types depends upon stakeholders’ reactions. In general, citizens tend to attach

outputs and outcomes to specific programs (Taro 2015). If so, then we are back to comparing the harmful consequences of effectiveness and efficiency information.

According to Wholey and Hatry (1992), public managers fear that elected officials, interest groups, and the media may use outcome information as fodder for attacks. The possible misuse of negative findings is a risk that comes with performance information (Wholey and Hatry 1992). This fear is not unjustified because it is a common phenomenon in politics (and in human behavior) that negative information produces more activity and impact than positive information (Rozin and Royzman 2001). This negativity bias encourages the avoidance of bad publicity and can influence the willingness to provide performance information.

Because outcomes are not under the control of politicians and public managers, transparency can generate bad publicity and adverse effects by putting poor results in the spotlight every time a partially uncontrollable outcome goal is not achieved. The effects of poor results depend on whether the public sector is applying full transparency (c.f. Rousseau 1772), direct transparency (c.f. Bryan 2010), or indirect transparency (c.f. Hood 2007). For these reasons, the ability to control information is linked closely to the chosen state of transparency.

Information can be controlled in two ways: by inhibiting information production or by controlling what information is delivered and to whom. The inability to determine who gets the information may lead to situations where outcome measuring will be inhibited because this is the only way to ensure that information about the negative results does not end up in the hands of opponents.

*Contingent argument: outcome information may expose more severe errors in public sector actions than output information. The inability to control the distribution and production of the outcome information in a transparent setting may attract too much blame and trigger dysfunctional behavior, conflicts, and blame games (**controllability of information and conflict orientation**).*

By evaluating the credit-claiming and blame-avoiding opportunities in different situations, it becomes apparent that politicians or public managers may choose to be loss averse, risk averse, risk neutral, or risk seeking. Depending on this choice, the arguments introduced in this research can be valued differently. For example, a risk-seeking politician might not care about the possibility of outcome measurement producing bad publicity or conflicts in the institution; by comparison, loss-averse politicians may care a great deal and make a choice accordingly. Typically, people are more loss averse (Tversky and Kahneman 1992), and politicians often choose to be risk averse (Weaver 1986). The constituencies and beneficiaries may also prefer that the results of the government program go unmeasured because this measurement could demonstrate that the program has actually been ineffective, of little value, or unimportant in achieving the desired effect or impact in the society (Kristensen et al. 2002).

*Contingent argument: loss-averse politicians and public managers try to avoid conflicts, professional and political disasters, resource wasting, and legitimacy losses. They will not promise to deliver outcome information because outcome results may be ticking time bombs that are beyond their control, at least partially (**loss aversion**).*

1.4 Conclusions and Implications

As a main contribution, we found several contingent arguments relevant to politicians and public managers. These arguments can be examined when the transition from output to outcome measurement is undermined (see Fig. 1.2). The arguments are linked to each other, and together, they form a complex network of issues that may influence the decision-makers to reject or resist outcome-based performance management.

We organized the wide array of arguments under two more general topics. We looked at whether or not outcome measurement provides value for money and improves legitimacy. However, we are fully aware that the arguments could have been organized differently. In order to develop our theoretical–conceptual framework, we now reorganize all the contingent arguments presented in Fig. 1.2 under broader categories:

1. Information (information need and information system).
2. Controllability (controllability of results and information).
3. Legitimacy.
4. Nature of outcome measurement (nature of outcomes).
5. Political conflict (competition for resources, conflict orientation, and loss aversion).

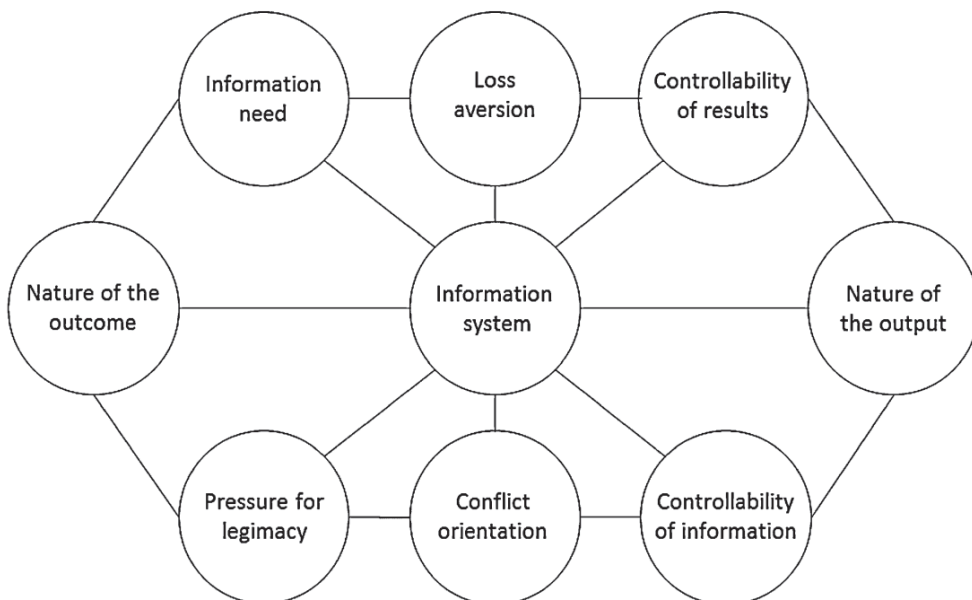


Fig. 1.2 The contingent arguments to ignore outcome measurement

These categories can be used to assess relevant contingent arguments that may inhibit the development of outcome measurement. Past and future research on outcomes can also be classified according to the above categories.

We do not assume that all the contingent arguments are present or assessed at the point of decision-making. In fact, it is probably more likely that some arguments are not even recognized by the decision-maker. We only assume that if at least one contingent argument is acknowledged and considered by politicians or public managers, outcome measurement may be rejected. We also acknowledge that these arguments can be valued differently by various politicians and public managers, and this valuing most likely varies among different decision-making situations. We also do not exclude the possibility of emergence: The whole could somehow differ from the parts (c.f. Morowitz 2002), and the final decision may deviate from the decision made purely based on weighting and calculating all the arguments, either favoring or opposing outcome measurement.

Finally, these contingent arguments can potentially enrich our theoretical and analytical view of the institutional practices and problems of developing outcome measurement in public administration. Therefore, we suggest that these contingent arguments be taken as proposals for the future research endeavors in the area of public financial management and public sector performance measurement (see Table 1.1). If these contingent arguments are supported by the empirical evidence in future research, they can be obstacles preventing the implementation outcome measurement. Taken into consideration the importance of outcome information to the stakeholders of the public sector, this threat cannot be taken lightly.

<i>Future research questions</i>
What type of information need can the outcome information satisfy, and how should the use and context of use be designed based on this need? (information need)
How would the current information systems have to be expanded or transformed if cross-sectional outcome measurement is to be connected to day-to-day management? (information system)
Do voters, political opposition, and other stakeholders understand that the government has only partial control over the outcomes, or do they exploit bad outcome results in order to gain advantage in elections and political decision-making by using blame games? (controllability of results)
Does transparency positively or negatively affect incentives to measure outcomes? (controllability of information)
Are voters, politicians, and public managers attaching outcome measurement to legitimacy in general? (control of legitimacy)
Do politicians, public managers, and voters recognize the complexity of outcomes and the political power associated with outcome measurement? (nature of outcome)

In which situations can outputs reflect outcomes adequately and accordingly? (nature of outputs)
How high is outcome measurement in the hierarchy of needs when resource allocations are considered by politicians and public managers? (opportunity costs/competition for resources)
How does a conflict-oriented environment affect incentives to measure outcomes? (conflict orientation)
Are public managers and politicians risk averse, loss averse, risk neutral, or risk-loving when it comes to deciding whether or not to implement the outcome measurement? (loss aversion)

References

- Agranoff, R., & McGuire, M. (2001). American federalism and the search for models of management. *Public Administration Review*, 61(6), 671–681.
- Anderson, G. F., Hurst, J., Hussey, P. S., & Jee-Hughes, M. (2000). Health spending and outcomes: Trends in OECD countries, 1960–1998. *Health Affairs*, 19(3), 150.
- Anthony, R. N., & Young, D. W. (1988). *Management control in nonprofit organizations*. Homewood, IL: Irwin.
- Ashforth, B. E., & Gibbs, B. W. (1990). The double-edge of organizational legitimation. *Organization Science*, 1(2), 177–194.
- Bandy, G. (2011). *Financial management and accounting in the public sector*. Abingdon, OX: Routledge.
- Bouckaert, G. (1993). Measurement and meaningful management. *Public Productivity and Management Review*, 17(1), 31–43.
- Bovaird, T. (2014). Attributing outcomes to social policy interventions—‘Gold standard’ or ‘fool’s gold’ in public policy and management? *Social Policy and Administration*, 48(1), 1–23.
- Bryan, F. M. (2010). *Real democracy: The New England town meeting and how it works*. Chicago: University of Chicago Press.
- Carlin, T., & Guthrie, J. (2003). Accrual output based budgeting systems in Australia The rhetoric-reality gap. *Public Management Review*, 5(2), 145–162.
- Chalmers, D. (2008). *Indicators of university teaching and learning quality*. Canberra: Australian Learning and Teaching Council, the Australian Government Department of Education.
- Chan, Y. C. (2004). Performance measurement and adoption of balanced scorecards: A survey of municipal governments in the USA and Canada. *International Journal of Public Sector Management*, 17(3), 204–221.
- Chaston, I. (2011). *Public sector management: Mission impossible?* Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Cunningham, G. M., & Harris, J. E. (2001). A heuristic framework for accountability of governmental subunits. *Public Management Review*, 3(2), 145–165.
- Currstine, T., Park, C. K., & Emery, R. (2008). Budgeting in Portugal. *OECD Journal on Budgeting*, 8(3), 7.

- Dervin, B. (1983). Information as a user construct: The relevance of perceived information needs to synthesis and interpretation. In S. Ward & L. Reed (Eds.), *Knowledge structure and use: Implications for synthesis and interpretation* (pp. 153–184). Philadelphia: Temple University Press.
- Dugan, R. E., & Herson, P. (2002). Outcomes assessment: Not synonymous with inputs and outputs. *The Journal of Academic Librarianship*, 28(6), 376–380.
- Ferlie, E., Pollitt, C., & Lynn, L. E. (2005). *The Oxford handbook of public management*. Oxford: Oxford University Press.
- General Accounting Office. (1997). *Managing for results: Opportunities for continued improvements in agencies performance plans*. Washington, DC: Government Printing Office. GAO/AIMD-99-215.
- Gleason, J. M., & Barnum, D. T. (1982). Toward valid measures of public sector productivity: Performance measures in urban transit. *Management Science*, 28(4), 379–386.
- Greenwood, R., Hinings, C. R., & Ranson, S. (1977). The politics of the budgetary process in English local government. *Political Studies*, 25(1), 25–47.
- Grizzle, G. A., & Pettijohn, C. D. (2002). Implementing performance based programme budgeting. *A System-Dynamics Perspectives, Public Administration Review*, 62(1), 51–62.
- Guba, E. G., & Lincoln, Y. S. (1998). Competing paradigms in qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd edn., pp. 191–215). Thousand oaks: Sage Publications.
- Hahn, M., Lawson, R., & Lee, Y. G. (1992). The effects of time pressure and information load on decision quality. *Psychology and Marketing*, 9, 365–378.
- Harrison, J., Rouse, P., & De Villiers, C. J. (2012). Accountability and performance measurement: A stakeholder perspective. *Journal of Centrum Cathedra: The Business and Economics Research Journal*, 5(2), 243–258.
- Hatry, H. P. (1997). Where the rubber meets the road: Performance measurement for state and local public agencies. *New Directions for Evaluation*, 1997(75), 31–44.
- Hatry, H. P. (2006). *Performance measurement: Getting results*. The Urban Institute: Washington, DC.
- Hatry, H. P. (2005). Results matter: Suggestions for developing country's early outcome measurement effort. In A. Shah (Ed.), *Handbook on public sector performance review*. Washington, DC: World Bank.
- Heinrich, C. J. (2002). Outcomes-based performance management in the public sector: Implications for government accountability and effectiveness. *Public Administration Review*, 62(6), 712–725.
- Hodges, R. (2012). Joined-up government and the challenges to accounting and accountability researchers. *Financial Accountability and Management*, 28(1), 26–51.
- Hogwood, B. W., & Gunn, L. A. (1992). *Policy analysis for the real world*. Oxford: Oxford University Press.
- Hood, C. (2007). What happens when transparency meets blame-avoidance? *Public Management Review*, 9(2), 191–210.

- Irwin, T. (1996). An analysis of New Zealand's new system of public sector management. Public management occasional papers no 9. Performance management in government. Contemporary Illustrations. Paris: Organisation for Economic Co-operation and Development.
- Jackson, P. (2012). Value for money and international development: Deconstructing myths to promote a more constructive discussion. Paris: Organisation for Economic Cooperation and Development.
- Kennedy, M. M. (2007). Defining a literature. *Educational Researcher*, 36(3), 139–147.
- Kunda, Z. (1990). The case for motivating reasoning. *Psychological Bulletin*, 108, 480–498.
- Keever, L., Treleaven, L., Sykes, C., & Darcy, M. (2012). Made to measure: Taming practices with results-based accountability. *Organization Studies*, 33(1), 97–120.
- Kristensen, J. K., Groszyk, W., & Bühler, B. (2002). Outcome-focused management and budgeting. *OECD Journal on Budgeting*, 1(4), 7–34.
- Kurunmaki, L., Miller, P., Keen, J., & Institute of Chartered Accountants in England and Wales. (2003). Health act flexibilities: Partnerships in health and social care. London: Centre for Business Performance.
- Kurunmaki, L., & Miller, P. (2011). Regulatory hybrids: Partnerships, budgeting and modernizing government. *Management Accounting Research*, 22(4), 220–241.
- Lewin, Kurt. (1936). *Principles of topological psychology*. New York: McGraw-Hill.
- Lowe, T. (2013). New development: The paradox of outcomes—the more we measure, the less we understand. *Public Money and Management*, 33(3), 213–216.
- Marks, M. M. (2005). In S. Mathison (Ed.), *Encyclopedia of evaluation* (287). Thousand Oaks, CA: Sage.
- Martin, L. L. (1997). Outcome budgeting: A new entrepreneurial approach to budgeting. *Journal of Public Budgeting, Accounting and Financial Management*, 9(1), 108.
- Mascarenhas, R. C. (1996). Searching for efficiency in the public sector: Interim evaluation of performance budgeting in New Zealand. *Public Budgeting and Finance*, 16, 13–27.
- Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *The Canadian Journal of Program Evaluation*, 16(1), 1.
- Mayne, J. (2007). Challenges and lessons in implementing results-based management. *Evaluation*, 13(1), 87–109.
- McGill, R. (2001). Performance budgeting. *International Journal of Public Sector Management*, 14(5), 376–390.
- McPhee, I. (2005). Outcomes and outputs: Are we managing better as a result?. In CPA National Public Sector Convention, 20.
- Miller, H. T., & Fox, C. J. (2007). *Postmodern public administration*. New York: M.E. Sharpe.
- Midwinter, A. (2009). New development: Scotland's Concordat—An assessment of the new financial framework in central–local relations. *Public Money and Management*, 29(1), 65–70.
- Modell, S. (2001). Performance measurement and institutional processes: A study of managerial responses to public sector reform. *Management Accounting Research*, 12(4), 437–464.

- Morowitz, H. J. (2002). *The emergence of everything: How the world became complex*. Oxford: Oxford University Press.
- Newcomer, K. E. (2007). Measuring government performance. *International Journal of Public Administration*, 30(3), 307–329.
- Newcomer, K. E. (2015). From outputs to outcomes. In Guy & Rubin (Eds.) *Public administration evolving: From foundations to the future* (pp. 125–158). New York: Routledge.
- OECD. (2013). *Together for better outcomes: Engaging and involving SME taxpayers and stakeholders*. Paris: Author.
- Peirce, C. S. (1998). *The essential Peirce selected philosophical writings*. Bloomington, IN: Indiana University Press.
- Pollitt, C., & Bouckaert, G. (2004). *Public management reform: A comparative analysis*. Oxford: Oxford University Press.
- Robichau, R. W., & Lynn, L. E., Jr. (2009). The implementation of public policy: Still the missing link. *Policy Studies Journal*, 37(1), 21–36.
- Rosen, E. D. (1993). *Improving public sector productivity: Concepts and practice*. Newbury Park, CA: Sage Publications.
- Rossi, P. (1997). Program Outcomes: Conceptual and measurement issues. In E. Mullen & J. Magnabosco (Eds.), *Outcomes measurement in human services* (pp. 20–34). Washington, DC: NASW Press.
- Rousseau, J. J. (1772/1985). *The government of Poland*, translated, with an introduction and notes by Willmoore Kendall, Indianapolis: Hackett Publishing Co.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320.
- Ryan, C., & Walsh, P. (2004). Collaboration of public sector agencies: Reporting and accountability challenges. *International Journal of Public Sector Management*, 17(7), 621–631.
- Schalock, R. L. (2001). *Outcome-based evaluation*. New York: Kluwer Academic/Plenum Publishers.
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 18(2–3), 277–310.
- Smith, P. C. (1996). *Measuring outcome in the public sector*. New York: Taylor and Francis.
- Snell, S. A. (1992). Control theory in strategic human resource management: The mediating effect of administrative information. *Academy of Management Journal*, 35(2), 292–327.
- Spoon, J. J., & Klüver, H. (2014). Do parties respond? How electoral context influences party responsiveness. *Electoral Studies*, 35, 48–60.
- Stewart, J., & Walsh, K. (1994). Performance measurement: When performance can never be finally defined. *Public Money and Management*, 14(2), 45–49.
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610.
- Talbot, C. (2010). *Theories of Performance: Organizational and service improvement in the public domain*. Oxford: Oxford University Press.

- Taro, Kulli. (2015). *The attribution problem in performance measurement in the public sector: Lessons from performance audits in Estonia*. Tallinna: TUT Press.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- U.S. Department of Health and Human Services. (1997). *Strategic Plan*. Washington, DC: Author.
- Van Dooren, W., & de Walle, Van. (2008). *Performance information in the public sector: How it is used*. Basingstoke, Hampshire: Palgrave Macmillan.
- Van de Walle, S., & Van Dooren, W. (2010). How is information used to improve performance in the public sector? Exploring the dynamics of performance information. In K. Walshe, G. Harvey, & P. Jas (Eds.), *Connecting knowledge and performance in public services: From knowing to doing* (pp. 33–55). Cambridge: Cambridge University Press.
- Van der Valk, W., & Van Iwaarden, J. (2011). Monitoring in service triads consisting of buyers, subcontractors and end customers. *Journal of Purchasing and Supply Management*, 17(3), 198–206.
- Van Maanen, J., & Schein, E. H. (1979). *Toward a theory of organizational socialization*. Cambridge, Mass: M.I.T.
- Van Thiel, S., & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public Performance and Management Review*, 25(3), 267–281.
- Vedung, E. (1997). *Public policy and program evaluation*. New Jersey: Transaction Publishers.
- Weaver, R. K. (1986). The politics of blame avoidance. *Journal of Public Policy*, 6(04), 371–398.
- Wholey, J. S., & Hatry, H. P. (1992). The case for performance monitoring. *Public Administration Review*, 52, 604–610.
- Yeung, O. M., & Mathieson, J. A. (1998). *Global benchmarks: Comprehensive measures of development*. Washington, DC: Brookings Institution Press.
- Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond survival: Achieving new venture growth by building legitimacy. *Academy of Management Review*, 27(3), 414–431.



Exploring challenges of boundary-crossing performance dialogues in hybrids

Tomi Rajala¹ · Harri Laihonen¹ · Jarmo Vakkuri¹

Published online: 27 August 2019
© The Author(s) 2019

Abstract

Literature on performance management in hybrid organisations remains lacking, even though they increasingly are providing public services. This study created new knowledge on performance management in hybrid organisations by answering a question that addressed what kinds of challenges are in hybrid organisation's boundary-crossing performance dialogues. Boundary-crossing performance dialogues here refer to performance management discussions between representatives of public and private sector organisations forming the hybrid. To answer the proposed question, a case study approach based on interviews and documents was chosen. Using inductive content analysis, the study identified several challenges that occurred during the boundary-crossing dialogues. This contributes to existing research on performance management in hybrid organisations. For practitioners, the results provide insights for tackling the specific problems of performance management in hybrid organisations.

Keywords Hybrid organisation · Performance dialogue · Performance management · Performance governance · Performance information

1 Introduction

Hybrid organisations increasingly are providing public services (e.g., McGuire 2002; Christensen and Laegreid 2011). A typical hybrid organisation comprises public and private actors who, through collaboration and cooperation, provide public services for citizens (Johanson and Vakkuri 2017). Extant literature claims

✉ Tomi Rajala
tomi.rajala@tuni.fi

Harri Laihonen
harri.laihonen@tuni.fi

Jarmo Vakkuri
jarmo.vakkuri@tuni.fi

¹ Faculty of Management and Business, Tampere University, 33014 Tampere, Finland

that traditional entity-based performance management systems are inadequate in hybrid organisations and, therefore, performance management systems based on inter-organisational collaboration are being adopted by these types of organisations (Agostino and Arnaboldi 2018). Here, *performance management* means a ‘continuous process of identifying, measuring and developing the performance of individuals and teams and aligning performance with the strategic goals’ (Aguinis 2009, p. viii). Adopting inter-organisational performance management systems aims to turn performance management into performance governance that operates within and across public, private and third sector interfaces and relies on inter-sectoral co-processes (Bouckaert and Halligan 2008). One way to create inter-organisational collaboration in performance management is to use dialogue that engages all necessary actors in developing an organisation’s performance (Agostino and Arnaboldi 2015). In this article, dialogue that addresses performance management tasks is called *performance dialogue*.

More specifically, performance dialogue is a phenomenon in which ‘participants jointly interpret performance information and discuss it while identifying the actions needed to manage the performance according to this information’ (Rajala et al. 2018). Performance dialogue utilises the methods of dialogue and concentrates on performance information (e.g., Moynihan 2005). Performance information includes data about resources, workload, outputs, processes, outcomes and efficiency (Hatry 2006), and it systematically is collected, produced and shared (Van Dooren et al. 2015). Boundary-crossing performance dialogues here refer to performance management discussions between representatives of public and private sector organisations operating in the hybrid; this definition of boundary-crossing dialogues was borrowed from the work of Rajala and Laihonon (2019).

Generally, embedding dialogue in performance management can be beneficial in many ways, as it enables learning and development of organisational practices (Laihonon and Mäntylä 2017). However, performance dialogue simultaneously imposes many challenges on organisations, such as a lack of dialogue skills and motivation to conduct dialogues (Rajala et al. 2018). In this research, we were interested in the specific difficulties of hybridity associated with boundary-crossing performance dialogues that occur between representatives of different organisations forming the hybrid organisation. Therefore, we asked the following question: What kinds of challenges are in hybrid organisation’s boundary-crossing performance dialogues? This research question remains unanswered until now. Although Rajala et al. (2018) have studied performance dialogue problems in both hybrid and public organisations, they did not identify challenges of hybridity in boundary-crossing performance dialogues within hybrid organisations. To answer our question, we utilised a case study approach and content analysis. By identifying several challenges that occurred in hybrid organisations’ boundary-crossing performance dialogues, we contribute to conversations addressing performance dialogue and performance management practices in hybrid organisations (e.g., Hodges 2012; Rajala et al. 2018). Our results also offer reasons why it is difficult to attain the type of performance governance that is achieved through collaboration between public, private and third sector actors, according to Halligan et al. (2012). We also argue that our results

provide useful insights into conversations about why performance management does not reach its full potential in hybrid organisations.

The rest of this study is organised as follows. In the next section, we provide more precise conceptual definition of *hybrid* and we define performance management in hybrids. In the third section, research methods and context are explained. In the fourth section, we present an analytical framework for the content analysis, while the fifth section shows how this original framework was iterated and developed during the empirical analysis. The fifth section also provides the empirical analysis. In the final section, conclusions are presented.

2 What are hybrid organisations?

Previous research has demonstrated that *hybridity* is not an easy concept to define (Johanson and Vakkuri 2017). Theoretically, hybrids have been considered an entity that one cannot explicitly explain (Ménard 2004). Almost everything can become hybrid, as its conceptualisations are associated with exploring ‘impure’ forms of social organising (Philoppopoulos-Mihalopoulos 2012). Thus, hybrid forms of organising have been defined in various ways.

Extant literature attaches many characteristics to hybrid organisations. First, hybrid organisations can produce public and private goods (Johanson and Vakkuri 2017). Second, mixed ownership in hybrid organisations is common (Hansmann 1996). *Mixed ownership* indicates that a hybrid organisation has both public and private sector owners (Thynne 2011). Third, different institutional logics, goal congruency and incongruence typically exist in these types of organisations (Johanson and Vakkuri 2017; Kreps and Benoît 2011; Reay and Hinings 2009; Pache and Santos 2013; Ebrahim et al. 2014). Fourth, hybrid organisations obtain funding from a variety of sources (Hodge and Greve 2007). Fifth, differentiated forms of economic and social control are exercised in hybrid organisations (Power 1997; Kelly 2005; Jordana and Levi-Faur 2004). For purposes of this research, an organisation formed by different types of legal organisations is called a *hybrid organisation*, i.e., it comprises public and private or third sector (i.e. voluntary sector) organisations, and these organisations have a common goal that they aim to achieve via collaboration.

3 Performance management in hybrid and non-hybrid organisations

The concept of performance management forms from two major aspects: performance measurement and performance information usage that aims to enhance management and other organisational activities (e.g., Aguinis 2009, p. viii.). In public, private and third sector organisations, performance measures viewed in different performance management models can be divided into the same measurement categories: input, process, workload, output, outcome, productivity and cost-effectiveness measures (e.g., Politt and Bouckaert 2004; Micheli and Kennerley 2005; Thomas et al. 2008; Schläfke et al. 2012; Van Dooren et al. 2015). Performance measures that are used to report a hybrid organisation’s value to the surrounding community also belong to these seven

performance measurement categories (Liu et al. 2014; Vo et al. 2016). The differences between performance measures used in public, private, third sector and hybrid organisations boil down to the question of what types of measures are used, e.g., output and outcome measures. For example, arrested criminals can be an output measure in the public sector, whereas manufactured cars can be an output measure at a car production plant, part of the private sector.

When considering performance evaluations, public, private, third sector and hybrid organisations are trying to evaluate performance aspects that also can be grouped under the seven aforementioned measurement categories (e.g., Lee and Nowell 2015). Here, the only difference is that performance evaluations are non-routine events that are designed in a case-by-case manner, whereas performance measurement is a routine event that occurs more or less the same way at different time points because the objective is to produce comparable information (McDavid and Hawthorn 2006).

The complexity of performance measurement and evaluation differs between the hybrid and non-hybrid organisation types. Performance measurement has been viewed as more complex in the public and third sectors than in the private sector (van Helden and Reichard 2016). The fact that no market-price information on several public sector activities exists makes it difficult to evaluate government interventions and public services' impact and attribution (Johanson and Vakkuri 2017). It also has been argued that hybrids, as organisational networks, are more complex to measure than public and third sector organisations, although problems in evaluating and measuring network performance bear a resemblance to complications viewed in the evaluations of public and third sector organisations (Provan and Milward 2001). The private sector networks are also more complex in terms of performance measurement and evaluation compared to a private sector organisation (Kamminga and Van der Meer-Kooistra 2007). At the organisational level, the constituents of public, private and third sector organisations have conflicting views on what is valuable to the organisation and its actions (Cyert and March 1963; Agranoff and McGuire 2001). Assessing networks, such as hybrid organisations, involves more constituents, and this complicates the evaluation, as more conflicting views on the network's value can exist (Johanson and Vakkuri 2017).

Besides the number of constituents, measuring network performance involves more analysis levels because network performance must be considered, and one cannot focus merely on the organisational and community levels, as is the case with private, public and third sector organisations (Provan and Milward 2001). Using performance information for performance management purposes is more complicated as well because developing the performance of individuals and teams and aligning performance with strategic goals involve more actors.

4 Research setting

This study applies an explorative case study approach (e.g., Yin 2009). We chose this approach because we wanted to focus on one organisation, as this enabled us to use an in-depth exploration and multiple perspectives in describing the

chosen hybrid organisation's complexity and uniqueness (e.g., Simons 2009). Using more cases would have limited our opportunity to understand the richness related to the challenges of boundary-crossing performance dialogue in hybrid organisations. Thus, the case selection served the theory-building objective that we aimed to reach (e.g., Thomas 2011). Indeed, the chosen case enabled us to develop the current theory addressing performance dialogue challenges. Overall, our research can be considered an instrumental case study because it provides a general understanding of the challenges present in boundary-crossing dialogues addressing a hybrid organisation's performance. Thus, we aim to understand how things might be in general, not just how things are in the particular organisation that we examined.

The examined hybrid organisation could be defined as the most extreme, as the scope of services provided was diverse and these services were consumed by heterogeneous clientele. In this research, we focussed on the performance dialogues occurring in the steering and coordination groups, which were the hybrid organisation's two main administrative organs. The steering group prepared budget proposals, solved conflicts, monitored goal achievement and carried into practice the hybrid organisation's procedures. The coordination group made operational decisions and prepared matters for the steering group and service market manager. These administrative organs involved representatives from all nine service units operating in the hybrid organisation. These service units were as follows:

1. A government agency that provides basic social security for residents of Finland.
2. A federation of municipalities that produces healthcare services.
3. A private company that generates healthcare services.
4. A third sector organisation, along with municipal and voluntary workers, that arranges art events.
5. A library (a municipal service unit that offers versatile collections and helps information seekers).
6. A maternity and child health clinic (a municipal service unit that supplies social and health services).
7. A youth service (a municipal service unit that provides low-threshold guidance to youths and young adults ages 16–29).
8. A mental health and substance abuse service clinic (a municipal service unit that helps people over age 18 with mental health and substance abuse problems).
9. A citizen service (a municipal service unit serving citizens by distributing information, selling tickets, etc.).

The hybrid organisation's main goal was to make it easier for citizens to use public sector services, which was the biggest reason why it was formed. The second objective was to create an innovation platform that would improve public service production. The hybrid organisation aimed to fulfil its goals by combining the ideas of accessibility and customer orientation into a new technology and public space design that supported public service production and use. The hybrid organisation attempted to provide better services to citizens by generating

collaborations among different service units. Indeed, active searching for synergies among different units and new ways to collaborate was built into the hybrid's *modus operandi*. For example, many health-related services were part of the hybrid, and they collaborated to create effective service paths for citizens. Creating effective service paths included joint decisions on different service units' tasks in each service path. Other examples of partnerships included collaborations between the library and youth services, joint briefing sessions and other health education events that the library and health service units arranged for citizens. The steering and coordination groups' activities served the hybrid's objectives, and within these groups, the boundary-crossing performance dialogues that we focussed on in this research occurred.

It was typical for the hybrid's personnel to view citizens' needs in a holistic manner, with service paths between different service units generated accordingly. Thus, personnel in the hybrid saw the citizen not only as their own customer, but also as other service units' customer in the hybrid. From citizens' perspective, it was convenient that the often-used services were provided in same building, lowering the threshold to use public services, according to reports (Kenno 2017).

Our data collection method mainly entailed expert interviews (e.g., Meuser and Nagel 2009), which provided more room for a more in-depth explorative approach, as follow-up questions can be asked and clarifying statements may be requested. Interviews were carried out between November 2016 and March 2017. The semi-structured interviews' objective was to determine managers' perceptions of current challenges in boundary-crossing performance dialogues. The shortest interview lasted about 32 min, while the longest one took 70 min. Generally, most interviews lasted about an hour, but all interviews were conducted face to face, audiotaped and transcribed verbatim. Each of the 11 managers was interviewed only once. Only one researcher was sent to conduct each interview. The interviewees were selected through purposive sampling (Patton 2002) and in collaboration with city representatives so that we could locate public managers who participated in the dialogues in which we were interested.

In the hybrid organisation, we interviewed three members from the steering group: the administration and development director, a representative from the library and a representative from a government agency. Then we interviewed five members of the coordination group (the service market manager and representatives from the maternity and child health clinic, library and mental health and substance abuse service clinic). Finally, we interviewed three people from the city's central administration: the director of human resources, the services development director and the project manager. Anonymity for the interviewees was guaranteed and secured; thus, special attention has been paid to preserving anonymity during the sampling and reporting phase. Other researchers listened to the interviews, and feedback was provided to the interviewer if necessary. No biases arising from the interviewer were noticed in the process, and any feedback related mostly to follow-up questions. To triangulate and cross-check the interviews' findings, five documents from the hybrid organisation were analysed (see Table 1).

As an analysis method, inductive content analysis was used. At first, the theoretical framework developed by Rajala et al. (2018) was adopted and used in the

empirical analysis. However, during the analysis process that took place after all interviews were conducted, we iterated and corrected this framework after going through 50 of the interview transcripts. This method is known as *inductive category development*, in which new theoretical categories are created from collected empirical data (Mayring 2000). By utilising inductive category development, it was possible to provide more space for interactions between the researchers and research subjects. This space was very useful because our research topic had been unexplored in previous literature, and we did not know exactly what we were looking for when we entered the case organisation. Conducting the study via inductive content analysis gave us the chance to iterate the research setting. Overall, the whole research process can be summarised in 11 steps:

1. Review literature on performance management and identify research gap.
2. From the literature, adopt theoretical categories for the content analysis.
3. Create questions for interviews based on these categories and select semi-structured interviews as the interview type.
4. Choose interviewees with the case organisation by using purposive sampling and secure interview permissions.
5. Pre-test the interview questions with colleagues and iterate each one of them.
6. Send interview questions to the interviewees beforehand and set interview dates.
7. Conduct all interviews using the chosen questions and record and transcribe each interview.

Table 1 The empirical documents used in this study

Document	Content
A report on an ethnographic study	This study examined the capabilities and challenges of co-production in the hybrid. A total of 100 study subjects participated. The study was based upon interviews and observations, and a private company conducted it. The examined municipality published the report
Presentation of the management model, governing structures, and joint goals of the hybrid	The document describes actors involved in the governance and management of the hybrid and their tasks. The jointly agreed-upon goals of the organisations forming the hybrid are also presented. This document is an internal document of the municipality
The annual budget of the municipality	This document reveals what the hybrid reported to the external stakeholders in official municipal documents. It reported the number of customers and costs per customer
The quarterly report of the municipality	This document reveals what the hybrid reported to the external stakeholders in official municipal documents. It reported the number of customers and costs per customer
A report describing the hybrid as an innovation platform	This 50-page report provided by the municipality describes the hybrid and its functions and problems in detail

8. Analyse 50% of the interview transcripts with theoretical categories using the computer program Atlas.ti.
9. Develop theoretical categories based on the conducted content analysis.
10. Using Atlas.ti, conduct a new content analysis of all interview transcripts and documents with the developed theoretical categories.
11. Report the results.

4.1 Analytical framework: known challenges of performance dialogues

We adopted an analytical framework, presented in Table 2, for the content analysis. This framework comprises seven challenge categories, listed in Column 1. The categories are based on key concepts, described in Column 2. In Column 3, these categories' precise content in our research setting is explained further with example statements excerpted from the interviews. This framework was presented originally by Rajala et al. (2018), and it initially was developed from reasons for non-use of performance information. Extant literature used mental models (McGrath 1999), power (Bass and Riggio 2006; Verhoest et al. 2004), information (Bouckaert 1993; Poister 2003), information systems (Abdel-Maksoud et al. 2015), organisational culture (Moynihan 2005) and organisational structure (Goh 2002) to explain non-use of performance information. Because boundary-crossing performance dialogues are based on performance information use, this theoretical framework seems relevant.

4.2 Challenges of performance dialogues in hybrid organisations

As we went through 50 percent of the interview transcripts after all the interviews were conducted, we noticed that the analytical framework was too general. To describe the challenges of boundary-crossing performance dialogues within the hybrid organisation, we further developed the framework by using our empirical data. In our case, increasing the intensions (e.g., Blackburn 2016, p. 170, for a definition of *intensions*) of the concepts enabled us to describe the challenges of boundary-crossing performance dialogues that were evident in interviewees' statements. We increased intensions by creating subcategories under those categories presented by Rajala et al. (2018). Our subcategories had more characteristics than the categories adopted in the previous research. As a result, we were able to report the challenges of boundary-crossing performance dialogues in the chosen hybrid organisation in a more analytical manner. These subcategories are discussed in detail in the following sub-sections.

4.3 Mental models, motivation and power

First, we developed two subcategories for the category of mental model. The first identified subcategory was named 'language barriers between member organisations'. This challenge indicated the lack of a common language among representatives of member organisations participating in the boundary-crossing performance dialogues (interviewee 3). A representative from a private company expressed this

Table 2 The preliminary framework for the content analysis (taken from Rajala et al. 2018, but see also Van de Walle and Van Dooren 2010; Van Dooren et al. 2015; Rajala 2019)

Challenge category	Definition of the key concept in the category	Example statements of challenges excerpted from the interviews
Mental model	Mental models are psychological representations demonstrating how something works in the real, hypothetical or imaginary world (Johnson-Laird 1998). Explicitly stated values, perceptions, explanations and other knowledge are all mental models that can be examined (Rajala et al. 2018)	Performance dialogue is not viewed as valuable. Performance information is not understood. Participants report lack of a common language. Comprehension of performance dialogues is lacking. Different perceptions of performance cause harmful conflicts. Worldviews are colliding
Motivation	Motivation describes what stimulates a person to action or inaction (Pervin 2003)	Motivation to conduct a performance dialogue is lacking. Motivation to action or inaction that damages a performance dialogue is too high
Power	Power is the ability to take action or inaction (Foucault 1977)	Action or inaction makes it difficult to conduct performance dialogue
Information	Information is propositional knowledge that is stored, sent, received or manipulated in any medium (Dretske 1985)	Information is inadequate and unsatisfying in some way, creating challenges for performance dialogues
Information system	A group of interconnected components that gather, retrieve, process, store and share information to support organisational activities (Laudon and Laudon 2016)	The usability of the information system does not meet the user's standards, and poor usability is complicating the performance dialogue because some or all of the performance information remains unused because of the poor usability
Organisational culture	Organisational (sub)culture comprises habits (i.e., action or inaction), as well as social rules that guide the use of performance information (Van Maanen and Schein 1979)	Performance dialogues fail to gain acceptance within (sub)cultures. The (sub)culture allows people to hinder performance dialogues
Organisational structure	How tasks associated with main goals are divided among branches, sections, divisions, departments, positions and jobs in the organisation's structure (Rajala et al. 2018)	Organisational structures restrict the development of performance dialogues by allowing individuals to do harmful things to these dialogues or by preventing dialogues from occurring

issue in the following manner (document 1): ‘We do not speak the same language, as there are so many actors, but this can be resolved’. When asked about the challenges of boundary-crossing performance dialogues, one interviewee stated that one problem is that:

‘Our [referring to the representatives from the member organisations] language is not necessarily the same....In my opinion, this [lack of a common language] has been a challenge...generating a common language. This is a key factor...’.

Complications in communication arose as a result. In a hybrid organisation, each member organisation has developed its own performance jargon before joining the hybrid, so it was difficult sometimes to determine a ‘common language’ in the organs of the hybrid organisation (interviewee 8). This caused complications in information production and information interpretation in the hybrid organisation.

We labelled the second subcategory ‘conflicting mindsets of the member organisations’. Here, the participants brought their different organisational values and perceptions to the boundary-crossing performance dialogue, and these collided (interviewee 1). The member organisations’ conflicting mindsets led to disagreements about what information should be produced and why (interviewee 11). Statements in the empirical data described the different organisational mindsets in the boundary-crossing performance dialogues this way (interviewee 7, document 1):

‘When we talked about customer service desks [in the hybrid organisation], there were eight different perceptions on what is a customer service desk. Even in such a general matter, there were very different perceptions’.

‘One challenge from the perspective of companies is that expectations do not match up. Actors of different size do not understand that start-ups try to sell their services and try to receive payments from testing these services. Cities want that they are paid for operating as platforms [for testing these services offered by start-ups]. These conversations start well, but do not reach a conclusion, as hopes and goals are different’.

We identified two types of motivational challenges: extrinsic (interviewee 9) and intrinsic (interviewee 10). Extrinsically motivated people are driven by external factors, such as rewards and punishments. People who are extrinsically motivated can conduct actions even though the task might not be in and of itself rewarding. This happens because the reward or punishment associated with these actions acts as a motivator (Frey and Osterloh 2001). The extrinsic motivation challenges were caused by the hybrid organisation’s inability to sanction sufficient disciplinary actions should a member organisation operate in a way that degraded the performance or performance dialogue. In boundary-crossing performance dialogues, we named challenges relating to extrinsic motivation as ‘lack of inter-organisational sanction systems’. Indeed, no incentive systems existed that could be used to force member organisations to do something. As one interviewee stated, ‘I do not know

how we could force to participate those who do not participate even though they are invited [to the meetings]'. Sanctioning also refers to accountability systems that are based on proper monitoring of actions through information systems. However, these monitoring and information systems were lacking in the hybrid organisation, as the following quotations demonstrate (interviewee 2, document 1):

'I do not know exactly what happens to the matters that the steering group has decided. Are they executed or not?...This we do not monitor very much'.

'It depends on the person how the information flows in the service unit, as there is no control for it'.

Compared with extrinsically motivated individuals, intrinsically motivated individuals are driven by interest or enjoyment in the task itself. Thus, intrinsic motivation exists within the individual and does not rely on external pressures, as extrinsic motivation does (Frey and Osterloh 2001). Unfortunately, intrinsic motivation occasionally was lacking in the hybrid organisation because member organisations were not that interested in each other's results, and collaboration was viewed as undesirable or uninspiring. Thus, the subcategory relating to intrinsic motivation was called 'lack of interest in other member organisations'. One interviewee described this challenge as follows (interviewee 9, document 1):

'I'm not interested in what happens in the organisations of [mentions the name of the municipality]. Instead, I'm interested in what happens in [mentions his own organisation by name]'.

'Some employees feel that co-development generates an extra burden that takes time away from their jobs. Some of the employees feel that they have too much to do in their current jobs, so co-development does not raise enthusiasm'.

The problem in the hybrid organisation was that it was not easy to create topics that would interest all the member organisations because these organisations had very different tasks. One manager operating in the hybrid described this by stating that 'all the information is not considered as interesting by everyone'. Another interviewee added that 'if there is a meeting...not producing added value or its purpose is somehow vague, one can perhaps skip it'.

When we examined the empirical data by using the category of power, we conceived a subcategory titled 'powerlessness of the member organisations'. Powerlessness was considered to be a lack of influence over the other member organisations, each of which had an equal amount of power in boundary-crossing performance dialogues. Powerlessness in dialogues points to a situation in which a person or multiple people feel powerless when they attempt to break detrimental deadlocks that inhibit improvements in current performance dialogue practices or their outcomes. For example, the companies thought that one of the

main challenges is ‘how to get people interested in co-development’ (document 1). Interviewees themselves used the word ‘powerless’ to describe their incapacity to change things in performance dialogues. Of course, the fact that a manager is feeling powerless means that someone else must have the power to make the change that the manager desires. Thus, we do not mean that a total absence of power exists when we talk about powerlessness. Rather, powerlessness describes the manager’s feelings. In the hybrid organisation, all the representatives in the boundary-crossing performance dialogues could have felt powerless when dead-lock situations between the representatives occurred. The following quote captures this phenomenon in the hybrid organisation (interviewee 11): ‘Although I am a manager in here, I do not have any power to make [certain] decisions [relating to performance dialogues]’. Another statement from the same interview added:

‘Efficiency is required from all of us as a whole [talking about the goals of the hybrid organisation]....It is an interesting goal [for the hybrid organisation] because I cannot decide on their budgets [refers to the budgets of the other member organisations]. The budget I can decide is so small even in percentages that trimming it won’t do anything’.

4.4 Information systems and information

One subcategory was identified under the category of information systems. We conceptualised this category as ‘member organisations as information system silos’. When information system silos occurred, the participants had different access rights to information systems in different organisations. Each of the member organisations formed its own information system silo in the hybrid organisation. Indeed, the member organisations already had developed their own information systems to serve their organisational goals. The information systems had been developed before any of the organisations became part of the hybrid organisation, and they were mutually distinct (interviewee 11). These separately conducted development processes meant that integration of these information systems was such a complex process that the hybrid organisation decided to abandon the idea altogether, even though integration could offer better information about the costs and effects of the hybrid organisation (interviewee 6). Two people expressed the existence of these information system silos as follows:

‘We do not get the financial figures from our member organisations [operating within the hybrid organisation]. The representatives have been ordered not to give their numbers’ (interviewee 11).

‘the biggest challenges are, in my opinion, in the information flows if information does not transfer quickly enough...but these can be solved as soon as [information] systems become more easily approachable’ (document 1).

Because the member organisations formed information system silos, multiple information asymmetries existed between participants in the hybrid organisation's boundary-crossing performance dialogues (interviewee 3). This affected the performance dialogue in many ways: Time was taken away from the actual use of performance information and expended instead on levelling the pre-existing information asymmetries (interviewee 6).

The subcategory under the concept of information was named 'inability to aggregate the member organisation data'. This subcategory pointed to a lack of aggregated data on member organisations, according to the information user. Producing aggregated information about inputs, processes, workload, outputs, efficiency and outcomes of the organisation network was difficult at the level of the hybrid organisation. The hybrid organisation's main issue was the current data collection procedures. According to one public manager, 'the units collect it differently [she had talked about output information collection in the member organisations]', and 'we have tried to achieve comparability, but we did not achieve this' (interviewee 6).

4.5 Organisational culture and structure

Two subcategories under the main category of organisational culture were found: 'clash of organisational cultures' and 'prevailing culture among the member organisations'. The challenge in the hybrid organisation was that it had to deal with member organisations' clashing organisational cultures. Of course, each member organisation had its own internal cultural clashes, but they also had cultural clashes with the other member organisations when boundary-crossing performance dialogues were implemented. The following quotation captures the clash of organisational cultures arising from the attempt to implement boundary-crossing performance dialogues: 'There are participants from different work cultures, and I do understand that some people do not get it that their presence in the meetings could be useful to others' (interviewee 2). Below are other quotations from the hybrid organisation that summarise this challenge (interviewee 7, document 1):

'The most central [challenge] is that we cannot harmonise master data [in the hybrid]....Information is incompatible, and so is the information production....Different [service] units [within the hybrid]...have different cultural frameworks that create mindsets guiding the information production....Common understanding in information utilisation cannot be reached because the mindsets are different'.

'The public sector is very cautious, and failures are not tolerated....This is a big challenge when one is talking about the culture of experimentation. The culture of the city can be a challenge, and it needs to be developed so that co-development can function'.

The prevailing culture permitted behaviour that was detrimental to performance dialogues and allowed it to become a part of the organisational culture. As a sign of detrimental behaviour permitted by the prevailing culture, one interviewee stated that

‘the weakness of the steering group...is a low attendance rate in meetings, as there are three people in them at best’. As nine service units exist in the hybrid, many representatives of different organisations did not attend. More challenges relating to the prevailing culture were viewed in the documents, and one document contained the following statement: ‘Some employees are persistently staying in their own teams, and they do not try to get to know others, even if one asks nicely’.

Because member organisations were equal in terms of power, the hybrid organisation had to accept the fact that dissent among member organisations on some issues could prevent certain types of performance dialogues. Indeed, the lack of sanction systems meant that detrimental behaviour toward the performance dialogues had to be tolerated. One manager described this situation by stating, ‘I do not know how we can force those to participate who do not participate [in boundary-crossing performance dialogues] now’ (interviewee 8).

We found three subcategories of challenges associated with organisational structures. First, we observed ‘inter-organisational territorialism’. *Territorialism* here is understood as human behaviour characterised by the defence of a particular area, sphere of activity or influence. In short, territorialism is possessiveness. People defended their organisations (territories) by blocking new information systems or tasks suggested to them by someone else operating in the hybrid organisation. In a hybrid organisation, examples of inter-organisational territorialism were found that blocked collaboration, and this inter-organisational territorialism was even understood and accepted by other representatives of member organisations. An interviewee described this by stating that ‘I can understand why [mentions the name of one service unit] does not understand how they could collaborate with [the] child health clinic or library’. Another sign of inter-organisational territorialism was one statement from the documents (document 1):

‘Many state that health services are detached from other services in the service market (i.e., the hybrid organisation). The common aim is that this segregation would not prevent collaboration in the future’.

Within the hybrid organisation, each member organisation had its own territory outlined by organisational boundaries. The member organisations’ representatives also had the option to exercise territorialism because the representatives had an equal amount of power in dialogues. The member organisations even encouraged their agents to be territorial whenever their interests were in jeopardy (interviewee 11). This can be viewed from earlier quotations demonstrating how joint performance information systems could not be implemented because member organisations either did not provide their information or did not want to expend the effort to change their current information systems. Based on the interviews, a hybrid organisation brings inter-organisational territorialism into the organisation (interviewee 1) and makes it an internal matter, rather than an external one.

Second, we labelled a subcategory ‘representative line-up rotation in boundary-crossing performance dialogues’, which refers to a cycle in which representatives of the member organisation change constantly in the dialogues or do not

show up, which, in turn, creates difficulties for the performance dialogue. In the hybrid organisation, one interviewee stated:

‘The representatives [of some member organisations] have changed every time....In fact, I do not even know what is the correct line-up in the steering group’.

‘The staff is changing at a high rate. It is burdensome’.

Rotation was named as a challenge in the hybrid organisation. The hybrid organisation could not force anyone from the member organisations to attend the meetings (interviewee 1). According to an interviewee, ‘the line-up has constantly changed’ in the steering group (interviewee 2). This was viewed as problematic because important voices were being left out of the performance dialogues, and the constant rotation hampered continuity.

Third, we identified a phenomenon that we conceptualised as ‘incompatible tasks of the member organisations’. In this subcategory, performance dialogues in the hybrid organisation were incompatible with member organisation(s)’ tasks. Therefore, the connection between the topic addressed in the performance dialogue and the representative(s)’ tasks was weak, or sometimes even non-existent. Due to these incompatible tasks, representatives sometimes viewed performance dialogues as a waste of time. Consider this statement from interviewee 9:

‘I do not participate in every steering group meeting. I look at the meeting agenda. There are meetings that relate very little to [mentioned his organisation’s name] and to these, I do not necessarily participate....I feel that I do not have anything to give or receive from these meetings’.

The organisations’ incompatible tasks mainly resulted from the municipality’s dominant role in the hybrid organisation and because all matters relevant to the municipality were not relevant to the other member organisations. Many municipal units were involved in the hybrid organisation; therefore, municipal issues were addressed more often in the dialogues. As one manager put it, ‘there is the world of [mentioned the name of his organisation], and there is the world of the city, and the city is not so interested in our figures’ (interviewee 9). One of the documents (document 1) revealed that ‘time pressures associated with service unit tasks made it more difficult for the service units to participate [in] the activities of the hybrid organisation’.

4.6 Summarising empirical findings

The saturation of findings was achieved in this study (see Table 3), as empirical data demonstrates that boundary-crossing performance dialogues between public and private sector actors are confronted with the challenges reported in this study. Different performance management practices between public, private and third sector organisations partly explain why it was difficult to create common performance management systems for the hybrid organisation. For example, differences existed in

Table 3 Saturation of findings

Challenge of boundary-crossing performance dialogue	How many interviewees mentioned this challenge?	How many documents mentioned this challenge?
Language barriers between member organisations	8	2
Conflicting mindsets of the member organisations	10	2
Lack of inter-organisational sanction systems	6	2
Lack of interest in other member organisations	8	2
Powerlessness of the member organisations	8	2
Member organisations as information system silos	9	2
Inability to aggregate the member organisation data	10	2
Clash of organisational cultures	11	2
Prevailing culture among the member organisations	9	2
Inter-organisational territorialism	9	2
Representative line-up rotation in boundary-crossing performance dialogues	8	2
Incompatible tasks of the member organisations	9	2

the organisations' mental models concerning performance. Public, private and third sector organisations also were dissimilar in terms of the information that they collected, and they even applied different methods to collect the same kind of information. Moreover, the information systems in these organisations deviated remarkably from each other. Finally, the organisational cultures and structures supporting their performance were quite different because each organisation served different purposes, and the professions included in these organisations also differed. The hybrid organisation itself did not have any formal power over the member organisations, and proper motivation schemes were not adopted in the hybrid setting. Thus, reconciling differences and finding a common ground for mutual performance management were based on member organisations' willingness. When this willingness was not forthcoming, deadlocks occurred, and the lack of formal power and motivation schemes in the hybrid organisation was viewed as a challenge.

5 Conclusions

Hybrid organisations contain many organisational disparities that can cause conflicts in performance management. This makes performance dialogues especially relevant to these types of organisations because such dialogues often are used to transform conflicts into cooperation. Considering the hybrid organisation's characteristics requires active inter-organisational dialogue when performance management systems are being designed and used. For example, using the performance dialogues to understand goal incongruence between the organisations comprising the hybrid organisation can be very beneficial to the success of

performance management. However, creating a functional performance dialogue also can be challenging in many ways.

In this study, we sought to identify the challenges of boundary-crossing performance dialogue in a hybrid organisation. This issue is relevant not merely due to performance dialogues' important role in all organisations (e.g., Agostino and Arnaboldi 2015, 2018), but also particularly because of our limited understanding of performance dialogues' impacts on different forms of hybrid organisations. Accordingly, this paper has explored the problems of boundary-crossing performance dialogues in the context of particular hybrid organisation. As a result, we found many challenges and described them in a detailed way in our analysis. With this analysis, the study revealed relationships between inter-organisational factors and obstacles with boundary-crossing performance dialogues. Conflicting mind-sets between member organisations, member organisations as information system silos, inter-organisational territorialism, and clashes of organisational cultures are all examples of inter-organisational factors that can cause issues in performance dialogues. Some of the problems which can arise include inhibition of innovation, and limitations placed on the learning and knowledge-sharing potential which is normally embedded in boundary-crossing performance dialogues. In general, the inter-organisational factors described in this study were harmful due to the fact that they were detrimental to an environment open to respectful communication of views on performance. As respectful and open communication is a bedrock of every effective dialogue (Rajala et al. 2018), issues emerge when these practices are lacking. Thus, the relationships identified by this study create new insights into the systems and practices of hybrid performance management.

It seems intuitive to claim that the challenges of boundary-crossing performance dialogues in hybrid organisations have something to do with the hybrids' general characteristics (e.g., Johanson and Vakkuri 2017). Because this study did not focus on describing how the hybrid organisation's characteristics can create the challenges that we identified, a need exists to examine this in future research. As an example, future studies could address how different forms of social and economic control create, e.g., clashes between organisational cultures, or how different funding methods shape the hybrid organisation's performance management system. Furthermore, it can be concluded that additional theoretical work addressing the conceptual relationships between the challenges of boundary-crossing performance dialogues and the general characteristics of hybrids is needed.

Because the empirical examinations showed what caused the challenges of boundary-crossing performance dialogues, they also presented a model describing how these challenges occurred. However, further testing is needed before this model can be accepted as a general theory on boundary-crossing performance dialogues. As a methodological contribution, we created a conceptual framework that can be used to study the challenges of boundary-crossing performance dialogues occurring between public and private sector actors in hybrid organisations. In this research, we tested two conceptual frameworks in the analysis section, which revealed that more analytical frameworks would be useful in understanding the characteristics of the challenges associated with boundary-crossing performance dialogues. This

conceptual framework also proposes new concepts for the performance dialogue theories (e.g., Moynihan 2005; Rajala et al. 2018; Rajala and Laihonon 2019).

The results of this study have several managerial implications. First, because boundary-crossing performance dialogues are platforms that reflect the state of the organisation, they suffer from the same general issues which challenge the hybrid form of organising. These problems are related to motivational challenges and incompatible information systems, as well as the lack of common mindsets and language between people working in different areas of the hybrid organisation. Inadequate organisation culture and structures are also an issue. Second, as the complications of forming hybrids engender costs, practitioners should consider carefully whether the benefits of establishing such an organisation outweigh the costs arising from it. Here, the central question is whether it makes more sense for these organisations to work independently or as a hybrid. Finally, practitioners should also keep in mind that boundary-crossing performance dialogues are only one coordination instrument amongst others which can be utilised to shape hybrids. However, boundary-crossing performance dialogues can also be a vital coordination tool when regularly used to solve difficult and complicated disagreements.

In this study, we used a case study approach, and for this reason, the results cannot be generalised (e.g., Abdel-Kader et al. 2009). Other limitations of the study include the limited number of interviewees and the fact that we did not interview representatives from the private sector. The views of the private sector actor were examined from secondary data (i.e., documents produced by the hybrid organisation) as these contained direct quotations from company representatives. When one is conducting qualitative content analysis, the observer-expectancy effect always runs the risk of biasing results. To tackle doubts concerning observer-expectancy effect, we have provided quotations from the challenges of boundary-crossing performance dialogues. From these quotations, the reader should be able to discern whether we described the challenge arising from the interviews correctly. The analytical framework created in this study was built from the descriptions we used to depict the challenges of boundary-crossing performance dialogues that interviewees described to us. Because saturation of interview findings (see Table 3) confirmed our results about the challenges of boundary-crossing performance dialogues, and the analysed documents provided additional support for the interview findings, we are confident that specific challenges of boundary-crossing performance dialogues existed in the examined hybrid organisation. However, we do acknowledge that the way that we chose to describe the challenges seen in the interview is our construction that attempts to convey interviewees' views. As the created analytical framework is our construction, it means that constructivism describes our research approach.

Despite these limitations, we think that our research has provided fertile ground for future research to address and more fully test our theoretical ideas about the challenges of performance dialogues in hybrid settings. Thus, we call for more research on the ideas proposed here so that more knowledge on dialogues as performance management practices in hybrid and public organisations can be gathered. This new knowledge is valuable in attempts to move to the type of inter-organisational performance governance described by Bouckaert and Halligan (2008).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Interview questions

Thematic interview questions categorised according to the theoretical framework.

General view on performance dialogues

1. How would you describe the dialogues about performance?

Performance information and performance information system

2. Are there problems related to performance information and the systems providing it?

Organisational structure and motivation

3. Who participates when organisational goals and performance indicators are determined, and actions are being decided based on the performance information?
4. Are all necessary participants present? Why/Why not?

Mental models

5. Is the performance information provided useful to you?
6. Are there different interpretations/views about the performance information available?

Power and organisational culture

7. Do people listen and respect different interpretations/views about the performance information?
8. How do you resolve conflicting views and develop conversational culture in the organisation?

References

Abdel-Kader, M. G., Lin, E. Y.-C., & Connect, Palgrave. (2009). *Performance measurement of new product development teams: A case of the high-tech sector*. Basingstoke: Palgrave Macmillan.

- Abdel-Maksoud, A., Elbanna, S., Mahama, H., & Pollanen, R. (2015). The use of performance information in strategic decision making in public organisations. *International Journal of Public Sector Management*, 28(7), 528–549.
- Agostino, D., & Arnaboldi, M. (2015). The New public management in hybrid settings: New challenges for performance measures. *International Review of Public Administration*, 20(4), 353–369.
- Agostino, D., & Arnaboldi, M. (2018). Performance Measurement Systems in Public Service Networks. The What, Who and How of Control. *Financial Accountability & Management* p 1-14.
- Agranoff, R., & McGuire, M. (2001). American federalism and the search for models of management. *Public Administration Review*, 61(6), 671–681.
- Aguinis, H. (2009). *Performance management*. Upper Saddle River: Pearson Prentice Hall.
- Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership*. Mahwah, NJ: Erlbaum.
- Blackburn, S. (2016). *The Oxford dictionary of philosophy*. Oxford: Oxford University Press.
- Bouckaert, G. (1993). Measurement and meaningful management. *Public Productivity & Management Review*, 17, 31–43.
- Bouckaert, G., & Halligan, J. (2008). *Managing performance. international comparisons*. Oxon, UK: Routledge.
- Christensen, T., & Lægreid, P. (2011). Complexity and hybrid public administration—theoretical and empirical challenges. *Public Organisations Review*, 11(4), 407–423.
- Cyert, R. M., & March, J. G. (1963). *A behavioural theory of the firm*. Englewood Cliffs: Prentice Hall.
- Dretske, Fred. (1985). Precise of knowledge and the flow of information. In Hilary Kornblith (Ed.), *Naturalizing epistemology* (pp. 169–187). Cambridge, MA: MIT Press.
- Ebrahim, A., Batilana, J., & Mair, J. (2014). The governance of social enterprises: Mission drift and accountability challenges in hybrid organisations. *Research in Organisational Behavior*, 34, 81–100.
- Foucault, M. (1977). *Power/Knowledge*, edited by Colin Gordon. New York: Pantheon Books.
- Frey, B. S., & Osterloh, M. (Eds.). (2001). *Successful management by motivation: Balancing intrinsic and extrinsic incentives*. Springer Science & Business Media.
- Goh, S. C. (2002). Managing effective knowledge transfer: An integrative framework and some practice implications. *Journal of Knowledge Management*, 6(1), 23–30.
- Halligan, J., Sarrico, C. S., & Lee Rhodes, M. (2012). On the road to performance governance in the public domain? *International Journal of Productivity and Performance Management*, 61(3), 224–234.
- Hansmann, H. (1996). *The ownership of enterprise*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Hatry, P. (2006). *Performance measurement: Getting results*. Washington DC: The Urban Institute.
- Hodge, C., & Greve, C. (2007). Public private partnerships: An international performance review. *Public Administration Review*, 67(3), 545–558.
- Hodges, R. (2012). Joined-up government and the challenges to accounting and accountability researchers. *Financial Accountability & Management*, 28(1), 26–51.
- Johanson, J., & Vakkuri, J. 2017. *Governing Hybrid Organisations. Exploring Diversity of Institutional Life*. Routledge. ISBN 9781138655829.
- Johnson-Laird, P. (1998). Imagery, visualization, and thinking. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 441–467). New York: Academic Press.
- Jordana, J., & Levi-Faur, D. (2004). *The politics of regulation: Institutions and regulatory reforms for the age of governance*. Cheltenham, UK: E. Elgar.
- Kamminga, P. E., & Van der Meer-Kooistra, J. (2007). Management control patterns in joint venture relationships: A model and an exploratory study. *Accounting, Organizations and Society*, 32(1–2), 131–154.
- Kelly, J. M. (2005). The Dilemma of the unsatisfied customer in a market model of public administration. *Public Administration Review*, 65(1), 76–84.
- Kenno. (2017). *Etnographic research on service market*. Helsinki: Kenno Oy.
- Kreps, T. A., & Benoît, M. (2011). ‘Doing Well by Doing Good’? Ambivalent moral framing in organisations. *Research in Organisational Behavior*, 31, 99–123.
- Laihonen, H., & Mäntylä, S. (2017). Principles of performance dialogue in public administration. *International Journal of Public Sector Management*, 30(5), 414–428.
- Laudon, K. C., & Laudon, J. P. (2016). *Management information system*. Pearson Education India.
- Lee, C., & Nowell, B. (2015). A framework for assessing the performance of nonprofit organizations. *American Journal of Evaluation*, 36(3), 299–319.
- Liu, J., Love, P., Smith, J., Regan, M., & Sutrisna, M. (2014). Public-private partnerships: a review of theory and practice of performance measurement. *International Journal of Productivity and Performance Management*, 63(4), 499–512.

- Mayring, P. (2000). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (7th ed.). Weinheim: Deutscher Studien Verlag.
- McDavid, J. C., & Hawthorn, L. R. L. (2006). *Program evaluation & performance measurement: An introduction to practice*. Thousand Oaks: SAGE Publications.
- McGrath, R. G. (1999). Falling forward: real options reasoning and entrepreneurial failure. *Academy of Management Review*, 24(1), 13–30.
- McGuire, M. (2002). Collaborative public management: Assessing what we know and how we know it. *Public Administration Review*, 66(s1), 33–43.
- Ménard, C. (2004). The economics of hybrid organisations. *Journal of Institutional and Theoretical Economics*, 160, 345–376.
- Meuser, M., & Nagel, U. (2009). The expert interview and changes in knowledge production. In A. Bogner, B. Littig, & W. Men (Eds.), *Interviewing experts* (pp. 17–42). London: Palgrave Macmillan.
- Micheli, P., & Kennerley, M. (2005). Performance measurement frameworks in public and non-profit sectors. *Production Planning & Control*, 16(2), 125–134.
- Moynihan, D. (2005). Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), 203–216.
- Pache, A.-C., & Santos, F. (2013). Inside the Hybrid Organisation: Selective coupling as a response to competing institutional logics. *Academy of Management Review*, 38(4), 972–1001.
- Patton, M. (2002). *Qualitative research & evaluation methods*. California: Sage.
- Pervin, L. (2003). *The science of personality* (2nd ed.). New York: Oxford University Press.
- Philoppopoulos-Mihalopoulos, A. (2012). Suspension of suspension: Notes on the hybrid. In N. Åkerstrøm Andersen & I.-J. Sand (Eds.), *Hybrid forms of governance: Self-suspension of power* (pp. 12–29). Chippenham: Palgrave Macmillan Publisher.
- Poister, T. H. (2003). *Measuring performance in public and nonprofit organizations*. San Francisco, CA: Wiley.
- Pollitt, C., & Bouckaert, G. (2004). *Public management reform: A comparative analysis*. Oxford: Oxford University Press.
- Power, M. (1997). *The audit society. Rituals of Verification*. Oxford: Oxford University Press.
- Provan, K. G., & Milward, H. B. (2001). Do networks really work? A framework for evaluating public-sector organizational networks. *Public administration review*, 61(4), 414–423.
- Rajala, T. (2019). Mind the information expectation gap. *Journal of the Knowledge Economy*, 10(1), 104–125.
- Rajala, T., & Laihonen, H. (2019). Managerial choices in orchestrating dialogic performance management. *Baltic Journal of Management*, 14(1), 141–157.
- Rajala, T., Laihonen, H., & Haapala, P. (2018). Why is dialogue on performance challenging in the public sector? *Measuring Business Excellence*, 22(2), 117–129.
- Reay, T., & Hinings, C. R. (2009). Managing the rivalry of competing institutional logics. *Organisation Studies*, 30(6), 629–652.
- Schläfke, M., Silvi, R., & Möller, K. (2012). A framework for business analytics in performance management. *International Journal of Productivity and Performance Management*, 62(1), 110–122.
- Simons, H. (2009). *Case study research in practice*. London: SAGE.
- Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative Inquiry*, 17(6), 511–521.
- Thomas, A., Barton, R., & Chuke-Okafor, C. (2008). Applying lean six sigma in a small engineering company—a model for change. *Journal of Manufacturing Technology Management*, 20(1), 113–129.
- Thynne, I. (2011). Ownership as an instrument of policy and understanding in the public sphere: Trends and research Agenda. *Policy Studies*, 32(3), 183–197.
- Van de Walle, S., & Van Dooren, W. (2010). How is information used to improve performance in the public sector? Exploring the dynamics of performance information. In K. Walshe, G. Harvey, & P. Jas (Eds.), *Connecting knowledge and performance in public services* (pp. 33–54). Cambridge: Cambridge University Press.
- Van Dooren, W., Bouckaert, G., & Halligan, J. (2015). *Performance management in the public sector*. Routledge: Taylor & Francis Group.
- van Helden, J., & Reichard, C. (2016). Commonalities and differences in public and private sector performance management practices; a literature review. In M. J. Epstein, F. Verbeeten, & S. K. Widener (Eds.), *Performance measurement and management control: contemporary issues* (Vol. 31, pp. 309–352). Emerald: Egmore.

- Van Maanen, J., & Schein, E. (1979). *Towards a Theory of Organizational Socialization*. Cambridge: MIT Press.
- Verhoest, K., Peters, B. G., Bouckaert, G., & Verschuere, B. (2004). The study of organizational autonomy: a conceptual review. *Public Administration and Development*, 24, 101–118.
- Vo, A. T., Christie, C. A., & Rohanna, K. (2016). Understanding evaluation practice within the context of social investment. *Evaluation*, 22(4), 470–488.
- Yin, R. K. (2009). *Case study research—design and methods* (4th ed.). Thousand Oaks, CA: Sage.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Tommi Rajala MA, has graduated at the University of Tampere. He has specialized in local public economics. Currently Rajala works as a project researcher in the hybrid organization project and as a lecturer in public financial management at the Faculty of Management and Business. As a PhD student, he is researching performance public sector outcome measurement. Along with these tasks, Rajala is part of a research project focusing on performance dialogues. Previously he has worked in municipalities and Statistics Finland with government productivity statistics.

Harri Laihonon is a research director at Faculty of Management and Business, Tampere University, Tampere, Finland. Harri's research interests lie in the fields of knowledge management and performance management. Within these fields, his current research projects investigate how performance information can be more efficiently used to support decision-making and organizational learning through performance dialogue. At moments, Harri is operating mostly in the public sector but the phenomena are general. Service systems and hybrid organizations provide extremely interesting contexts for the research in different sectors.

Jarmo Vakkuri is Professor at the Faculty of Management and Business, Tampere University, and the director of the research group on Public Financial Management. His research concentrates for instance on performance measurement in public administration, theories of decision-making, hybrid organizations and governance, and urban policy-making. Vakkuri is the director of the HYPER project (Performance measurement for hybrid governance), and one of the directors of the "Beyond MALPE coordination: integrative envisioning" (BEMINE), both funded by the Academy of Finland.

Rajala, T. (2020) Blame avoidance strategies in governmental performance measurement. *Financial Accountability & Management*. 36: 278–299. <https://doi.org/10.1111/faam.12225>

Blame avoidance strategies in governmental performance measurement

Tomi Rajala, University teacher, Ph.D. student

University of Tampere, School of Management, FI-33014 University of Tampere, Finland

Email: tomi.rajala@uta.fi

Abstract

Performance measurement and blame avoidance are significant forces that shape the development of the public sector. Unfortunately, extant literature has not paid much attention to blame avoidance in performance measurement. Thus, this article aims to show how blame avoidance strategies can be embedded in performance measurement. This case study's results provide theoretical ideas and empirical examples that demonstrate how a particular performance measure—central government productivity—enabled blame avoidance. These results will help practitioners and academics view blame avoidance aspects in performance measurement.

Keywords: blame avoidance, performance management, performance measurement

Introduction

Performance measurement and blame avoidance are essential aspects of public sector activities. “Performance measurement,” that is, the act of measuring performance with indicators, often is presented as a tool that leads to improved public sector performance (Van Dooren & Van de Walle, 2011) and transparent government (Johnsen, 2005). Extant blame avoidance literature presents blame avoidance as a central feature of public management and administration (Hood, 2011). In this paper, “blame avoidance” refers to the act of minimizing the expected blame that one must face when something unwanted happens in the public sector domain (Hood, 2014). Avoiding blame when service production operations in the public sector fail is a typical example of blame avoidance (Rajala, Laihonon, & Vakkuri, 2018).

Blaming, as a phenomenon, needs at least two actors: a blamer and a scapegoat. The scapegoat needs to do something blameworthy, and the blamer must be able to observe the scapegoat's blameworthy actions. Blame cannot happen without information that describes someone doing something blameworthy. Performance measurement provides information about the public sector's blameworthy actions. Indeed, performance measurement is related to accountability (Klout, 1999) and to the question of who should be blamed when performance falls short (Bovens, 2010). Public sector actors operate in an environment characterized by a negativity bias (Charbonneau & Bellavance, 2012), and performance measurement, as an accountability mechanism, may be used to blame public managers (Flynn, 1986) or politicians (James, 2010). Because blame is a risk associated with performance measurement (Wholey & Hatry, 1992), risk management can be adopted (e.g., Hood, 2002) and based on blame avoidance strategies (Hood, 2011).

Until now, scientific research on performance measurement has focused on performance measures' design (Wisniewski & Stewart, 2004), as well as implementation (Collier, 2006) and use of such measures (Ho & Chan, 2002). Measurement systems' content also has been examined (Van Peurse, Pratt, & Lawrence, 1995). However, extant literature has not considered what type of measurement system content is an indication of the use of blame avoidance strategies.

The blame avoidance literature does not provide an answer to this question either, focusing mainly on blame avoidance strategies used in political and administrative functions to mitigate, delegate, or hide blame arising from public sector failures (Hood, 2011). These strategies explain how public officials and politicians try to avoid blame (Hinterleitner, 2017). However, none of these strategies requires performance measures, which partly explains why blame avoidance strategies used in measurement solutions are an underexamined topic (e.g., Hood, 2011; Weaver, 1986). To address this research gap, this article studies how performance management systems' design may facilitate blame avoidance strategies and how governments use these strategies in performance measurement systems.

As a result, this article demonstrates that performance measurement solutions can actualize blame avoidance strategies. This contributes to discourse on performance measurement and blame avoidance, as neither research stream has demonstrated how blame avoidance strategies can be incorporated into performance measures. A case study applying a hypothetico-deductive method is utilized to achieve these results.

For practitioners and academics, the contributions offer new ways to understand performance measurement as a system that utilizes blame avoidance strategies in its functions. This understanding is important because blame avoidance inhibits the ability to see reasons for failures; thus, it makes learning more difficult in the public sector. The theoretical ideas presented here also offer fresh starting points for future studies to examine blame avoidance in other types of performance measures, not just in productivity measures. Moreover, this study left many blame avoidance strategies uncovered; therefore, new research is needed to address these.

In the next section, the research gap in the previous literature is presented in more detail and the created theoretical hypotheses on blame avoidance strategies in performance measurement are demonstrated. In the third section, the research method is described. The fourth section provides an empirical analysis, and the article ends with a discussion and conclusions.

Theoretical background

Using blame avoidance strategies in performance measurement refers to what extant literature calls "blame avoidance behavior," which several studies have examined. Blame avoidance behavior, as a research subject, was examined first by Weaver (1986), who said blame avoidance behavior existed in the political domain because politicians used blame avoidance strategies for two reasons: Politicians are loss averse, and their constituents had a negativity bias. According to Weaver (1986), loss averse

politicians aim to avoid blame more than they want to claim credit. Therefore, politicians are willing to give up opportunities to claim credit. As one can see, Weaver's focus focused on political decision-making, not on performance measurement systems enabling blame avoidance.

Since Weaver's (1986) seminal article, blame avoidance behavior has been studied in two fields: comparative welfare state research and public policy and administration (Hinterleitner, 2017). In comparative welfare state research, Pierson (1994, 1996) studied blame avoidance behavior and how it was used to pursue out of favor reforms. In Pierson's (1994) thinking, politicians have only two objectives: to advance their political agendas and ensure their re-election. If a politician's political agenda shifts from expansion to cutback politics, cutbacks can cause losses for politically important interest groups. Therefore, advancing a political agenda that calls for cutbacks can imperil re-election prospects. According to Pierson (1994), blame avoidance behavior is used to reconcile tension between retrenchment and electoral retribution to ensure re-election despite cutbacks.

Since Pierson's (1994, 1996) work, other scholars' studies have examined ideas about blame avoidance behavior suggested by Pierson and have confirmed their relevance (Lindbom, 2007; Mortensen, 2012; Vis, 2009; Vis & van Kersbergen, 2007; Wagschal & Wenzelburger, 2008; Wenzelburger, 2011; Zohlnhöfer, 2007). What is essential from the present study's perspective is that comparative welfare state research has not examined how performance measures can actualize blame avoidance strategies because they focus on blame avoidance behavior associated with cutback policies.

In addition to studies on welfare state retrenchment, blame avoidance behavior has been examined in public policy and administration research (Hinterleitner, 2017), which views blame avoidance behavior as a more widespread behavioral phenomenon in the political sphere than comparative welfare state studies, which view blame avoidance behavior as a tool used to achieve retrenchment while ensuring re-election (e.g., Hood, 2011; Pierson, 1994). Politicians are forced to use blame avoidance strategies not only when they exercise cutback policies, but also in other situations in which their political agendas are jeopardized. In fact, public policy and administration scholars assume that public accountability "comes in many guises" (Bovens, 2007, p. 454). As Hinterleitner (2017, p. 248) states, "public actors are generally held accountable by different actors in different forums, for different things and in different situations." Hinterleitner's (2017) statements indicate that blame avoidance behavior is used in many other situations in which political agendas are threatened, not only in the context of cutback policies.

Depending on the situation, extant literature has identified how public actors can use anticipatory and reactive forms of blame avoidance (Sulitzeanu-Kenan, 2006). In the anticipatory form, proactive public actors try to keep problems off the agenda before stakeholders detect them and react by voting for change in government. The anticipatory form has been studied to some extent (Fiorina, 1982, 1986; Hood, 2007; Horn, 1995). The reactive blame avoidance occurs after a problematic issue has surfaced

on the public agenda, and the blame arising from this issue must be addressed (Hinterleitner, 2017). Many studies have addressed reactive blame avoidance forms (Boin, Hart, McConnell, & Preston, 2010; Brändström & Kuipers, 2003; Hinterleitner & Sager, 2015). Unfortunately, the research on public policy and administration does not address how blame avoidance manifests itself in performance measurement systems (e.g., Hinterleitner, 2017).

Literature on accountability has considered how performance measures enhance accountability (Tilbury, 2006). These studies are useful because the “sanction-imposing” sense of accountability relates to the type of blame avoidance that refers to “outcome rather than activity” (Hood, 2014). Here, “accountability” means sanctions against those accountable, and this accountability remains unachieved when blame avoidance occurs. Unfortunately, the research gap in accountability literature is that no studies are examining blame avoidance strategies used in performance measurement systems. For this reason, current knowledge on the relationship between accountability and blame avoidance behavior is lacking.

Literature on performance management and measurement has addressed blame avoidance behavior only to a limited extent, but these studies have not examined how performance measures can actualize specific blame avoidance strategies (e.g., George, Desmidt, Nielsen, & Baekgaard, 2017; Nielsen & Baekgaard, 2015). Rajala et al. (2018) noted how outcome measures may attract more bad publicity than output measures; thus, outcome measures sometimes may be avoided in performance measurement systems. Many outcomes also require a longer time frame than, for example, outputs, which might appeal to those who need short-term success to avoid accusations of inefficiency. When compared to outcomes, outputs usually offer also more control over results, as they comprise goods and services manufactured in the public sector, whereas outcomes describe the effects that outputs directly and indirectly cause in society (Rajala et al., 2018). Thus, having more control over outputs can make output measures a less risky option.

These arguments that consider differences between output and outcome measures point out that certain elements in performance measures can encourage one to adopt blame avoidance strategies in performance measurement systems. These elements can be unseen or well-known. As performance measures can have many unintended effects (Smith, 1995), it is possible that some elements that encourage blame avoidance behavior are unseen in the design phase. However, if designers of performance measures know the blaming potential embedded within performance measures, they can build blame avoidance strategies into performance measures, assuming that the context enables this type of designing and that the designers have the desire and know how to do so.

The need to adopt blame avoidance strategies also relates to the purpose of measurement and to the surrounding operational environment. For example, if a measurement system’s purpose is to provide accountability (e.g., Behn, 2003), the need to adopt blame avoidance strategies may be more prevalent

compared with performance measurement systems that seek to enhance learning. The measurement context also plays a role, as Johnsen (2012), Nicholson-Crotty, Theobald, and Nicholson-Crotty (2006), and Nicholson-Crotty, Nicholson-Crotty, and Fernandez (2017) have implicated that using blame avoidance may be more likely in some administrative sectors and policy areas than others if the public sector is risk averse (see also George et al., 2017).

Considering the different literature streams that have addressed blame avoidance, the present study can bridge the aforementioned apparent research gap (see, e.g., Hinterleitner, 2017). New understanding on blame avoidance behavior in performance measurement systems can be gained by using the descriptions of blame avoidance behavior found in extant literature and testing whether this type of behavior also occurs in performance measurement systems.

Blame avoidance in performance measurement: theoretical hypothesis

To create theoretical hypotheses to be tested, a citation analysis was used to identify highly cited publications and authors presenting blame avoidance strategies in the public policy and administration domain, which was chosen because research in this field has stated that blame avoidance strategies can be used on many different kinds of occasions, whereas comparative welfare state studies have examined blame avoidance behavior only as an instrument to pursue retrenchment (see, e.g., Hood, 2011; Pierson, 1994).

As a process, “citation analysis” entails counting the number of times a research article is cited by other studies to measure an author or publication’s impact on the field of study (Miller, 2009). As a search term, “blame avoidance” was used. Because no single citation analysis tool collects all publications and their cited references, in this study, many resources—including Web of Science, Scopus, and Google Scholar—were used. As a result, one author, Christopher Hood, had the most citations in the public policy and administration field. Hood (2011) addresses behavioral strategies used to prevent or mitigate blame that public sector activities generate. In the field of study that addresses blame avoidance, Kent Weaver had more citations, but his work focuses on the political domain, so Hood was chosen.

Because known blame avoidance strategies do not rely on the use of performance measures (see, e.g., Hinterleitner, 2017; Weaver, 1986), how these blame avoidance strategies could be applied in a performance measurement context had to be inferred. Four known strategies were adopted for the present study: herding, secrecy, reorganizing, and abstinence (e.g., Hood, 2011). As a theoretical contribution, this article proposes how these four blame avoidance strategies can be used in the context of performance measurement.

The first blame avoidance strategy addressed here is “herding,” in which citizens cannot tell who is responsible for poor results because so many hands are involved in the public service process (Hood,

2011). In the literature, this type of situation often is referred to as the problem of many hands (Thompson, 2014). As a blame avoidance strategy, herding uses collective decision-making and group work as a means to avoid blame. In herding, it is essential to stay with the herd so that no one will be held accountable individually for failures (Hood, 2011).

The problem of many hands can be incorporated into performance results with maneuvers conducted in performance measurement design. For example, performance information can be aggregated to a level in which it presents only the performance of a particular group of people or agencies operating within the public sector, for example, output measures can be defined in such a way that they present only group actions to the public. This way, the burden from poor performance results will not fall on the shoulders of any individual or agency. When a performance measurement system fulfills the herding strategy's requirements, citizens can report only that group performance is lacking in certain areas. However, citizens cannot identify the individual(s) or agency behind the failure.

Performance measurement systems can execute herding as a blame avoidance strategy when the blamer cannot disaggregate group performance results to a level at which it is possible to see root causes of aggregated performance results. "Root causes" in this paper refer to individual agency's contributions to group results. The inability to disaggregate performance results can result from limited access rights to performance data or a lack of understanding of how the aggregation of performance results can be disaggregated. To conclude, what extant blame avoidance literature calls herding is the same thing as aggregation in performance measurement literature whenever blamers cannot disaggregate group results to see how individual actors or agencies contributed to group results. In this study, individual actors are understood as agencies that comprise the government. Besides aggregating information, herding can be adopted by using measurement scales that make different units look similar in terms of performance.

In the design phase, before any blame arises, herding created by aggregating data or fixing the measurement scales is an anticipatory form of blame avoidance in which proactive public actors try to keep problems off the agenda before stakeholders detect them and react by voting for change in government. Herding, as a reactive form of blame avoidance, would point to two different types of cases when performance measurement systems are being reorganized. In the first case, measurement scales are changed after the blame has risen so that every agency starts to look similar in terms of results. In the second case, calculation procedures involving data aggregation are adopted after accusations have been provided. In the first case, changing the scales creates convergence. This is different from what Smith (1995) refers to as convergence, in which measured units want to be average and similar, instead of excellent or poor. Smith's convergence is herding but different kind of herding than changing scales. Changing scales to create convergence is performance measurement technique and blame-avoidance behavior seen in measurement design and redesign. Agencies who want to be average and therefore

change their behavior based upon measured performance are exercising blame avoidance behavior related the measurement use, not measurement design or redesign. To conclude, the following hypothesis is developed from the theoretical ideas presented concerning herding:

Hypothesis 1: The government uses herding by aggregating data, which avoids blaming individual agencies for poor performance.

Generally, “secrecy” refers to intentional noninformation (Bok, 1983), and restricting information is one way to avoid blame (Hood, 2011). Designing performance measures involves choosing how phenomena are conceptualized and how activities are recorded, calculated, analyzed, and reported. Designing a performance measurement system in a particular way opens up opportunities for keeping secrets. Reasons behind results can be hidden in lengthy calculations and evaluations if citizens do not have access to all these calculations and evaluations. Moreover, blame avoidance can set the tone for how things are calculated or evaluated, and this method can be kept secret.

It is possible to design the amount of secrecy related to a particular performance indicator’s results if citizens do not demand full transparency in performance measurement, or they do not understand how secrecy can be built into performance indicators. Poor results for problem units can be buried in indicators that show positive development to the outside world, as long as the number of well-performing units outweighs the number of problem units when performance numbers of different units are being summed up. Unless citizens demand specific performance information about problem units, both aggregating performance information and leaving out problematic information can keep poor-performing units out of the public eye.

Secrecy, as a blame avoidance strategy, can take anticipatory and reactive forms. If secrecy is adopted before any blame arises, it is the anticipatory form of secrecy. However, if secrecy is added to a performance measurement system when performance measures are reorganized as a reaction to poor results, it is the reactive form. From the above ideas relating to secrecy, the following hypothesis is created:

Hypothesis 2: The government uses secrecy by limiting publication of or access to data, which avoids blaming individual agencies for poor performance.

Moving the blame target is the fundamental idea behind the defensive reorganization that Hood (2011) addresses. According to Hood, those who seek to level blame at a public sector organization:

“can always be answered by the counterclaim that the critics have not understood the complexities of the system or that they are hopelessly out-of-date and whatever faults they

are highlighting belong to an older system whose faults have now been overcome in the latest round of reorganization.”

Coe (2003) has captured this reorganization in a performance measurement context in his research examining performance report cards in the United States. According to Coe (2003), readers of performance report cards could not compare report card grades over time because indicators had changed, with some organizations even changing them annually. Lombardi, Craig, Capaldi, and Gater (2000) reported similar changes in performance measures in the United States, and Park and Cho (2014) reported changes in performance measures in Korea. Lonti and Gregory (2007) tracked changes in performance measures in New Zealand and argued that many measures had changed during the time interval examined. Some authors even have suggested changing performance measures periodically, as this prevents gaming the system in contracting out situations (Heinrich & Choi, 2007).

To summarize the past findings, it seems that a performance measurement’s designer can reorganize, that is, change indicators used to report public sector performance. The unit of analysis can be changed, as well as the overall phenomenon being captured with the indicator (e.g., Lonti & Gregory, 2007). These changes create a moving target for critics.

Out-of-date performance information can be countered by stating that the new measurement system better describes the performance in which we are interested. Scapegoats and the performance measurement system’s designers can use these counterarguments. If people criticize the performance indicators, the performance measurement system’s designer can state that the critics are looking at old indicators and should be looking at new indicators that will provide new data in the near future. Similarly, constant changes in indicators offer explanations to public managers responsible for the performance being measured. For example, a public manager whose performance is measured can state that the indicators were changed because they did not adequately reflect actual performance and that critics should not use these indicators if they want to avoid misperceptions.

Reorganization of performance measures is an anticipatory form of blame avoidance if performance measures are changed before any blame actually occurs. Here, one anticipates that not changing performance measures will generate blame in the future; therefore, it is better to change measures before the blame arises. Reorganizing performance measures is a reactive form of blame avoidance when measures are changed as a reaction because these measures revealed poor governmental performance or functioned poorly as indicators of performance. Overall, from the arguments addressing reorganization of performance measures, the following hypothesis can be derived:

Hypothesis 3: A government reorganizes data by changing the measurement system, which avoids blaming the government or individual agencies for poor performance.

As Hood (2011) notes, in the heart of abstinence is the choice not to provide service to citizens. On some occasions, not providing services produces less blame than providing services (Hood, 2011). Thus, abstinence refers to the decision to shut down a service that generates too much blame to mitigate any blame that public actors might face. In this case, the public actor has the right to stop providing this service (Hood, 2002).

Performance measurement can be considered a service that provides citizens with information about the public sector's performance. In the context of performance measurement, abstinence means that the performance remains unmeasured whenever measuring the performance would produce more blame than not measuring it. The logic behind abstinence is the following: If measured results would generate more blame than not measuring results, then inhibiting measurement would lead to lower blame levels. Inhibiting measurement also would eliminate the prospect of complaints from citizens alleging that performance measures inaccurately reflect performance. However, a lack of performance measurement can elicit complaints; thus, a blame avoidance frontier exists in performance measurement (see Figure 1).

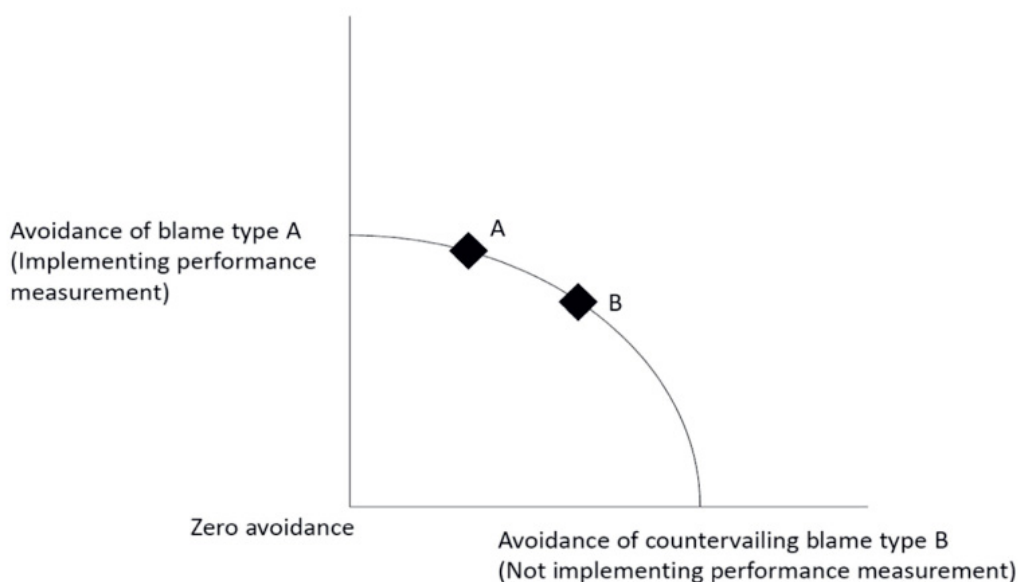


Figure 1. The blame avoidance frontier in performance measurement (adopted and modified from Hood, 2011 and Wiener, 1998)

Abstinence can be either an anticipatory or reactive form of blame avoidance. If performance measures are terminated before any blame arises, abstinence is an anticipatory form of blame avoidance. If performance measures are terminated after blame has risen, then abstinence is a reactive form that is

often applied when performance measurement system is being reorganized. Drawing from the theoretical ideas presented above, the fourth hypothesis used in this study is the following:

Hypothesis 4: The government terminates performance measures that generate blame, which avoids blame for the government or individual agencies' poor performance.

Research methods

The present study answers the proposed research question by using Popper's (1959) hypothetico-deductive method. The universal claim that this study aims to falsify is the following: Performance measurement does not incorporate blame avoidance strategies. To falsify the proposed claim, this study creates four hypotheses that predict the use of specific blame avoidance strategies in performance measurement. If even one hypothesis is accepted, it falsifies the universal idea that performance measurement does not use blame avoidance strategies. If all four hypotheses are rejected, then it can be claimed that this research did not find any support for the claim that blame avoidance strategies are used in performance measurement.

Finland's central government was used as a case-study organization, and the productivity data used in this study were acquired from Statistics Finland (e.g., 2006, 2007, 2008). The data present productivity figures from central government agencies during the 2006–2012 period. Other empirical data used in this study include national news articles (n =112), webpages (n =10), and documents (n =135) that Statistics Finland and other central government agencies have published. The research subject was chosen because the chosen statistic's measurement methods were criticized by National Audit Office of Finland (2011), and the productivity results generated public blame.

Content analysis was used as an analytical method. In a content analysis, descriptions of blame avoidance behavior described in the four theoretical hypotheses operate as coding categories, and the empirical data are coded accordingly. The coding aimed to locate public actors' actions that can be identified as blame avoidance behavior. The content analysis was used to reveal how blame avoidance is embedded in a performance measure. However, the objective of the content analysis was not to demonstrate how frequently blame avoidance is embedded in performance measures.

Empirical analysis

This study's empirical analysis is divided into four sections. The first deals with herding in performance measurement, whereas the second is devoted to secrecy in performance data publishing. The third section examines procedures that reorganize performance measures to create moving targets for critics. The last section describes how abstinence works in performance measurement systems.

Herding as a blame avoidance strategy in performance measurement

The statistics on central government productivity described volume changes in government agencies' total productivity and work productivity, as well as changes in outputs, labor inputs, and total inputs. The data were published online on the level of the whole government and administrative sectors

(Statistics Finland, 2013). Individual agencies' productivity figures were not published (Ministry of Social Affairs and Health, 2008), although they were collected (Statistics Finland, 2006, 2007, 2008, 2009, 2010a). For the sake of simplicity, this article only uses administrative sectors and central government's total productivity to demonstrate herding. The actual total productivity calculation process entailed the following: The numbers of individual agencies were tallied to calculate administrative sectors' total productivity, and administrative sectors' total productivity data were tallied to get the government's total productivity (see Figure 2). Thus, statistics on central government productivity incorporated herding as a blame avoidance strategy by publishing aggregated results and by not granting access to individual agencies' productivity figures (e.g., Statistics Finland, 2010a).

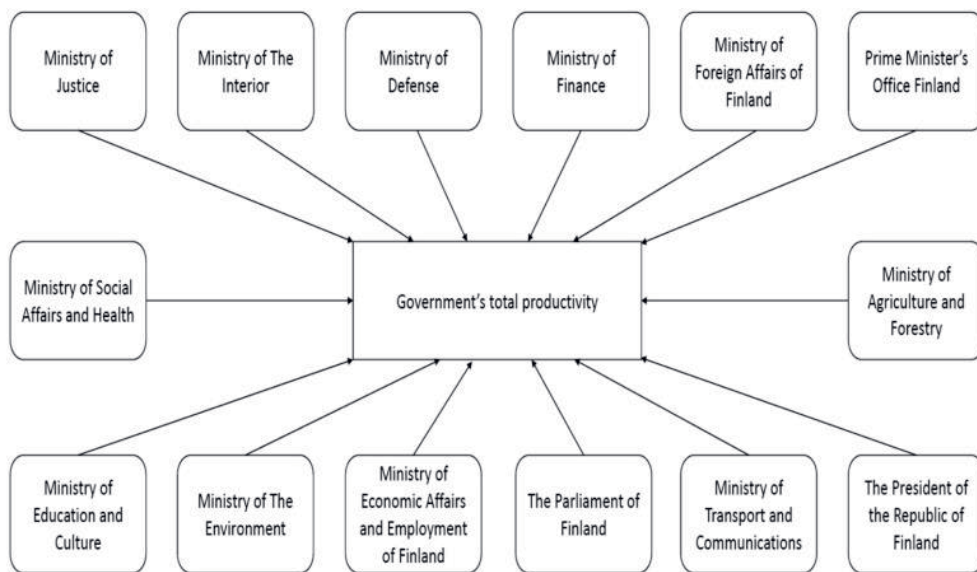


Figure 2. The 14 administrative sectors that were used to calculate the government's total productivity

Publishing productivity data on the total level meant, for example, that in the year 2012, all total productivity figures from 76 government agencies were tallied with a weighting system and that only one number was published online, representing the central government's total productivity (Statistics Finland, 2013). This one number presented the whole government's total productivity across 76 governmental agencies and their total productivity results. Producing the administrative sector's total productivity meant that productivity numbers of the agencies belonging to a certain administrative sector were tallied with a weighting system and that only one number was published online, representing the administrative sector's total productivity. For example, the total productivity of the Ministry of the Environment was calculated from the total productivity numbers of four agencies (see Figure 3).

By having access only to government and administrative sector-level data on productivity, the public could not see how individual government agencies contributed to the whole government or

administrative sector's total productivity. Thus, herding's effects emerged, and the public could not tell how different government agencies contributed to the government or administrative sector's total productivity. Citizens could ask for this information from Statistics Finland, but in the best-case scenario, they would need to pay a fee to get the information (e.g., Statistics Finland, 2018). However, it still was not guaranteed that one could get this information because producing it would have required manual calculations, and Statistics Finland could refuse to provide information that is not directly produced during normal processes relating to central government productivity statistics.

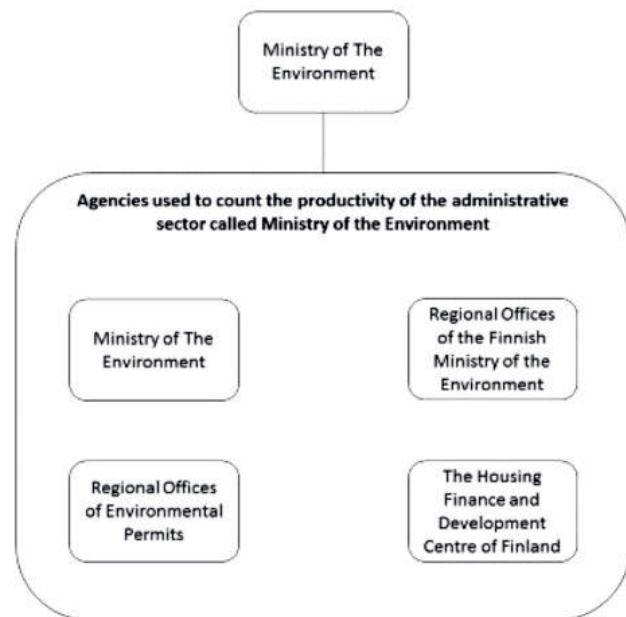


Figure 3. The government agencies that form the administrative sector known as the Ministry of the Environment

Manual calculations were required because the computer program automatically calculated and recorded productivity figures from agencies and administrative sectors, and it did not record which agencies contributed most to government productivity. Moreover, a weighting system was used in the calculations, which inhibited the possibility of inferring from the agency-level data how much a certain agency contributed to administrative sector's total productivity and government's total productivity. For example, when one tallied government agencies' productivity figures to get the numbers describing productivity development within administrative sectors, a weighting system based on total costs' proportions was applied. Because agencies' costs varied annually, the weights used also varied. A similar weighting system based on total costs' proportions was applied when administrative sectors' productivity numbers were tallied to determine the government's total productivity. The computer

program that calculated government productivity data did not record the weighting of different agencies in different years. Thus, the data showing which agencies were most responsible for the government's productivity results were limited.

To conclude, the productivity calculations incorporated herding in two stages. In the first stage, herding, as a blame avoidance strategy, was applied by using the weighting system and aggregating individual agencies' total productivity numbers into a single number representing the administrative sector's total productivity. In each administrative sector, many agencies were grouped together, and their work was presented through one number, representing a particular administrative sector's productivity development. Once this aggregation was done and the administrative sector's total productivity was published online, the public could not tell how well individual agencies were doing their jobs and how they contributed to the administrative sector's total productivity. In the second stage, herding, as a blame avoidance strategy, was applied for the second time by aggregating the administrative sector's total productivity numbers into a single number that represented the government's total productivity. Once again, the public could not tell from this published figure how well individual agencies were doing their jobs and how they contributed to the whole government's productivity data.

Based on the empirical evidence, it seems that data were aggregated in the statistics (Statistics Finland, 2010b) and the media did not blame individual agencies, although the Finnish government's productivity was criticized in the news articles (e.g., Helsingin Sanomat, 2010a; Mainostelevisio, 2005). From the above, it follows that hypothesis one can be accepted as it claimed that the government uses herding by aggregating data, which avoids blaming individual agencies for poor performance.

As Hautakangas and Heikkinen (2008) state in an article published in Statistics Finland, the central government's productivity statistics do not tell why productivity developed the way it did; thus, the statistics do not provide information on how to affect productivity results. This information must be sought from other information sources (Hautakangas, & Heikkinen 2008). Indeed, aggregating agencies' productivity results and limiting the access to data describing the exact productivity calculation hid reasons for the productivity development and generated herding effects in the measurement system. In this context, it is important to note that restricted information is a typical feature of secrecy, which is another blame avoidance strategy (Hood, 2011).

Besides Finland, aggregation also has been typical in other countries. For instance, the United Kingdom, Australia, and New Zealand have measured and reported government productivity at the national level, and all aggregated their productivity numbers (e.g., Campbell & Foxtton, 2016; Industry Commission, 1997; New Zealand Productivity Commission, 2018). The gathered data are aggregated at the sectoral level, and data on individual agencies and organizations were not reported in these three countries.

Using secrecy as a blame avoidance strategy

The principles of secrecy were actualized in statistics on central government productivity as demonstrated in the section addressing herding. The year 2010 provides another example case from this. The Finnish government's total productivity rose over 2% in 2010 (Statistics Finland, 2010a). However, out of the 66 agencies that this statistic comprised, in 33 agencies, productivity fell (see Appendix). The published statistic only showed administrative sectors and the government's productivity development as a whole, while not reporting agencies' productivity numbers. Thus, the public could not see that 33 government agencies trended downward in their productivity development in 2010. The public could see that four administrative sectors trended downward in their productivity development, but could not see which agencies were causing these downward trends (e.g., Statistics Finland, 2010a). The eight administrative sectors that showed positive development in their productivity numbers included agencies that had downward trends in their productivity development, but the public could not see these agencies' data in central government productivity statistics. Thus, in 2010, it was possible that poor results from problem units were buried in indicators that showed positive development to the outside world because well-performing units outweighed the number of problem units in eight administrative sectors, as well as at the whole-government level.

Of course, the secrecy only works in this case if no one from the public asks for the data comprising productivity numbers from individual agencies. However, some factors hindered public curiosity. For example, a service fee had to be paid, and a written service agreement with Statistics Finland was needed to procure such data. The service fee was set according to the work hours and hourly rate of the worker collecting the data, the amount of data that needed to be delivered, and how difficult it would be to collect or produce the data. The hourly rate was 110 euros, and the price for the data used in this study was approximately 500 euros.

The service fee and agreement may discourage citizens from seeking out government agencies' productivity development data, in which case, secrecy, as a blame avoidance strategy, is actualized in the statistics on central government productivity. Another factor contributing to secrecy was that the government's productivity publications did not inform the reader about the option of accessing individual agencies' productivity information. These publications only advised contacting statistics Finland for further inquiries about such statistics (e.g., Statistics Finland, 2010a). As the media did not blame individual agencies in the examined news articles, it is justified to accept the hypothesis two. The hypothesis two claimed that the government uses secrecy by limiting publication of or access to data, which avoids blaming individual agencies for poor performance.

Secrecy relating to productivity data is not an unfamiliar phenomenon in other countries concerning the collection and reporting of such data at the national level. For example, New Zealand, Australia, and the United Kingdom do not publish agency-level data in their government level statistics (e.g., Campbell & Foxton, 2016; Industry Commission, 1997; New Zealand Productivity Commission, 2018).

Reorganizing performance measures to avoid blame

To see how reorganization, as a blame avoidance strategy, exhibited itself in central government productivity statistics, consider the following quotation (Statistics Finland, 2012):

“Statistics recording changes in central government productivity have been produced since 1995. The coverage of the statistics has increased annually, which means that the same units are not included in each measuring point within the time series... The productivity development in the period between 1995–2004 is not fully comparable with the productivity development that occurred after 2004. The change in the weighting method caused the change in comparability. The weighting method was changed because the statistics on central government productivity were developed... From the beginning of the year 2010, the statistics on central government productivity do not include universities that left the state sector in that year. Thus, the time series is not fully comparable with previous years from that point on.”

The above quotation indicates that the unit of analysis measured in the central government productivity statistics was changing constantly. By looking at the data that government agencies reported to Statistics Finland (2006, 2007, 2008, 2009, 2010a), one also can see that many outputs changed within agencies, that is, the production of some welfare services was shut down, and new welfare services emerged in the calculations. Changes in what was being produced also altered the unit of analysis when a particular government agency's productivity was being assessed. Developing output indicators was viewed as a continuous process in central government productivity statistics (Hautakangas & Heikkinen, 2008). Therefore, changes were constant in output indicators, which modified central government productivity statistics. Another change was that governments' business accounting procedures were reformed in 1998. The reforms changed one information source and its content because central government productivity statistics used the government's business accounting data to retrieve agencies' cost information. The calculation methods used to determine coverage of the statistic also were changed in 2004 (Hautakangas & Heikkinen, 2008).

These constant changes in the unit of analysis and calculation methods offered blame avoidance opportunities to public managers and statisticians. Thus, the reorganization provided a moving target for critics wishing to blame government performance or the reliability and validity of statistics on central government productivity. The reorganization inhibited the possibility to observe productivity trends, and therefore critics could not use trend information to criticize government productivity development. As a result of the constant reorganization, the actual productivity figures received less attention in public conversations as these conversations started to concentrate more on whether or not public sector productivity can be measured (Helsingin Sanomat 2010a). Based on the presented empirical evidence, it is justified to accept hypothesis three. This hypothesis stated that government reorganizes data by changing the measurement system, which avoids blaming the government or individual agencies for poor performance.

It is important to note that reforming productivity measurements is not only just a Finnish phenomenon. Australia, New Zealand, and the United Kingdom also reported frequent changes in their national-level productivity measuring methods. These changes can relate to data collection, unit of analysis, and calculation methods (e.g., Campbell & Foxtton, 2016; Industry Commission, 1997; New Zealand Productivity Commission, 2018).

Abstinence

Central government productivity statistics generated blame toward the government through published results showing downward productivity trends. National news outlets reported results from central government productivity statistics annually, so poor results received media attention (e.g., Helsingin Sanomat, 2010b; Mainostelevisio, 2005). Citizens also used the statistical data to criticize the government. A caption from an opinion piece in the most prominent Finnish newspaper captured the common criticism among citizens toward the government in the following way (Helsingin Sanomat, 2010b):

“Statistics Finland has measured the productivity of the public sector for 15 years, and the results are not flattering. State productivity weakened again last year despite the productivity program, or perhaps as a result of the program.”

Moreover, calculation methods used to determine central government productivity statistics also were criticized. For example, a National Audit Office of Finland (2011) report criticized the measurement system itself. The report cited problems in quality assurance, and that the measurement system did not describe the development of service production adequately. Moreover, some government agencies and institutions were unable to produce the required information (National Audit Office of Finland, 2011).

The government chose to cease production of central government productivity statistics in 2014 (Statistics Finland, 2014a); thus, they chose the abstinence route. The need to save 4 million euros through budget cuts was cited as the principal reason for this decision (Statistics Finland, 2014b). The termination meant that citizens could no longer use statistics to criticize the government’s productivity development, and news outlets could no longer report on the government’s productivity or any possible downward trends. Thus, blame generated from the release of productivity statistics vanished. Abstinence also limited blame toward the government’s productivity measurement methods.

Criticism toward the termination was not substantial, and the statistic quietly ceased to exist. After the termination of central government productivity statistics, the media has not reported central government productivity figures anymore as no-one produces such figures. When productivity has been in the news in recent years, the focus has been on cutback policies and these policies’ effects. Cutback policies can affect personal satisfaction and unemployment levels, as well as services and costs. Of course, critical

voices in public discourse still maintain that the government is being unproductive, but these critics do not have central government productivity statistics to back up their claims anymore. The lack of measured evidence makes these claims more contestable in current public conversations; thus, criticism of the Finnish government has weakened, as hard statistical data carry much more weight in public discourse.

Was incorporation of blame avoidance strategies in productivity measurement inevitable?

In most cases, it is possible to avoid the formation of a state in which government's productivity-measurement system incorporates blame avoidance strategies. For example, when considering secrecy as a blame avoidance strategy, the central government actors designing the statistics on central government productivity had a choice concerning what numbers were published. Statistics Finland collected productivity data from individual agencies and calculated their productivity numbers, but did not publish these numbers, although it could have published the data if the government wanted to do so. Statistics Finland also collected data on municipalities, and one could get input and output data on individual municipalities from Statistics Finland, but one could not get any agency data as part of administrative sectors and the central government's productivity numbers. Thus, the central government chose to embed secrecy into central government productivity statistics.

Publishing data only on the administrative sector and central government levels was another choice in the measurement design. Statistics on municipalities also aggregated data to describe, for example, municipalities' total inputs, but these statistics included individual municipalities' data, though only aggregate data could have been published. When considering reorganization of performance measures, reorganization sometimes was given and other times chosen. It was given when agencies chose to change their cost calculating methods or improve their productivity measurement practices, or when political decision makers imposed new tasks on agencies. However, it was Statistics Finland's choice to change the weighting system used in productivity measurement, eliciting adverse effects on comparability.

When it comes to abstinence, the central government wanted to pursue productivity development, but opted mostly to use inputs (i.e., costs or staff-years) as a proxy for productivity (Ministry of Finance, 2017). However, some ministries still voluntarily provide information about their productivity development, measured as the ratio between outputs and inputs (e.g., Ministry of Transportation and Communications, 2017). Even when central government productivity statistics were produced, some of the agencies provided calculated productivity information produced by Statistics Finland in their financial statements, while others did not (e.g., Ministry of Finance, 2010; Ministry of Transportation and Communications, 2010). This shows that abstinence is a choice in performance measurement systems.

Discussion

Many assume that performance measurement is to be used for holding someone accountable (Behn, 2003). Productivity measurement in Finland traditionally has been used for accountability purposes, which can explain why this study found blame avoidance strategies in productivity measurement. In Finnish politics, productivity has been a very appealing notion for a long time, with its importance emphasized in the eyes of politicians and the public. One study described this fascination with government productivity as follows (Kork, Mänttari, & Vakkuri, 2015):

“Productivity appears to be the panacea for dealing with difficult challenges within the public service system in Finland. Demographic changes, evolving needs of the social and healthcare services, and problems of long-term financial sustainability have all been addressed as major challenges... The perceived solution is ‘doing more with less,’ that is, economizing the uses of public finances. While this appears straightforward, it is a highly contested aim in practice.”

Politicians demanded productivity and measures to track it. Prime Minister Vanhanen started the government productivity program in 2003, and with it, the development of productivity measurement. However, many saw the original productivity program as inefficient, and it was terminated in 2011 after heated debate. The following news article describes the debate quite accurately (Kerkkänen, 2010):

“The state productivity program began to really work in 2007. When people retired, jobs have not been filled at the same pace. According to the National Audit Office of Finland, the productivity program has not reached its goals, that is, work productivity has not increased as desired, and the value of the program seems to come from the reduction of the staff. The Parliamentary Audit Committee believes that the entire program should be discontinued in its present form. ‘It is not a productivity improvement program, but a program for reducing the number of personnel years. We want this mechanical model to be terminated,’ says Matti Ahde, chairman of the committee (and a member of the Social Democratic Party).

The public official responsible for the program’s implementation, Tomi Hytönen from the Ministry of Finance, does not accept the allegations about the ineffectiveness of the program. According to him, the State Audit Office assessed the program incorrectly when it found that the productivity of the public sector fell in 2009 compared to the previous year. According to him, the results must be looked at in a longer time frame. ‘The productivity program has worked. The state uses less labor than before, and I have not heard that the level of service has somehow collapsed,’ Hytönen concluded.

The number of state employees has dropped from about 90,000 to 82,000, and in the next few years, there will be a further 5,000 jobs gone. The figures do not include universities and apprenticeships. The reduction comes first and foremost through the retirement of baby boomers.

According to the Ministry of Finance, the implementation of the program is necessary to meet the needs of the private sector. Antti Palola, chairman of the trade union Pardia, believes that the reasoning does not work here. 'I can't really believe that the private sector will have a huge labor shortage. People have been saying for almost 30 years that a labor shortage is coming,' he says. Heikki Taulu, Akava's employment policy advisor, estimates that the program has poor timing. The private sector reduced its recruitment during the recession, and the state's productivity program undermines employment in state government. 'It would be good to put a productivity program on hold at least until unemployment has fallen,' he feels. At the end of November, the Cabinet Committee on Economic Policy stated that the program had progressed largely in line with its objectives and should be continued."

The study's results are in line with studies by Johnsen (2012), Nicholson-Crotty et al. (2006), and Nicholson-Crotty et al. (2017), as these researchers have shown that blame avoidance strategies may be more prevalent in some policy areas compared with others. Indeed, the Finnish central government measures many things, but it was the central government's productivity statistic that was terminated, while many other performance measures remained intact. Productivity policies have raised lots of attention, as the quoted news article demonstrates. The termination was perhaps not so surprising, as the productivity measure and program were both criticized.

This study's results raise several new research questions. For example, future research could focus on how political conflicts shape blame avoidance strategies used in different administrative sectors and policy areas (see e.g., Johnsen, 2012; Nicholson-Crotty et al., 2006). One also could examine how the measurement's purpose reflects blame avoidance strategies used in performance measurement (e.g., Behn, 2003). As the productivity measurement was used mainly for accountability purposes, it might induce more use in blame avoidance strategies than if productivity measures were serving some other purposes.

Conclusions

This research examined blame avoidance behavior in performance measurement system in order to contribute to the extant performance management and accountability literature (see, e.g., Rajala et al.,

2018). The addressed research questions asked how the design of performance measurement systems may facilitate blame avoidance strategies and how governments use these strategies in performance measurement. As a result, the use of four blame avoidance strategies in performance measurement was theorized, and the empirical data demonstrated the use of these strategies and confirmed the four hypotheses that predicted such use. Based on these results, one can reject the claim that performance measurement does not incorporate blame avoidance strategies. In the case examined, blame avoidance in performance measurement was enabled through the following strategies:

1. Herding, in which the government aggregated measures of multiple units into one figure.
2. Reorganization, in which the government reorganized the measures' content constantly.
3. Secrecy, in which the government hid calculations and evaluation methods, and delivered different performance information to public sector actors and external citizens.
4. Abstinence, in which the government terminated criticized performance measures that demonstrated poor results.

This study has many limitations. First, only four types of blame avoidance strategies were examined, while many such strategies were left unexamined (see, e.g., Hood, 2011; Weaver, 1986). Second, one cannot claim that the examined case is representative of other cases that apply blame avoidance strategies in performance measurement. Other research contexts can manifest different approaches to blame avoidance, as the level of political conflicts, purpose of measurement, organizational designs, and policy areas may differ (see, e.g., Behn, 2003; Nicholson-Crotty et al., 2017; Stone, 1995). Thus, generalizing this study's results would be an unreliable inference. However, this study's findings can provide research questions and hypotheses that future research can investigate. Third, the study focused only on productivity measurement, and other measurement types were not addressed (see e.g., Hatry, 2006). Different measurement types may facilitate different blame avoidance strategies.

Another study limitation was that it did not ask designers of the measurement system whether they had blame avoidance motives when they designed the performance measurement system. The performance measurement system's designers unintentionally could have created a measurement system that operates according to the principles of specific blame avoidance strategies. The purpose of measurement also can require particular measurement techniques, such as aggregation; thus, it can be rational from this perspective to use aggregation in the measurement system. However, the designer's motives in this case are irrelevant because the measurement system also can incorporate blame avoidance strategies unintentionally, and this research only aimed to show that blame avoidance behavior can be incorporated into performance measurement systems.

As blame avoidance can be embedded in performance measurement unintentionally or intentionally, it is important to be aware of blame avoidance aspects in performance measures. This study's results can

help practitioners and academics examine blame avoidance aspects in performance measurement systems. These measurement aspects enable public sector actors to change blame avoidance practices if they wish to do so. By showing how blame avoidance strategies are incorporated into productivity measurement, this study also has contributed to extant performance management and accountability research and studies examining blame avoidance behavior (see, e.g., Hinterleitner & Sager, 2015; Hood, 2011; Kloot, 2009). From a methodological perspective, the study demonstrated how blame avoidance behavior in performance measurement systems can be examined empirically.

Because this research only focused on a few blame avoidance strategies, it left quite a few strategies unexamined (see, e.g., Hinterleitner, 2017; Hood, 2011; Weaver, 1986). Future research could explore these unexamined blame avoidance strategies, as well as further examine those covered in this study. Future studies also could examine the relationship between accountability and blame avoidance in the context of performance measurement. For example, it would be interesting to know whether performance measures that enhance accountability also can enable blame avoidance, or are they mutually exclusive features in a performance measurement system. Ultimately, this study's results can offer only a starting point in the quest to understand the relationships among accountability, blame avoidance behavior, and performance measurement systems. Thus, more research is needed, as many unexplored questions and aspects of this topic remain to be examined.

References

- Behn, R. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 588–606. <https://doi.org/10.1111/1540-6210.00322>
- Bok, S. (1983). *Secrets: On the ethics of concealment and revelation*. New York, NY: Vintage Books.
- Boin, A., Hart, P., McConnell, A., & Preston, T. (2010). Leadership style, crisis response, and blame management: The case of Hurricane Katrina. *Public Administration*, 88(3), 706–723. <https://doi.org/10.1111/j.1467-9299.2010.01836.x>
- Bovens, M. (2007). Analyzing and assessing accountability: A conceptual framework. *European Law Journal*, 13(4), 447–468. <https://doi.org/10.1111/j.1468-0386.2007.00378.x>
- Bovens, M. (2010). Two concepts of accountability: Accountability as a virtue and as a mechanism. *West European Politics*, 33(5), 946–967. <https://doi.org/10.1080/01402382.2010.486119>
- Brändström, A., & Kuipers, S. (2003). From “normal incidents” to political crises: Understanding the selective politicization of policy failures. *Government and Opposition*, 38(3), 279–305. <https://doi.org/10.1111/1477-7053.t01-1-00016>
- Campbell, S., & Foxtton, F. (2016). *Public service productivity*. London, UK: Office for National Statistics.
- Charbonneau, E., & Bellavance, F. (2012). Blame avoidance in public reporting: Evidence from a provincially mandated municipal performance measurement regime. *Public Performance & Management Review*, 35(3), 399–421. <https://doi.org/10.2753/PMR1530-9576350301>
- Coe, C. (2003). A report card on report cards. *Public Performance & Management Review*, 27(2), 53–76.

- Collier, P. (2006). Costing police services: The politicization of accounting. *Critical Perspectives on Accounting*, 17(1), 57–86. <https://doi.org/10.1016/j.cpa.2004.02.008>
- Fiorina, M. (1982). Legislative choice of regulatory forms: Legal process or administrative process? *Public Choice*, 39(1), 33–66. <https://doi.org/10.1007/BF00242147>
- Fiorina, M. (1986). Legislator uncertainty, legislative control, and the delegation of legislative power. *Journal of Law, Economics, and Organization*, 2(1), 33–51.
- Flynn, N. (1986). Performance measurement in public sector services. *Policy & Politics*, 14(3), 389–404. <https://doi.org/10.1332/030557386782628172>
- George, B., Desmidt, S., Nielsen, P. A., & Baekgaard, M. (2017). Rational planning and politicians' preferences for spending and reform: Replication and extension of a survey experiment. *Public Management Review*, 19(9), 1251–1271. <https://doi.org/10.1080/14719037.2016.1210905>
- Hatry, H. P. (2006). *Performance measurement: Getting results*. Washington, DC: The Urban Institute.
- Hautakangas, S., & Heikkinen, J. (2008). *Tuottavuuskehitys julkisten palveluiden tuotannossa*. Helsinki, Finland: Talous & Yhteiskunta.
- Heinrich, C., & Choi, Y. (2007). Performance-based contracting in social welfare programs. *American Review of Public Administration*, 37(4), 409–435. <https://doi.org/10.1177/0275074006297553>
- Helsingin Sanomat. (2010a). Julkisen sektorin tuottavuuden mittaamisesta lihava riita. Retrieved from <https://www.hs.fi/kotimaa/art-2000004754244.html>
- Helsingin Sanomat. (2010b). Hallinnon tehostamiset ovat vain luoneet byrokraatiaa. Retrieved from <https://www.hs.fi/mielipide/art-2000004752488.html>
- Hinterleitner, M., & Sager, F. (2015). Avoiding blame: A comprehensive framework and the Australian home insulation program fiasco. *Policy Studies Journal*, 43(1), 139–161.
- Hinterleitner, M. (2017). Reconciling perspectives on blame avoidance behavior. *Political Studies Review*, 15(2), 243–254. <https://doi.org/10.1111/psj.12088>
- Ho, S., & Chan, Y. (2002). Performance measurement and the implementation of balanced scorecards in municipal governments. *Journal of Government Financial Management*, 51(4), 8–19.
- Hood, C. (2002). The risk game and the blame game. *Government and Opposition*, 37(1), 15–37. <https://doi.org/10.1111/1477-7053.00085>
- Hood, C. (2007). What happens when transparency meets blame-avoidance? *Public Management Review*, 9(2), 191–210. <https://doi.org/10.1080/14719030701340275>
- Hood, C. (2011). *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton, NJ: Princeton University Press.
- Hood, C. (2014). Accountability and blame-avoidance. In M. Bovens, R. E. Goodin, & T. Schillemans (Eds.), *The Oxford handbook of public accountability* (pp. 603–616). Oxford, UK: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199641253.013.0007>
- Horn, M. (1995). *The political economy of public administration: Institutional choice in the public sector*. Cambridge, UK: Cambridge University Press.

- James, O. (2010). Performance measures and democracy: Information effects on citizens in field and laboratory experiments. *Journal of Public Administration Research and Theory*, 21(3), 399–418. <https://doi.org/10.1093/jopart/muq057>
- Johnsen, Å. (2005). What does 25 years of experience tell us about the state of performance measurement in public policy and management? *Public Money and Management*, 25(1), 9–17. <https://doi.org/10.1111/j.1467-9302.2005.00445.x>
- Johnsen, Å. (2012). Why does poor performance get so much attention in public policy? *Financial Accountability and Management*, 28(2), 121–142.
- Industry Commission. (1997). *Assessing Australia's productivity performance*. Canberra, Australia: AGPS.
- Kerkkänen, T. (2010). Valtion tuottavuusohjelma halutaan lopettaa - "Ei lisännyt tuottavuutta". Retrieved from <https://yle.fi/uutiset/3-5680686>
- Kork, A., Mänttari, P., & Vakkuri, J. (2015). Complexities of productivity policies in the Finnish public sector: Knowing how to do more with less? *Financial Accountability & Management*, 31(4), 395–414. <https://doi.org/10.1111/faam.12062>
- Kloot, L. (1999). Performance measurement and accountability in Victorian local government. *International Journal of Public Sector Management*, 12(7), 565–584. <https://doi.org/10.1108/09513559910308039>
- Kloot, L. (2009). Performance measurement and accountability in an Australian fire service. *International Journal of Public Sector Management*, 22(2), 128–145. <https://doi.org/10.1108/09513550910934538>
- Lindbom, A. (2007). Obfuscating retrenchment: Swedish welfare policy in the 1990s. *Journal of Public Policy*, 27(2), 129–150. <https://doi.org/10.1017/S0143814X07000645>
- Lombardi, J., Craig, D., Capaldi, E., & Gater, D. (2000). *The top American research universities*. Gainesville, FL: The Center, University of Florida.
- Lonti, Z., & Gregory, R. (2007). Accountability or countability? Performance measurement in the New Zealand public service, 1992–2002. *Australian Journal of Public Administration*, 66(4), 468–484. <https://doi.org/10.1111/j.1467-8500.2007.00557.x>
- Miller, M. (Ed.). (2009). *21st century criminology: A reference handbook*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412971997>
- Ministry of Finance. (2010). *Financial statement*. Helsinki, Finland: Author.
- Ministry of Finance. (2017). *Financial statement*. Helsinki, Finland: Author.
- Ministry of Social Affairs and Health. (2008). *Sosiaali- ja terveystieteiden hallinnon tuottavuuden mittaamisen hanke*. Helsinki, Finland: Yliopistopaino.
- Ministry of Transportation and Communications. (2010). *Ministry of Transportation and Communications*. Helsinki, Finland: Author.
- Ministry of Transportation and Communications. (2017). *Ministry of Transportation and Communications*. Helsinki, Finland: Author.
- Mortensen, P. (2012). "It's the central government's fault": Elected regional officials' use of blame-shifting rhetoric. *Governance*, 25(3), 439–461. <https://doi.org/10.1111/j.1468-0491.2012.01585.x>

- Mainostelevisio. (2005). Valtion tuottavuus koheni. Retrieved from <https://www.mtv.fi/uutiset/talous/artikkeli/valtioon-tuottavuus-koheni/2115556>
- New Zealand Productivity Commission. (2018). Measuring state sector productivity: Final report of the measuring and improving state sector productivity inquiry (Vol. 2). Wellington, New Zealand: Productivity Control.
- Nicholson-Crotty, S., Nicholson-Crotty, J., & Fernandez, S. (2017). Performance and management in the public sector: Testing a model of relative risk aversion. *Public Administration Review*, 77(4), 603–614. <https://doi.org/10.1111/puar.12619>
- Nicholson-Crotty, S., Theobald, N., & Nicholson-Crotty, J. (2006). Disparate measures: Public managers and performance-measurement strategies. *Public Administration Review*, 66(1), 101–113. <https://doi.org/10.1111/j.1540-6210.2006.00559.x>
- Nielsen, P., & Baekgaard, M. (2015). Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment. *Journal of Public Administration Research and Theory*, 25(2), 545–569. <https://doi.org/10.1093/jopart/mut051>.
- Park, S., & Cho, Y. (2014). The influence of executive selection factors on the performance of public sector organizations in Korea. *Public Performance & Management Review*, 37(3), 412–440. <https://doi.org/10.2753/PMR1530-9576370304>
- Pierson, P. (1994). *Dismantling the welfare state? Reagan, Thatcher, and the politics of retrenchment*. Cambridge, UK: Cambridge University Press.
- Pierson, P. (1996). The new politics of the welfare state. *World Politics*, 48(2), 143–179.
- Popper, K. R. (1959). *The logic of scientific discovery*. London, UK: Routledge.
- Rajala, T., Laihonon, H., & Vakkuri, J. (2018). Shifting from output to outcome measurement in public administration arguments revisited. In E. Borgonovi, E. Anessi-Pessina, & C. Bianchi (Eds.), *Outcome-based performance management in the public sector* (pp. 3–23). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-57018-1_1
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 18(2–3), 277–310. <https://doi.org/10.1080/01900699508525011>
- Sulitzeanu-Kenan, R. (2006). If they get it right: An experimental test of the effects of the appointment and reports of UK public inquiries. *International Journal of Public Administration*, 84(3), 623–653. <https://doi.org/10.1111/j.1467-9299.2006.00605.x>
- Statistics Finland. (2006). *Statistics on central government productivity*. Helsinki, Finland: Author.
- Statistics Finland. (2007). *Statistics on central government productivity*. Helsinki, Finland: Author.
- Statistics Finland. (2008). *Statistics on central government productivity*. Helsinki, Finland: Author.
- Statistics Finland. (2009). *Statistics on central government productivity*. Helsinki, Finland: Author.
- Statistics Finland. (2010a). *Statistics on central government productivity*. Helsinki, Finland: Author.
- Statistics Finland. (2010b). *Valtion tuottavuustilaston laatuseloste*. Helsinki, Finland: Author.
- Statistics Finland. (2012). *Valtion tuottavuustilaston laatuseloste*. Helsinki, Finland: Author.
- Statistics Finland. (2013). *Statistics on central government productivity*. Helsinki, Finland: Author.

- Statistics Finland (2014a). Tiedonkäyttäjää kuullaan tilastojen muutoksista [referred: 27.8.2018]. Retrieved from https://www.stat.fi/ajk/tiedotteet/2014/uutinen_029_2014-09-26.html
- Statistics Finland (2014b). Muutoksia tilastojen ja tuotteiden tuotannossa 2015 [referred: 27.8.2018]. Retrieved from https://www.stat.fi/ajk/tiedotteet/2014/uutinen_040_2014-12-18.html
- Statistics Finland. (2018). Maksullisuus ja hinnat [referred: 27.8.2018]. Retrieved from <https://stat.fi/tup/hinnat/index.html>
- Stone, B. (1995). Administrative accountability in the “Westminster” democracies: Toward a new conceptual framework. *Governance*, 8(4), 505–526. <https://doi.org/10.1111/j.1468-0491.1995.tb00225.x>
- National Audit Office of Finland. (2011). Tuottavuuden mittaaminen valtiolla. Helsinki, Finland: Edita Prima Oy.
- Thompson, D. (2014). Responsibility for failures of government: The problem of many hands. *American Review of Public Administration*, 44(3), 259–273. <https://doi.org/10.1177/0275074014524013>
- Tilbury, C. (2006). Accountability via performance measurement: The case of child protection services. *Australian Journal of Public Administration*, 65(3), 48–61. <https://doi.org/10.1111/j.1467-8500.2006.00493a.x>
- Van Dooren, W., & Van de Walle, S. (2011). Performance information in the public sector: How it is used. Houndmills, UK: Palgrave Macmillan. <https://doi.org/10.1007/978-1-137-10541-7>
- Van Peursem, K., Pratt, M., & Lawrence, S. (1995). Health management performance: A review of measures and indicators. *Accounting, Auditing & Accountability Journal*, 8(5), 34–70. <https://doi.org/10.1108/09513579510103254>
- Vis, B. (2009). Governments and unpopular social policy Reform: Biting the bullet or steering clear? *European Journal of Political Research*, 48(1), 31–57. <https://doi.org/10.1111/j.1475-6765.2008.00783.x>
- Vis, B., & van Kersbergen, K. (2007). Why and how do political actors pursue risky reforms? *Journal of Theoretical Politics*, 19(2), 153–172. <https://doi.org/10.1177/0951629807074268>
- Wagschal, U., & Wenzelburger, G. (2008). Roads to success: Budget consolidations in OECD countries. *Journal of Public Policy*, 28(3), 309–339. <https://doi.org/10.1017/S0143814X08000901>
- Weaver, K. (1986). The politics of blame avoidance. *Journal of Public Policy*, 6(4), 371–398. <https://doi.org/10.1017/S0143814X00004219>
- Wiener, J. (1998). Managing the iatrogenic risk of risk management. *Risk: Health, Safety, and Environment*, 9(1), 39–82.
- Wenzelburger, G. (2011). Political strategies and fiscal retrenchment: Evidence from four countries. *West European Politics*, 34(6), 1151–1184. <https://doi.org/10.1080/01402382.2011.572385>
- Wholey, J., & Hatry, H. (1992). The case for performance monitoring. *Public Administration Review*, 52(6), 604–610.
- Wisniewski, M., & Stewart, D. (2004). Performance measurement for citizens: The case of Scottish local authorities. *International Journal of Public Sector Management*, 17(3), 222–233. <https://doi.org/10.1108/09513550410530153>

Zohlnhöfer, R. (2007). The politics of budget consolidation in Britain and Germany: The Impact of blame-avoidance opportunities. *West European Politics*, 30(5), 1120–1138.
<https://doi.org/10.1080/01402380701617498>

Appendix 1

Productivity indexes of administrative sectors and agencies in 2010

Administrative sector and its acronym	Productivity index
	MISSING
Ministry of Foreign Affairs of Finland (MFAF)	DATA
Ministry of Justice (MJ)	97.6
Ministry of The Interior (MI)	101.3
Ministry of Defence (MD)	105.5
Ministry of Finance (MF)	100.4
Ministry of Education and Culture (MEC)	96.5
Ministry of Agriculture and Forestry (MAF)	106.1
Ministry of Transport and Communications (MTC)	101.8
Ministry of Economic Affairs and Employment of Finland (MEAEF)	98.8
Ministry of Social Affairs and Health (MSAH)	102
Ministry of Environment (ME)	102.5
Rest of the administrative sectors in total (see figure 2) (RAS)	78.4
	Productivity index
Agency and the acronym of the administrative sector it belongs to	index
	MISSING
The Parliament of Finland (RAS)	DATA
The national audit office of Finland (RAS)	78.41554
	MISSING
Finnish Institute of International Affairs (RAS)	DATA
	MISSING
The Office of the President of Republic of Finland (RAS)	DATA
	MISSING
The Prime Minister's Office (RAS)	DATA
	MISSING
Ministry of Foreign Affairs of Finland (MFAF)	DATA
	MISSING
Ministry of Justice (MJ)	DATA
Ministry of Justice/ special courts (MJ)	105.5723
	MISSING
Criminal Policy Institute (MJ)	DATA
	MISSING
Office of Data Protection Ombudsman (MJ)	DATA
Safety Investigation Authority (MJ)	107.2266
The Office of Bankruptcy Ombudsman (MJ)	80.53291
	MISSING
Juridical Administration's Center of Technology (MJ)	DATA
Legal Register Centre (MJ)	113.6976
	MISSING
Institute of Criminology and Legal Policy (MJ)	DATA
Supreme courts (MJ)	97.37602
Court of Appeals (MJ)	97.79143

Administrative courts (MJ)	101.3635
Supreme Administrative Courts (MJ)	105.5767
District Courts (MJ)	90.47879
Legal Aid Offices (MJ)	100.6142
National Administrative Office for Enforcement (MJ)	106.7205
Prosecutors (MJ)	101.2055
Criminal Sanctions Agency (MJ)	96.4629
The Training Institute for Prison and Probation Services (MJ)	MISSING DATA
Ministry of the Interior (MI)	MISSING DATA
The Finnish Border Guard (MI)	106.7171
Aland's Government Agency (MF)	MISSING DATA
National Bureau of Investigation (4MI)	112.0753
Administration's Centre of Technology (MI)	MISSING DATA
Finnish Security Intelligence Service (Mi)	MISSING DATA
National Traffic Police (MI)	107.9008
Police Technical Centre (MI)	MISSING DATA
Emergency Services College (MI)	95.50672
Police University College (MI)	128.0274
National Police Board of Finland (Mi)	MISSING DATA
Population Register Centre (MF)	MISSING DATA
Finnish Immigration Service (MI)	73.85306
Regional State Administrative Agencies (MF)	MISSING DATA
Emergency Response Centre Agency (MI)	95.86473
Regional Police Departments (MI)	99.05326
Ministry of Defence (MD)	MISSING DATA
The Finnish Defence Forces (MD)	106.0614
Construction Establishment of Finnish Defence Administration (MD)	95.43244
Ministry of Finance (MF)	MISSING DATA
State Treasury (MF)	89.16449
Customs (MF)	110.1117
Tax Administration (MF)	99.1242
VATT Institute for Economic Research (MF)	100.2735
Finnish Government Shared Services Centre for Finance and HR	MISSING DATA
Statistics Finland (MF)	102.3633
Ministry of Agriculture and Forestry (MAF)	MISSING DATA
Appeal Board for Agriculture trade (MAF)	89.614

Information Centre of the Ministry of Agriculture and Forestry (MAF)	MISSING DATA
National Land Survey of Finland (MAF)	108.2652
Finnish Forest Research Institute (MAF)	107.1172
Finnish Food Safety Authority (MAF)	107.5563
Research Centre for Agriculture and Food Economy (MAF)	97.21049
Finnish Game and Fisheries Research Institute (MAF)	111.3911
Finnish Geodetic Institute (MAF)	MISSING DATA
Agency for Rural Affairs (MAF)	MISSING DATA
Ministry of Transportation and Communications (MTC)	95.31476
Finnish Transport Agency (MTC)	101.5981
Finnish Transport Safety Agency (MTC)	104.5753
Ministry of Economic Affairs and Employment of Finland (MEAEF)	76.31986
Finnish Meteorological Institute (MTC)	101.8281
Finnish Communications Regulatory Authority (MTC)	104.3005
Technical Research Centre of Finland (MEAEF)	103.6305
Geological Survey of Finland (MEAEF)	82.82827
Visit Finland (MEAEF)	MISSING DATA
Finnish Patent and Registration Office (MEAEF)	98.53014
Tekes (MEAEF)	MISSING DATA
Finnish Competition Authority (MEAEF)	89.76966
Finnish Consumer Authority (MEAEF)	95.90476
Finnish Safety and Chemical Agency (MEAEF)	84.28944
National Metrology Institute of Finland (MEAEF)	MISSING DATA
Centre for Consumer Society Research (MEAEF)	89.19836
The Consumer Disputes Board (MJ)	110.7028
Energy Authority (MEAEF)	MISSING DATA
Center for Economic Development, Transport and the Environment (MEAEF)	MISSING DATA
Ministry of Social Affairs and Health (MSAH)	116.9754
Unemployment Benefit Appeal Board (MSAH)	111.6177
Social Security Appeal Board (MSAH)	106.1307
Radiation and Nuclear Safety Authority (MSAH)	MISSING DATA
National Institute for Health and Welfare / research (MSAH)	MISSING DATA
Reform schools of the State (MSAH)	89.2304
Mental Institutions of the State (MSAH)	95.13413
Finnish Medicines Agency (MSAH)	91.2058
National Supervisory Authority for Welfare and Health (MSAH)	MISSING DATA
Ministry of Education and Culture (MEC)	MISSING DATA

Finnish Matriculation Examination Board (MEC)	MISSING DATA
Student Financial Aid Appeal Board (MEC)	96.10485
Institute for the Languages of Finland (MEC)	MISSING DATA
National Audiovisual Institute (MEC)	MISSING DATA
The Finnish Board of Film Classification (MEC)	66.3947
Russia and Eastern Europe Institute (MEC)	MISSING DATA
National Repository Library (MEC)	78.624
Finnish National Agency for Education (MEC)	122.9401
Celia (MEC)	MISSING DATA
Arts Promotion Centre Finland (MEC)	MISSING DATA
National Archives of Finland (MEC)	101.1456
Academy of Finland – Administrative Branch (MEC)	MISSING DATA
Academy of Finland – Research (MEC)	MISSING DATA
Finnish Heritage Agency (MEC)	MISSING DATA
The Governing Body of Suomenlinna (MEC)	114.463
National Gallery (MEC)	81.39945
Finnish national agency for education / Special schools (MEC)	96.1924
Finnish national agency for education / Language scholl (MEC)	93.90396
Finnish national agency for education / Career Colleges (MEC)	MISSING DATA
Finnish national agency for education / Training facilities (MEC)	93.66107
Finnish national agency for education / Administrative branch (MEC)	MISSING DATA
Ministry of Environment (ME)	MISSING DATA
The Housing Finance and Development Centre of Finland (ME)	102.5031
Finnish Environment Institute (ME)	MISSING DATA

Note. Missing data indicates that the agency could not provide the needed figures. Values over 100 show upward trend in productivity and values below 100 demonstrate downward trend in productivity.

Post-print

Rajala, T. (2019), "Numerical performance information in presidential rhetoric: Comparing Estonia and Lithuania", *Journal of Accounting in Emerging Economies*, Vol. 10 No. 3, pp. 321-344.
<https://doi.org/10.1108/JAEE-10-2018-0119>

Numerical performance information in presidential rhetoric - Comparing Estonia and Lithuania

Tomi Rajala, University teacher, Ph.D. student

Tampere university, Faculty of Management and Business, FI-33014 Tampere University, Finland

email tomi.rajala@tuni.fi

Abstract

Purpose – Presidents have constitutional powers and are incentivized to use performance information that is essential to economic leadership practices. However, presidents have not previously been studied in this context. The purpose of this paper is to examine how two sitting presidents use numerical performance information in their speeches. A speech is a formal talk given to a large number of individuals at a particular instance.

Design/methodology/approach – Empirical data were obtained from 85 presidential speeches given by the president of Estonia and 35 by the president of Lithuania. The speeches were analyzed using qualitative and quantitative content analysis. Inductive inference, descriptive statistics and statistical tests were used to propose new theoretical ideas for future research.

Findings – Studied presidents used extensively numerical performance information, primarily outcome information. Also, the presidents used performance information differently, even though both presidents operated in a similar political context and had similar individual characteristics. The differences were in part explained by speech length but not speech context. Older age, doctoral degree, and longer administrative and political career were associated with lower use.

Practical implications – The study provides preliminary results on how presidents use performance information and what type of performance information is most useful in presidential speeches that address the nation and conduct economic leadership.

Originality/value – New analytical models are presented that can be used to study the intensity of performance information use in rhetoric. Conceptual definitions of the various levels of intensity in performance information use are also introduced. In general, presidential performance information use adds a new dimension to existing research.

Keywords Outcome information, Performance information use, Performance information user, Politician, President, Public performance management

Introduction

In many nations with a presidential system of government, the president is the head of the state (see Linz, 1990). Presidents as performance information users have long been neglected by accounting and performance management scholars (e.g. ter Bogt, 2001; Askim, 2009; Hausteijn et al., 2019). Performance information refers to data about performance and is systematically produced, collected, stored and used to improve the performance of public organizations (e.g. Pollitt, 2006; van Thiel and Leeuw, 2002). The content of performance information reflects what performance measures are measuring. Performance measures either measure past performance, current performance or forecast future performance (Bogan and English, 1994; van Helden and Reichard, 2019), and they can be divided into input, process, output, outcome, workload and productivity measures (Pollitt and Bouckaert, 2004; Hatry, 2006).

The literature has focused on local councilors, state legislators and members of parliament or congressional representatives as performance information users (Rhee, 2014; Lu and Willoughby, 2015; Grossi et al., 2016; Johanson et al., 2019), and no studies conducted in central and eastern Europe have focused on presidents as performance information users (see Nakrošis, 2008; Raudla and Savi, 2015). In the current study, presidents as performance information users are appropriate study subjects for two reasons. First, presidents typically have considerable constitutional powers (Linz, 1990), making them powerful public sector actors who influence national matters by making decisions based on performance information. Second, presidents have incentives to use performance information in their speeches when addressing the nation.

There are two incentives for including performance information in presidential speeches. First, performance information in rhetoric can guide the public's attention toward the performance areas that the president wants to focus on (see Cohen, 1995). As Behr and Iyengar (1985) illustrated, when a president gives speeches dedicated to a single problem in the public sector performance, the public tends to become more concerned about this problem. For presidents, numbers on performance are helpful for describing and framing the focus area. Second, presidents can provide economic leadership to the nation using rhetoric (Wood, 2004), and it is difficult to discuss economics without referring to economic performance information, such as economic growth, balance of trade, employment and inflation. Therefore, turns in the economy are usually described with performance information. Economic leadership is vital to gaining public approval, and presidents' rhetoric on performance provides cues for consumers and companies and affects the tone under which these actors make spending and investment decisions in the economy (Wood et al., 2005). Moreover, economic leadership operating through rhetoric affects the behaviors of consumers and companies and indicates that the president is leading macroeconomic performance. Macroeconomic performance gives credibility to the president and is a strong determinant of the president's job approval rating, which is essential to

presidential success (e.g. MacKuen, 1983; Ostrom and Smith, 1993; Edwards, 2003). Performance information can help achieve job approval and presidential success.

To fill the research gap concerning how presidents use performance information, this study examines the performance information use habits of two presidents. Previously, researchers have inquired about how politicians in local, state or central government use performance information (e.g. Ezzamel et al., 2004; Pollitt, 2006). This study adopts the following research questions from the literature (note that the original research questions used the word politician in the research questions, not the word president):

RQ1. Is performance information actually used by presidents (adopted from Raudla and Savi, 2015)?

RQ2. To what extent are presidents using performance information (adopted from Raudla, 2012)?

RQ3. What type of performance information is being used by presidents (adopted from Jorge et al., 2016)?

RQ4. Do some presidents make more use of performance information than others (adopted from Askim, 2007)?

RQ5. What explains differences between presidents in performance information use (adopted from Askim, 2007)?

The hypotheses related to these questions were derived from the literature and tested. As empirical data, this study analyzes the presidential speeches of two sitting presidents: Kersti Kaljulaid (Estonia) and Dalia Grybauskaitė (Lithuania). Qualitative and quantitative content analysis techniques are applied in the research, and new methods of assessing performance information use in speeches are developed. The findings show that the two presidents use performance information extensively and that this information mainly concerns outcomes. Thus, it is proposed that presidents might differ from other politicians in terms of performance information use, as they are not input oriented. Moreover, the examined presidents have significantly different performance information use habits even though similar contextual and individual attributes influenced their use. However, the presidents were not identical in terms of age, education, occupational background and length of their political career, and these differences are associated with dissimilar use. Therefore, more research is needed, as this study only examines two presidents. As a methodological contribution, this study develops methods to examine numerical performance information use and its intensity in rhetoric.

Next, the research context is explained in detail. In the third section, the relevant literature is reviewed, and the most relevant results are displayed. In the fourth section, the research method is explained. The fifth section reports the empirical analysis. The last section is devoted to discussion and conclusions.

Research context

As this is a comparative case study based on the most similar cases design, the chosen countries and presidents are as similar as possible (e.g. Gerring, 2007). The Soviet Union formerly occupied Estonia and Lithuania, which both share borders with Russia. These countries are geographically small, located in the Baltic area and have experienced similar historical development in recent decades (Auers, 2015). Estonia and Lithuania are also members of the United Nations, European Union and North Atlantic Treaty Organization, and they joined these organizations at the same time. Estonia adopted the euro currency in 2011 and Lithuania in 2015.

Similarities can also be seen between the two presidents. The political context in both countries is confrontational and volatile in terms of parties and party systems as these change constantly but are stable with regard to the political elite and central policies. Presidential powers are the same in both countries, except for powers associated with initiating legislation (Table 1). Modernization, westernization and democratization processes are similar in both countries (Auers, 2015). Finally, the overall socioeconomic and demographic developments of their countries are alike (OECD, 2016).

Table 1. Comparison of presidential powers (Auers, 2015; President of Estonia, 2018; President of Lithuania, 2018)

Powers	Estonia	Lithuania
Call early elections	Yes (partial)	Yes (partial)
Nominate prime minister	Yes	Yes
Initiate legislation	No (except for amendments to the constitution)	Yes
Has suspensive veto	Yes	Yes
Appoint commander-in-chief	No	No
Nominate chief justice of the supreme court	Yes	Yes
Make foreign policy	Yes (partial)	Yes (partial)
Appoint and recall diplomats	Yes (partial)	Yes (partial)
Ratify treaties	Yes	Yes

The presidents also share similar personal characteristics (Table 2). Both presidents are women and are “middle-aged” being between the ages of 45 and 65 (e.g. Erikson, 1982). The presidents both studied finance and economics at university, and their employment backgrounds include tasks involving financial management and economics. This level of similarity is used to test whether the demonstrated similarity predicts similar performance information use. If the presidents differ in their performance

information use, the differences seen in Tables 1 and 2 can be used to assess these dissimilar performance information uses. One key difference between the presidents not presented in Table 2 is that Grybauskaitė has more political experience because she has been the president of Lithuania from 2009 to 2019, whereas Kaljulaid has served only two years as president.

Table 2. Similarities in personal attributes of the presidents

Variable	Kaljulaid	Grybauskaitė
Age	49	62
Sex	Female	Female
Education	Master's degree in Business Management	Doctoral degree in Economics
Career	<ul style="list-style-type: none"> - Member of the European Court of Auditors 2004–2016 - CFO and CEO of the Iru Power Plant of Eesti Energia 2002–2004 - Worked in various Estonian companies 1994–1999 	<ul style="list-style-type: none"> - Head of the Agriculture Division at Vilnius High Party School 1983–1984; lecturer at the Department of Political Economy 1985–1990 - Secretary at the Lithuanian Institute of Economics under the Ministry of Economics of the Republic of Lithuania 1990–1991 - Program director in Prime Minister's office and Director of the European Department at the Ministry of International Economic Relations 1991 - Director of the Economic Relations Department at the Ministry of Foreign Affairs 1993 - Envoy Extraordinary and Minister Plenipotentiary at the Lithuanian Mission to the EU 1994 - Minister Plenipotentiary at the Lithuanian Embassy in the U.S. 1996–1999

Presidents as performance information users

Past studies have scrutinized whether or not politicians use performance information and if so, how and why they use it (e.g. Askim, 2007; Raudla, 2012). According to past results, politicians do use performance information but mostly in a limited fashion (see Table 3). It has also been claimed that contextual conditions and characteristics of individuals affect performance information use and provide reasons why performance information is actually used by politicians (Moynihan, 2005; Buylen and Christiaens, 2016). While research has focused more on public managers as performance information users than citizens and politicians (ter Bogt, 2004), Pollitt (2006) and van Helden (2016) have demanded more research on politicians' and citizens' performance information use.

By studying the presidential use of performance information, it is possible to contribute new knowledge to the field. In Table 3, none of the studies examining politicians' performance information focuses on

presidents. Although this study provides some insights on why presidents might be using performance information, the focus is on whether performance information is being used and how it is used by the presidents. The first research question asks whether performance information is actually used by presidents. Previous research has shown that politicians do use performance information, as every study examined in this literature review confirmed this finding (see Table 3). Thus, the following hypothesis is derived:

H1: Both the Estonian and Lithuanian presidents use performance information in their speeches.

The second research question asks to what extent presidents use performance information. Some studies have reported extensive use, while others have reported limited use among the politicians (see Table 3). Limited use has been more often cited than non-existent, moderate, high or extensive use (e.g., Chen and Smith, 2019). Thus, it is hypothesized that:

H2: Both presidents demonstrate limited performance information use in their speeches.

The third research question asks what type of performance information is being used by the presidents. Prior studies found that information regarding input, process, output, outcome, efficiency and cost-effectiveness are being used by politicians, with input information being the most used in the public sector (e.g. Moynihan, 2005). In regard to process, output, outcome, efficiency and accounting information, the intensity of use varies from non-use to extensive use depending on the information type, context and the user (e.g. Ho, 2005; Askim, 2009; Charbonneau and Bellavance, 2015; Giacomini et al., 2016). Thus, the literature justifies the following research hypothesis:

H3: Both presidents most often use input information in their speeches.

Table 3. Timeline of research focusing on politicians' performance information use

<i>Researcher(s) and government level under study</i>	<i>Is performance information actually ever used by politicians?</i>	<i>To what extent are politicians using performance information?</i>	<i>What type of performance information is being used and how intense is this use?</i>
ter Bogt (2001), local government	Yes	Limited use	Dutch local politicians focused relatively little on outputs
ter Bogt (2004), local government	Yes	Limited use	The output-oriented performance information is not seen as valuable by the aldermen, and they use it rarely
Ospina <i>et al.</i> (2004), central government	Yes	Limited use	The outcome and performance information is in limited use among politicians
Ezzamel (2004), central government	Yes	Limited use and high use	Accounting information is used in a limited fashion, while performance information use is widespread among politicians
Moynihan (2005), state government	Yes	Limited use and high use	Output and outcome information is used in a symbolic manner, and input information is used extensively in budgeting by politicians
Ho (2005), city (local) government	Yes	Varies from extensive use to no use	Input, output, outcome, and efficiency information use varies in different cities
Pollitt (2006), central and local government	Yes	Limited use	Information on the outputs and outcomes of public programs and organizations is seldom used by politicians
Sterck and Scheers (2006), central (federal) government	Yes	Limited use	Not much evidence that budgetary decision-making or legislative oversight functions would be based on performance information (output and outcome information)
Askim (2007), local government	Yes	High use	Surprisingly high use of performance (input, output, and outcome) information
Posner and Fantone (2007), central (federal) government	Yes	Information use varies	Use of performance information relating to program assessment rating tool (PART) varies
Nakrosis (2008), central government	Yes	Limited use	Use of output and outcome information is limited

Wang (2008), state and local government	Yes	Moderate use	Output and outcome information is incorporated into budgeting processes in moderate levels
Askim (2009), local government	Yes	Varies from high use to no use	Councilors' use of input, output, and outcome information varies depending on their individual attributes
Johansson and Siverbo (2009), local government	Yes	Varies from high use to no use	Quality, productivity, and efficiency benchmark information use varies depending on the local government
Joyce (2011), central (federal) government	Yes	Limited use	Historically, there has been little appetite in the Congress for evidence-based decision-making
Raudla (2012), central government	Yes	Limited use	Output and outcome information is used in a limited fashion
Kroll and Proeller (2013), local government	Yes	High and low use	Input, output, outcome, efficiency, quality, and effectiveness information use varies
Saliterer and Korac (2013), local government	Yes	High and extensive use	Mayors and head officials frequently use outcome and efficiency measures
Saliterer and Korac (2014), local government	Yes	High and extensive use	Mayors use outcome and efficiency measures on a regular basis
Raudla and Savi (2015), central government	Yes	Limited use	Politicians generally do not use output and outcome information
Lu and Willoughby (2015), state government	Yes	Moderate use	Budgeting for performance is not a strong or weak practice but in between
Charbonneau and Bellavance (2015), local government	Yes	Varies from high use to no use	Input, output, outcome, and efficiency information use varies depending on the local government
Guarini (2016), local government	Yes	Intensity of use was not studied	Accounting information is used in all three cases and demonstrates perverse performance information use
Grossi <i>et al.</i> (2016), local government	Yes	Limited use	The use of efficiency, effectiveness, and quality information is modest and varies

Byulen and Christiaens (2016), local government	Yes	Varies from extensive use to no use	The use varies depending on the type of financial information
Bjornholt <i>et al.</i> (2016), local government	Yes	More than half of respondents strongly agree that they use performance information, but the intensity of use is not addressed	Input, output, and outcome information is used
Van Helden (2016), local, state, and central government	Yes	Varies from extensive use to no use	Input, output, outcome, and efficiency information use varies depending on the study
Giacomini <i>et al.</i> (2016), local government	Yes	Varies from high to low use	Cost, efficiency, and effectiveness information use varies depending on the political conflict
Hijal-Mograhabi (2017), local government	Yes	Varies from high to no use	Output, outcome, and efficiency measures are often used to achieve accountability but not in budgeting
Kenk and Haldma (2019)	Yes	Varies from higher to lower use	Performance information use varies depending on the local government merger type
Ellul and Hodges (2019)	Yes	Varies from high to low use	Input information is often used whereas output, outcome and efficiency information is rarely used
Jorge, de Jesus and Nogueira (2019)	Yes	Varies from occasional use to no use	Accounting information use varies depending on the information type
Sinervo and Haapala (2019)	Yes	Varies from higher to lower use	Financial information use varies depending political experience and financial expertise

The fourth research question asks whether some presidents make more use of performance information than others do. Prior research has found that political and financial contexts (Byulen and Christiaens, 2016), educational and employment backgrounds (Askim, 2009; Ouda and Klischewski, 2019), the powers of the user (Cunningham and Harris, 2005; Bourdeaux, 2006), cohort and sex (ter Bogt, 2004) are determinants of politicians' performance information use. Because this study views the presidents as similar in terms of their political and financial context, education and employment background, powers, cohort and sex, it is assumed that the presidents use performance information similarly.

However, as the contexts in which the two presidents give their speeches are different (see Table 4); one would expect to see dissimilar performance information use according to Askim (2009) because similar people may use performance information differently when the context changes, as demonstrated by Giacomini et al. (2016). Moreover, the differences in characteristics presented in the Research context section can also anticipate different types of use. The following hypotheses were derived for the statistical tests:

H4: In their speeches, both presidents do not use the same amount of performance information on average (comparison of the means).

H5: The performance information use of the presidents varies differently between the different speeches (comparison of standard deviations).

H6: Proportions of speeches including performance information use are not equal when Kaljulaid's speeches are compared with Grybauskaitė's speeches (comparison of proportions of speeches).

H7: The presidents do not have the same ratio of performance information and other content in their speeches (comparison of proportions of content).

H8: The presidents do not use the same types of performance measures equally often (comparison of proportions of measures).

In H6, the proportion of speeches including performance information is measured by dividing presidential speeches including performance information with the total number of speeches a president made. In H7, the ratio of performance information and other content is calculated by dividing the amount of words that are numbers representing performance with the total number of words used by the president in the speeches.

The hypothesis testing performed in this study adds the presidential dimension to the previous research findings and shows how the performance information use habits of different actors can be compared via speeches (see H4–H8). If the presidents are not similar in terms of their performance information use, this study attempts to determine why, thus answering the fifth research question that asks what explains the differences in presidents' performance information use.

Table 4. Speech contexts in the presidential speeches

Context	Kaljulaid (Estonia)	Grybauskaitė (Lithuania)
Ceremony (dinners, funerals, award shows, anniversaries, opening ceremonies, etc.)	33 (38.8%)	26 (74.3%)
Conference	37 (43.5%)	4 (11.4%)
Lecture (lectures in universities and institutes)	5 (5.9%)	0 (0%)
Official speech (UN assemblies, new year's speech, state visits, opening session in parliament, etc.)	10 (11.8%)	5 (14.3%)

Research method

The empirical data were derived from 120 presidential speeches given between January 1, 2016 and October 29, 2018. The president of Estonia gave 85 speeches and the president of Lithuania gave 35 speeches. All of these speeches were available online at the time of writing this paper. The context of these speeches can be seen in Table 4. The written transcripts used in this study were written in English but one can also find them in other languages (e.g. Russian or the mother language of the country).

The collected speeches were given at a wide array of forums and events. The contexts in which the president made their speeches were significantly different after the three categories of conference, lecture and official speeches were merged into one so that the χ^2 test could be conducted ($p = 0.000$). This indicates that presidents made their speeches in very different forums, which could explain possible differences in performance information use. As the speeches are in English and shared on the internet, the audience for all these speeches is the part of the population who can access the internet. Therefore, the speeches can be identified as mass communication. According to Berger (1995), "mass communication refers to a process of spreading texts and messages to large audiences through the use of the media." There was no list of participants available so it was not possible to assess who were the live audience. In the speeches, the topics covered foreign and domestic policy issues as well as legal matters. Therefore, these speeches should provide information on how performance information is used in various policy contexts. As topics changed even during the speeches, it was not possible to examine within the word limits of this paper whether certain topics relate to performance information use or not. This should be done in a separate study.

For data management, all the speeches were first converted to PDF format and then converted to a Word document format to accelerate the collection and storage process. The Word files were then downloaded to Atlas.ti 8 software where the actual coding was performed. A content analysis method was adopted

for the coding and focused on quantitative performance information. In the first phase of coding, an auto-coding tool was used to find all the numerical information used in the speeches. This tool prevented the introduction of human errors in the search process, and a computer was used to locate all the numerical data in the speeches. The numerical data could be in numerical (e.g. 0–9) or written format (e.g. one, two, hundred, million, first, second, etc.). The computer searched for cardinal, ordinal and nominal forms of numbers (see Appendix 1 for the search words). Two search codes were written for the auto-coding tool so that all the numerical data would be identified by the computer. The codes considered all the inflected forms of written numbers.

The second phase of coding was reserved for cleaning up the data. A researcher manually removed all unrelated numerical data that did not include performance information. Dates and numbers presenting years were removed if they did not relate to performance information. Because performance measures report what is happening, where and when it is happening and to what extent it is happening (Hatry, 2006), coding some of the dates and years was required, as these show when the performance occurred. Other numbers were occasionally removed, including expressions such as “first and foremost,” “on the one hand” and “one thing to be considered.”

In the third phase, the numerical data were further categorized according to different types of performance information (see Table 5). In the statistical testing phase, the total instances of performance information use were calculated, and the following statistics describing the performance information use were produced: mean (M), median (Mdn), mode (Z), range (RNG), minimum (Min), maximum (Max) and standard deviation (SD) (e.g. Krippendorff, 2004). This study uses these descriptive statistics to test the hypothesis that both presidents use performance information in their speeches. To test the second hypothesis, which claims that both presidents demonstrate limited performance information use in their speeches, this study uses descriptive statistics and the categorization of performance information use intensity depicted in Tables 6 and 7. When considering the categorization in Table 6, it is debatable how many speeches from the total number of speeches should include performance information when performance information use is, for example, high. Future studies can define these categories differently, but for communication purposes, Table 6 describes exactly what is meant in this study by high performance information use.

Table 5. What is being observed from the presidential speeches (the analytical framework)

<i>What was coded?</i>	<i>Conceptual definition</i>	<i>Coding example</i>
Use of numerical input information	Inputs are human, physical, and time resources consumed in the organizational activities (Cheema, 2006). Physical resources are assets, such as land, buildings, equipment, and raw materials (Bernolak, 1997)	“Public spending for 1999 had to be reduced by 5% annually.” (speech 32)
Use of numerical process information	Process information shows how activities are reduced, connected and merged, and how they work in tandem when the actual service process occurs (Rice, 2006)	“Some services more or less happen totally automatically, because people were refusing to push buttons ... when a baby is born and the doctor enters the data of the birth ... into the e-health system and in the background, without the doctor pushing one button, the digital identity is created for this baby ... originally we wanted the doctors to push one button and say: this way I will create the registry for this baby. But they refused ... We accepted. And therefore we created a system which automatically does it, in the background ... doctors even do not know that they are the civil registry managers.” (speech 8)
Use of numerical output information	Output information reports the quality and/or quantity of the goods and services produced by the public sector (Morley et al., 2001)	“There is one country, Japan, which can already manage to run trains not even one minute late.” (speech 20)
Use of numerical workload information	Workload information describes the workload coming into the organization. Measuring customer queues is typical for workload indicators, as they describe how many customers will be served by the organization in the near future (Hatry, 2006)	“So our people got really angry with the government, you know for what? Because ... some of them had to go to an office to update the certificates. Most of them could still do it online, but it did not function in the first two minutes; sometimes you had to do it nine times, because everyone was trying to update at the same time. No one could have such a wide channel open all the time so that 700,000 citizens could update at the same time.” (speech 2)
Use of numerical outcome information	Outcome information describes how conditions, events, attitudes, and behavior in society changed after the output was delivered (Morley et al., 2001)	“Since 1991, our GDP per capita has grown 20 times.” (speech 37)
Use of numerical productivity information	Productivity describes the relationship between inputs and outputs (Sumanth, 1994)	“Looking narrowly at tax collection, ... in 2004 we spent one euro to collect 100 euros and in 2013 0.4 to collect 100.” (speech 69)
Use of numerical cost-effectiveness information	Cost-effectiveness measures portray the outcomes achieved from the given inputs (Levin and McEwan, 2001)	“We have been able to offer more efficient public services, and the efficiency gains from digital signature are estimated to be as much as 2% of GDP per year. This 2% benefits mostly simple people and SMEs [small to medium enterprises], as neither has the capacity to handle big bureaucracy.” (speech 13)

Table 6. Categorization of performance information use intensity

<i>Performance information use is:</i>	<i>Values in terms of percentages of presidential speeches</i>
Non-existent	0%
Limited	0.1–25%
Moderate	25.1–50%
High	50.1–75%
Extensive	75.1–100%

Table 7. Second categorization of performance information use intensity

<i>The category created from the Likert five-point scale</i>	<i>The values in percentages (the relative proportion of numbers representing performance information from the whole word content in speeches)</i>
President <i>never</i> utilizes performance information in speeches	0%
President <i>seldom</i> utilizes performance information in speeches	0.1–0.18%
President <i>sometimes</i> utilizes performance information in speeches	0.19–0.37%
President <i>frequently</i> utilizes performance information in speeches	0.38–0.55%
President <i>very frequently</i> utilizes performance information in speeches	0.56–0.74%

This study develops a second method to compare the intensity of performance information use in Kaljulaid’s and Grybauskaitė’s speeches. The second method is based on estimates showing how frequently numbers are used in the English language. To create this estimate, the total number of English words was taken from the Oxford Dictionary, and a list of the most common words in English was downloaded. It was found that there are more than 1,000,000 words in the English language. A list of the 4,340 most common words (Russell, 2016) was used as a sample to estimate the ratio of numbers and other words in the population of 1,000,000 words. There were 32 numbers in the list, indicating that 0.74 percent of words (about seven out of every 1,000) are numbers. Unfortunately, the list did not include any frequency counts for the words, so a more accurate estimate could not be calculated. Based

on this estimation, a five-point Likert scale was constructed to assess the level of performance information use (see Table 7). The figure 0.74 was divided into four parts, each representing one category of the Likert scale. Because “never” means that there are no numbers presenting performance information in the speech, 0.74 was only divided by 4, as there are four categories left after excluding “never”: seldom, sometimes, frequently and very frequently. This method (Table 7) was developed so that one knows exactly what is meant when, for example, performance information is used “sometimes” in presidential speeches.

Presidents typically use less complex language (Hart, 1984) and presidential speeches are spoken in a way that seems familiar and banal to the listener (e.g. Edelman, 2013). Indeed, many public speaking guidebooks instruct speakers to use the language of their audience (Lundgren and McMakin, 2018). Therefore, it is expected that only 32 words out of 1,000 words are numbers in presidential speeches, reflecting typical English language use. Because numbers can raise distrust among citizens and numbers are often abstract by nature in presidential speeches (Stec, 2014), it is not realistic to assume that “more quality statistics simply and directly imply more persuasion” (Baele et al., 2018). Presidents knowing this would not try to maximize their performance information use for the sake of maximization.

To test the third hypothesis, which claims that both presidents most often use input information in their speeches, this study uses descriptive statistics illustrating used performance information. When the similarity of the presidents as performance information users is tested with the five hypotheses, the research setting utilizes comparisons of proportions and comparisons between groups. Two groups were generated – one for each president’s speeches. As differences were found between the presidents, correlations were used to show how the speech length was associated with performance information use. Based on these correlation tests, another set of group comparisons was conducted. In these group comparisons, the length of the speeches was controlled. As a second control, context was controlled and all the hypotheses were once again retested. Overall, the research process consisted of seven steps. These steps are summarized in Figure 1.

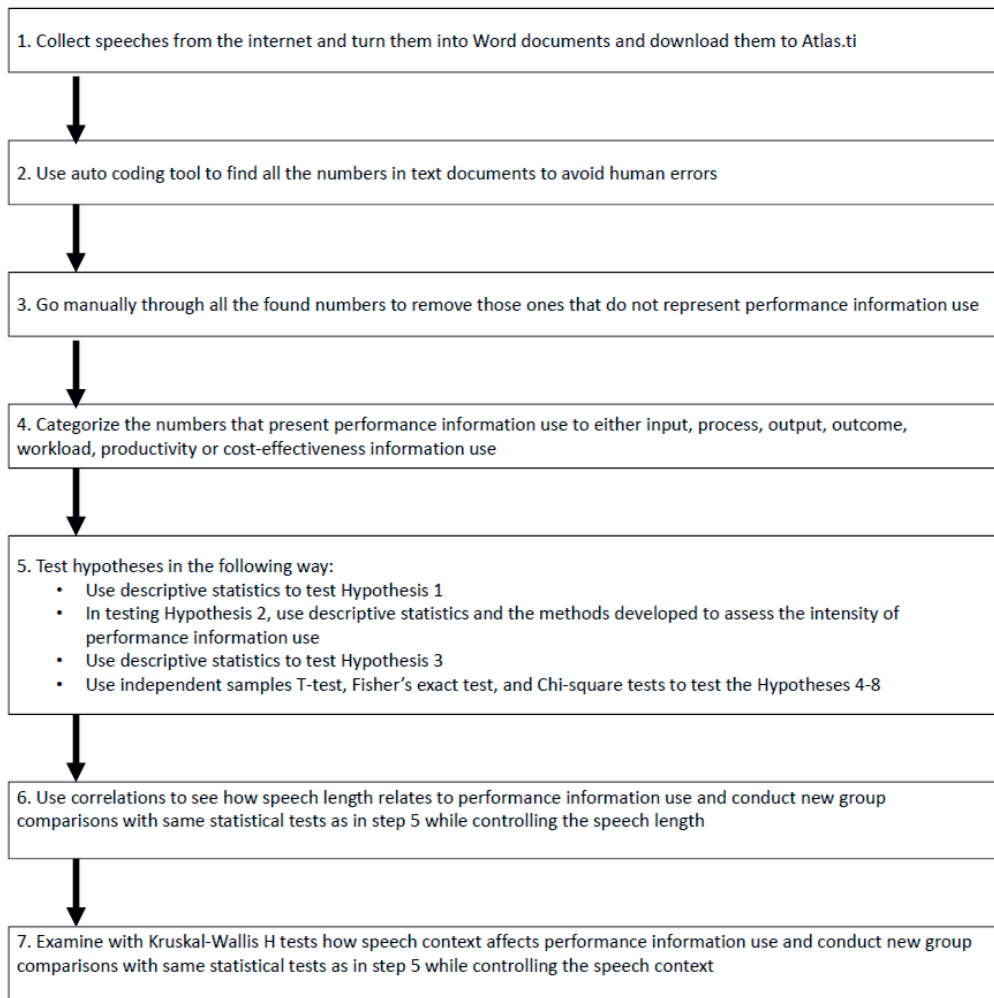


Figure 1. The research process

Describing performance information use in presidential speeches

The examined presidents did not use terms such as input, process, output, outcome, workload, productivity or cost-efficiency. They instead used performance measures in a less technocratic way. Quantitative performance information is often embedded in a story that describes the triumphs or failures of the government or foreign governments. The names of performance measures were not typically mentioned, and presidents did not state their sources of information, which is understandable as this would make the rhetoric appear clumsy. As an example of this type of performance information use, Lithuania's President Grybauskaitė stated in one of her speeches that “Women fill just 17 percent of the EU's IT [information technology] jobs. We must stop losing the world's brightest young minds”

(Speech 99). Grybauskaitė's quote was coded as outcome information according to our coding framework.

On some occasions, the presidents did use performance measures by name. Consider the following quotation from Estonia's President Kaljulaid, which was coded as cost-effectiveness information because it describes savings arising from outputs created with digital signing and benefits gained with lowered costs:

One visit which people had to undertake and sometimes wait for maybe one hour in the offices [...] Estonian people cannot tolerate this kind of relaxed attitude by the public sector that forces them to come to the office and queue at the office. [...] This actually means that this effect is measurable—it is 2% of the GDP just by signing digitally. We pay for our defense budget just from the savings we get from the digital environment. (Speech 2)

Table 8 provides the descriptive statistics of the quantitative content analysis. According to H1, both presidents use performance information. This hypothesis was confirmed as shown in Table 8. H2 claimed that both presidents demonstrate limited performance information in their speeches. However, the percentage of presidential speeches, including performance information and the relative proportion of numbers representing performance information from the whole word content in speeches in Table 8, shows extensive use, according to the interpretive framework shown in Table 6 and very frequent performance information use based on the scale presented in Table 7. For this reason, H2 is rejected.

H3 stated that both presidents most often use input information in their speeches. Table 9 summarizes how often a specific information type was used. The table demonstrates that outcome information is the most used performance measure in presidential speeches, as both presidents mainly applied this type of information. Thus, H3 is rejected, as the presidents' arguments seem to be outcome oriented. In fact, the presidents often talked about past and future achievements in terms of outcomes. The following quotations demonstrate the outcome orientation of the presidents:

In the 21st century, women must have equal legal status, equal pay, and equal political opportunities. And it is not only a matter of justice. Equal economic participation alone can add 28 trillion dollars to the global economy by 2025. (Grybauskaitė, Speech 99)

We exited the Soviet Union with an average salary of 30 dollars per month. It has now grown to almost 1,300 in 26 years. With the 30 dollars per month, we were very poor. (Kaljulaid, Speech 33)

Table 8. Key figures from the quantitative content analysis

Statistic	Kaljulaid (Estonia)	Grybauskaitė (Lithuania)
Total number of performance information uses in all speeches	1,164 (n = 85 speeches)	111 (n = 35 speeches)
Mean	14	3
Median	10	2
Mode	3	0
Min	0 (5 speeches i.e. 5.8% of all speeches)	0 (7 speeches i.e. 20% of all speeches)
Max	60	11
Range	60	11
Standard deviation	13	3
Percentage of presidential speeches including performance information	94%	80%
The relative proportion of numbers representing performance information from the whole word content in speeches	0.8%	0.7%

The outcome orientation is understandable because the presidents conduct economic leadership by addressing broad societal issues relating to their nations' well-being, and these issues are usually best described with outcome measures. Outcome measures typically capture different aspects of a nation's well-being better than input, process, output or workload information. Of course, cost-effectiveness measures could also be used, but there are currently more outcome measures available than cost-effectiveness measures. In addition, cost-effectiveness might also be more challenging to convey to ordinary people, and the presidents generally used language that is understandable by all.

Do the presidents use performance information similarly?

H4 argued that both presidents use different amounts of performance information in their speeches. Table 8 shows the total instances of performance information use in all speeches. Based on this information, Kaljulaid used more performance information during the examined time interval. However, Kaljulaid also had more speeches (n = 85) than Grybauskaitė (n = 35), which could explain

the difference between the two presidents. By looking at the average number of performance information uses per speech (i.e. the mean), it becomes evident that Kaljulaid used more performance information on average in her speeches (14 times per speech) compared to Grybauskaitė (3 times per speech).

An independent sample t-test was conducted to determine whether the mean values of the performance information are the same in the two groups (Kaljulaid in Group 1, Grybauskaitė in Group 2). As shown in Table 10, there was a statistically significant difference between Kaljulaid's and Grybauskaitė's performance information use, $t(104,06) = 7.003$, $p = 0.000$. Levene's test did not hold, and therefore equal variances were not assumed. The independent sample t-test revealed that Kaljulaid's speeches were significantly different from Grybauskaitė's speeches in terms of the average performance information per speech. Thus, H4, which claims that performance information use is different on average in Kaljulaid's and Grybauskaitė's speeches, is accepted.

Table 9. Types of performance measures used in speeches and their usage (see also Appendix 2)

Measures	<i>Kaljulaid</i>		<i>Grybauskaitė</i>		<i>Totals</i>
	Number of uses	%	Number of uses	%	
Cost-efficiency	13	1%	0	0%	13
Input	64	5%	4	4%	68
Outcome	843	72%	106	95%	949
Output	218	19%	1	1%	219
Process	8	1%	0	0%	8
Productivity	17	2%	0	0%	17
Workload	1	0%	0	0%	1
Totals	1,164	100%	111	100%	1,275

H5 proposed that the performance information use of the presidents varies differently between speeches. An examination of the minimum and maximum values revealed that the presidential performance information use fluctuates significantly between speeches. The minimum and maximum values in Table 8 demonstrate that Kaljulaid used performance information 60 times in one of her speeches, while five speeches did not include any performance information. Therefore, the range was 60. The fluctuation in performance information use was not so drastic in the speeches of Grybauskaitė, who used performance information 11 times in one of her speeches, while seven of her speeches did not include any performance information. The range here is therefore 11. Looking at the standard deviation confirms the findings of the stronger fluctuation in Kaljulaid's speeches compared to Grybauskaitė's. This fluctuation is significantly different, as shown by Levene's test for the equality of variances in Table 10. Therefore, H5 is accepted.

Table 10. Independent sample t-test

		Levene's test for the equality of variances		t-test for the equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									Lower	Upper
# Information use	Equal variances assumed	26.607	.000	4.730	118	.000	10.523	2.225	6.117	14.928
	Equal variances not assumed			7.003	104,1	.000	10.523	1.503	7.543	13.502

H6 suggested that the proportion of speeches including performance information is not equal when Kaljulaid's speeches are compared to Grybauskaitė's speeches. The ratio of speeches incorporating performance information and those that do not was significantly higher in Kaljulaid's speeches compared to the similar ratio number calculated from Grybauskaitė's speeches ($p < 0.05$, Fisher's exact test). Thus, H6 is accepted. The Fisher's exact test is used here because there is one cell that has an expected count of less than 5, and the data form a 2x2 nominal table.

According to H7, both presidents do not have the same ratio of performance information and other content in their speeches. The empirical data indicate that the presidents are not significantly different in their performance information use if one compares the relative proportions of numbers representing performance information from the whole word content in speeches. As shown in Table 8, we found 1,164 numbers (or words, as word processing tools see numbers) that described performance information in Kaljulaid's speeches, and 111 such numbers were present in Grybauskaitė's speeches.

In Table 8, the relative proportion of numbers representing performance information from the whole word content in speeches can be used to conduct comparisons between the presidents. Because the relative proportion of numbers representing performance information from the whole word content in speeches is 0.8 percent in Kaljulaid's speeches and 0.7 percent in Grybauskaitė's speeches, it seems that the balance between performance information and other information within the speeches is similar, although Kaljulaid used more performance information per speech on average. Kaljulaid used more

words in her speeches, so the balance between performance information and other information is the same when compared to Grybauskaitė, who used performance information fewer times per speech but whose speeches were also much shorter on average. The average length of Kaljulaid's speeches was 1,698 (measured by mean) or 1,456 (measured by median) words, and the average length of Grybauskaitė's speeches was 453 (mean) or 391 (median). A χ^2 test confirmed that Kaljulaid's and Grybauskaitė's performance information use habits are similar when the ratio of performance information and other content is examined ($p = 0.796$, χ^2 test). H7 is therefore rejected.

H8 projected that both presidents do not use the same types of performance measures equally often. According to Table 8, the types of measures used by Kaljulaid differed significantly from the performance measures used by Grybauskaitė ($p < 0.01$, χ^2 test). Thus, H8 is accepted. Table 8 describes the types of performance measures used and what measures were used more than other measures. The percentages in Columns 3 and 5 describe the relative proportions of the used measures.

Would the performance information use habits be similar if the presidents gave equally long speeches? Both Grybauskaitė's (Spearman $r = 0.436$, $p < 0.01$) and Kaljulaid's (Spearman $r = 0.648$, $p < 0.01$) performance information use correlated positively with speech length. By using 35 of Kaljulaid's speeches that were similar in length to Grybauskaitė's speeches, this study retested H4–H8. The means were significantly different ($p < 0.05$), and therefore, H4 was again accepted. However, an equality of variances was found ($p = 0.31$), and H5 was rejected. Moreover, the proportions of speeches including performance information use were equal ($p = 0.526$, χ^2 test), and the ratios of performance information and other content in the speeches were similar ($p = 0.781$, χ^2 test). Thus, H6 and H7 were rejected. After controlling for speech length, the types of measures used by Kaljulaid differed significantly from the performance measures used by Grybauskaitė ($p < 0.01$, χ^2 test) and H8 was accepted.

While controlling the speech length, H1–H3 were also retested. H1 was accepted as performance information was used by Kaljulaid 175 times and by Grybauskaitė 111 times. H2 was rejected as the presidents used performance information extensively and frequently (see Tables 6 and 7). From all the words, 0.7 percent were numerical performance information in Kaljulaid's and Grybauskaitė's speeches. In total, 85 percent of Kaljulaid's speeches and 80 percent of Grybauskaitė's speeches incorporated performance information. H3 was rejected because performance information was mostly outcome information (78.8 percent of performance information concerned outcomes in Kaljulaid's speeches and 95 percent in Grybauskaitė's speeches).

A Kruskal–Wallis H test was run to determine if there were differences in four different speech contexts in terms of distributions of speech lengths. These contexts were the ceremonial ($n = 59$), conference ($n = 41$), lecture ($n = 5$) and official speech ($n = 15$) speech groups (see Table 4). Both Kaljulaid's and Grybauskaitė's speeches were divided into these categories. Distributions of speech lengths were not similar for all contexts, as assessed by visual inspection of a boxplot. The mean ranks of speech lengths

were statistically significantly different between contexts, $\chi^2(3) = 47.935$, $p < 0.001$. Subsequently, pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. This post hoc analysis revealed statistically significant differences between the contexts as demonstrated in Table 11. Moreover, conference (mean = 1672) and lecture speeches (mean = 3582) involved more words than ceremonial speeches (mean = 697) and official speeches (mean = 872) on average.

Table 11. Pairwise comparisons between the different contexts and distributions of speech length within this context

<i>Sample 1 – Sample 2</i>	<i>Test statistic</i>	<i>Adj. Sig.</i>
Ceremonial – official speech	-13.9	1.000
Ceremonial – conference	-42.0	0.000
Ceremonial – lecture	-73.0	0.000
Conference – official speech	28.1	0.044
Lecture – official speech	59.1	0.066
Conference – lecture	-31.0	0.358

The first Kruskal–Wallis H test was replicated by using only Kaljulaid's speeches, and the results were similar to those in Table 11. Distributions of speech lengths in conference and lecture speeches differed from the distributions of speech lengths in ceremonial and official speeches. On average, Kaljulaid's lectures (mean = 3,582) were longer than conference speeches (mean = 1,802), which in turn were longer than official speeches (mean = 1,005). Official speeches were marginally lengthier than ceremonial speeches (mean = 1,004), according to mean values. Grybauskaitė's speech lengths were on average: ceremonial (mean = 309), conference (mean = 467), lecture (no lectures given) and official speech (mean = 607).

Another Kruskal–Wallis H test was run to determine if there were differences in four different speech contexts in terms of distributions of performance information use. Both Kaljulaid's and Grybauskaitė's speeches were divided into the four speech categories once again. Distributions of performance information use were not similar for all contexts, as assessed by visual inspection of a boxplot. The mean ranks of performance information use were statistically significantly different between contexts, $\chi^2(3) = 47.016$, $p < 0.001$. The pairwise comparison in Table 12 demonstrates that lectures and conference speeches are statistically significantly different from ceremonial and official speeches when speeches of both presidents are divided into categories based on different contexts. Similar results arise when the Kruskal–Wallis H test is run with the speeches of Kaljulaid. When the Kruskal–Wallis H test was used to assess Grybauskaitė's speeches in terms of performance information use, conference

speeches are statistically significantly different only from ceremonial speeches, not from official speeches.

Table 12. Pairwise comparisons between the different contexts and distributions of performance information use within the context

<i>Sample 1 – Sample 2</i>	<i>Test statistic</i>	<i>Adj. Sig.</i>
Ceremonial – official speech	-8.0	1.000
Ceremonial – conference	-42.2	0.000
Ceremonial – lecture	-67.1	0.000
Conference – official speech	34.2	0.007
Lecture – official speech	59.1	0.006
Conference – lecture	-24.9	0.777

To control the speech context, all the hypotheses were retested by using only the ceremonial speeches of the presidents. H1 was accepted as it predicted performance information use, as demonstrated in Kaljulaid using performance information 219 times and Grybauskaitė 65 times. H2 predicting limited use was rejected as both presidents demonstrated either extensive or high use and very frequent utilizations of performance information (see Tables 6 and 7). From Kaljulaid’s ceremonial speeches, 87.7 percent included performance information and 0.66 percent of all words were numerical performance information, whereas for Grybauskaitė, the findings were 73 and 0.8 percent, respectively. As both presidents used outcome information most frequently in their ceremonial speeches, H3 claiming that input information is used the most was rejected. From the performance information used, 80.8 percent was outcome information in Kaljulaid’s speeches and 93.8 percent in Grybauskaitė’s speeches.

There was a statistically significant difference between Kaljulaid’s and Grybauskaitė’s performance information use in ceremonial speeches, $t(51.5) = 7.003$, $p = 0.000$. Thus, H4 was accepted. The Levene’s test did not hold ($F = 9.7$, $p = 0.003$), and therefore, equal variances were not assumed. The Levene’s test indicates that the performance information use of the presidents varies differently between the different speeches, meaning that that H5 was accepted. H6 was accepted as the ratio of speeches incorporating performance information and those that do not were significantly higher in Kaljulaid’s speeches compared to the similar ratio number calculated from Grybauskaitė’s speeches ($p < 0.01$, χ^2 test). A χ^2 test confirmed that Kaljulaid’s and Grybauskaitė’s performance information use habits are similar when the ratio of performance information and other content is examined in the context of ceremonial speeches ($p = 0.199$, χ^2 test). As a result, H7 was rejected. H8 was accepted because the

types of measures used by Kaljulaid differed significantly from the performance measures used by Grybauskaitė ($p < 0.01$, χ^2 test). Overall, Table 13 summarizes the statistical tests and their results.

Table 13. Summary of the research results

<i>Research question</i>	<i>Hypothesis</i>	<i>Test</i>	<i>Retest with speech length controlled</i>	<i>Retest with context controlled</i>
Is performance information actually used by presidents?	1. Both presidents use performance information in their speeches.	Hypothesis accepted	Hypothesis accepted	Hypothesis accepted
To what extent are presidents using performance information?	2. Both presidents demonstrate limited performance information use in their speeches.	Hypothesis rejected (extensive and frequent use was found)	Hypothesis rejected (extensive and frequent use was found)	Hypothesis rejected (extensive, high and frequent use was found)
What type of performance information is being used by the presidents?	3. Both presidents most often use input information in their speeches.	Hypothesis rejected (outcome information was most used)	Hypothesis rejected (outcome information was most used)	Hypothesis rejected (outcome information was most used)
	4. In their speeches, both presidents do not use the same amount of performance information on average.	Hypothesis accepted	Hypothesis accepted	Hypothesis accepted
	5. The performance information use of the presidents varies differently between the different speeches.	Hypothesis accepted	Hypothesis rejected	Hypothesis accepted
Do some presidents make more use of performance information than others?	6. Proportions of speeches including performance information use are not equal when Kaljulaid's speeches are compared to Grybauskaitė's.	Hypothesis accepted	Hypothesis rejected	Hypothesis accepted
	7. Presidents do not have the same ratio of performance information and other content in their speeches.	Hypothesis rejected	Hypothesis rejected	Hypothesis rejected
	8. Presidents do not use the same types of performance measures equally often.	Hypothesis accepted	Hypothesis accepted	Hypothesis accepted

Discussion and conclusion

This study examined the numerical performance information use in presidential speeches and found extensive and frequent performance information use, with a prevalence of outcome information. This supports the results of scholars who have argued that politicians use performance information (e.g. Askim, 2007; Kroll and Proeller, 2013; Giacomini et al., 2016). The intensive use found also supports researchers who have reported politicians as active performance information users (Ezzamel et al., 2004; Saliterer and Korac, 2013; Charbonneau and Bellavance, 2015). The outcome orientation found in this study contradicts prior results stating that the public sector is oriented more toward input and output than outcome, as the measures mostly focus on inputs and outputs even though outcomes are seen as important (e.g. Ferlie et al., 2005; Moynihan, 2005; Sterck and Scheers, 2006; Ozturk and Swiss, 2008). However, a high use of outcome information has also been reported (e.g. Saliterer and Korac, 2014; Hijal-Moghrabi, 2017). In this study, the use of outcome information is not perhaps so surprising as outcome information describing the state of the nation in terms of well-being and economic growth is very useful when presidents conduct economic leadership to guide the behaviors of consumers and companies. Indeed, leading the macroeconomic performance often requires information about macroeconomic outcomes as these may be more relevant compared to outputs or inputs associated with certain programs and services. This topic requires further investigation.

The extensive and frequent use of performance information was associated with presidents who are middle-aged females and who had financial management and economics backgrounds. Both presidents also had similar powers and operated in political contexts that were very alike. These associations between performance information use and contextual and individual characteristics both support and contradict some of the findings of Askim (2009). Askim (2009) claimed that political conflict increases performance information use, which was supported by the findings of this study, as both Estonia and Lithuania have political tensions with Russia. There is also some tension between Russians and Lithuanians in Lithuania and Russians and Estonians in Estonia (see also Giacomini et al., 2016). Conflicts also arise in these countries because the style of national and local politics is also often confrontational in Estonia, while the three governing parties in Lithuania have very different interests.

Askim (2009) also found that higher education and age are associated with low use, and this finding was not supported by the results of this study showing that frequent and extensive use was associated to high age and education levels. One finding arising from the empirical analysis showed that presidents with employment backgrounds in financial management and economics frequently used performance information. However, this finding does not prove causality as there is no control group of presidents with different employment backgrounds. It only shows that in the studied cases, there were associations between performance information use and backgrounds in financial management and economics. The reader should take this result relating to backgrounds in financial management and economics as a description of how things were in the two countries, not as a proof of causality. Proving causality would

require another type of research setting and the results of this study only offer hypotheses to studies aiming to capture the causality. The ways different backgrounds contribute to performance information use is currently an underexamined topic that should be studied more based upon the preliminary findings shown here.

The two presidents also differed from each other in terms of performance information use, although they had many similarities in their individual backgrounds and operating contexts. This suggests that the level of similarity was not adequate, as the presidents were also different in many ways as described in the research context. These differences in individual characteristics can relate to the dissimilarities found in performance information use, as age, education, occupational background and political experience can all explain different use patterns (ter Bogt, 2004; Askim, 2007; Saliterer and Korac, 2014). Older age and higher education, different career paths and longer political career were associated with lower use in this study. Also, differences in socio-demographic factors, political situation and economic development can explain the different performance information use habits of the presidents (e.g. OECD, 2016).

This study also found that there can be significant differences within seemingly similar performance information use that is labeled as extensive use. Some of these differences vanished when the length of speech was controlled, but some differences remained. The study found correlations between speech length and performance information use in both presidents' speeches. This is natural, as the ratio between numbers and other words in the English language is fairly constant as this study demonstrated by sampling the most common words used in the English language and calculating this ratio between numbers and other words. As presidents try to speak the language of the ordinary people, it was here assumed that they also use numbers in a manner similar to the way other people generally do. The literature has not previously examined how speech length is associated with performance information use intensity (e.g. Buylens and Christiaens, 2016) and this requires further investigation.

Performance information use varied to some extent between different information users and speeches made in dissimilar contexts and provided more evidence for the already established notion that performance information use can fluctuate significantly depending on the use context and users (e.g. Johansson and Siverbo, 2009; Charbonneau and Bellavance, 2015). The study found that some speech contexts stimulated more use than others when presidents were examined individually, but the differences in performance information use between the presidents did not relate to differences in speech contexts as there were differences in use, although the speech context was controlled. Overall, this study indicates that theoretical models of performance information use should consider speech length and contexts as important variables that affect use.

The findings of this research expand upon current theories of performance information users by identifying a new political user group: presidents (e.g., van Helden and Reichard, 2019). Past studies

have focused on local councilors, state legislators and members of parliament or congressional representatives (Rhee, 2014; Lu and Willoughby, 2015; Grossi et al., 2016). Future research should study more on how presidents use performance information and compare their results to this study.

The research methods developed in this study may benefit future research by providing conceptual frameworks (see Tables 6 and 7) that can be used to assess the intensity of performance information use in rhetoric. The readers should keep in mind that the offered interpretation frameworks for performance information use intensity are not meant to be normative and they are very much debatable, but they should be thought of as conversation starters on how to define intensity of use. Indeed, the proposed methods invite scholars to discuss how high and low information use should be defined and measured, as the current research provides very few clues as to how high or low performance information use manifests itself in rhetoric (see e.g. Sterck and Scheers, 2006; Askim, 2007; Raudla, 2012; Buylen and Christiaens, 2016; Giacomini et al., 2016). Moreover, terms such as limited use or high use are vaguely defined in the literature (e.g. Ospina et al., 2004; Buylen and Christiaens, 2016). The proposed frameworks contribute conceptually by defining what is intensive performance information use, while they also provide a critique to the vaguely defined concepts of intensity used in the past literature. Currently, it is not clear whether Raudla's (2012) use of the term limited use means the same thing as in ter Bogt (2004). As interpretation depends on concepts that ensure their communicability, this is a problem in the current research.

Although the conceptual frameworks proposed here for intensity of use are debatable, this study has provided measurement techniques that make various usage levels of performance information more concrete and transparent. With these methods, future research can examine different aspects of performance information use in rhetoric, including what information is being used in speeches and why and how it is being used.

The second analytical model developed in this study used H4–H8 to create a comparative model that can be used to compare two speakers and their performance information use. The comparisons can be done with t-tests, Fisher's exact tests, χ^2 tests and Kruskal–Wallis H tests. As a theoretical contribution, the developed analytical models suggest that the intensity of performance information can be assessed from several dimensions. One can analyze how many speeches include performance information to determine intensity (see Table 6). One can also examine how much of the content of a speech is performance information (see Table 7). It is also possible to examine whether some types of performance information are used more intensely than others, as was the case in this study. Although one would not agree on the conceptual framework proposed in this study defining what high use is, the analytical model is still useful. For example, dividing the amount of words that are numbers representing performance with the total number of words used by the presidents in their speeches can be done in future studies, and the numbers can be compared to the ones seen in this study even if one does not

agree how frequent use was defined in this research. The numbers are comparable even if we disagree about the labels, such as frequent use or seldom used, given to these numbers.

The main limitations of this study are that it only focused on numerical information. This means that the actual use observed was underestimated. Future studies could use grounded theory and other qualitative settings to examine the performance information use of presidents. Another limitation was the small sample of just two presidents. Because of the sample size, only preliminary findings and suggestions for future research topics could be offered. The third limitation relates to the interpretative framework used in this study, which can be debatable as the previous studies have not established what high use means in practice. Therefore, the reader should look at the numbers and consider their interpretation.

The analysis of performance information in rhetoric is an underused research method. By studying presidential performance information use and its intensity in speeches with new methods, this study has added a new dimension to the existing research (e.g. Saliterer and Korac, 2014; Lu and Willoughby, 2015; Grossi et al., 2016). The findings of this study implicate that there is a lot of interesting theoretical and empirical work to be done on topics addressing performance information in rhetoric.

References

- Askim, J. (2007), "How do politicians use performance information? An analysis of the Norwegian local government experience", *International Review of Administrative Sciences*, Vol. 73 No. 3, pp. 453-472.
- Askim, J. (2009), "The demand side of performance measurement: explaining councilors' utilization of performance information in policymaking", *International Public Management Journal*, Vol. 12 No. 1, pp. 24-47.
- Auers, D. (2015), *Comparative Politics and Government of the Baltic States: Estonia, Latvia and Lithuania in the 21st Century*, Palgrave Macmillan, New York, NY.
- Baele, S.J., Coan, T.G. and Sterck, O.C. (2018), "Security through numbers? Experimentally assessing the impact of numerical arguments in security communication", *The British Journal of Politics and International Relations*, Vol. 20 No. 2, pp. 459-476.
- Behr, R. and Iyengar, S. (1985), "Television news, real-world cues, and changes in the public agenda", *Public Opinion Quarterly*, Vol. 49 No. 1, pp. 38-57.
- Berger, A. (1995), *Essentials of Mass Communication Theory*, Sage, CA.
- Bernolak, I. (1997), "Effective measurement and successful elements of company productivity: the basis of competitiveness and world prosperity", *International Journal of Production Economics*, Vol. 52 Nos 1-2, pp. 203-213.

Bjørnholt, B., Bækgaard, M. and Houlberg, K. (2016), "Does fiscal austerity affect political decision-makers' use and perception of performance information?", *Public Performance and Management Review*, Vol. 39 No. 3, pp. 560-580.

Bogan, C. and English, M. (1994), *Benchmarking for Best Practices: Winning through Innovative Adaptation*, McGraw-Hill, New York, NY.

Bourdeaux, C. (2006), "Do legislatures matter in budgetary reform?", *Public Budgeting & Finance*, Vol. 26, pp. 120-142, doi: 10.1111/j.1540-5850.2006.00841.x.

Buylen, B. and Christiaens, J. (2016), "Talking numbers? Analyzing the presence of financial information in councilors' speech during the budget debate in Flemish municipal councils", *International Public Management Journal*, Vol. 19 No. 4, pp. 453-475.

Charbonneau, É. and Bellavance, F. (2015), "Performance management in a benchmarking regime: Quebec's municipal management indicators", *Canadian Public Administration*, Vol. 58 No. 1, pp. 110-137.

Cheema, D. (2005), *Operations Research (for B.Com., B.Tech., BBA, MBA, BCA, MCA and other classes)*, Laxmi Publications, Bangalore. Chen, H. and Smith, S.H. (2019), "School board directors' information needs and financial reporting's role", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 578-595, available at: <https://doi.org/10.1108/JPBAFM-09-2018-0097>

Cohen, J. (1995), "Presidential rhetoric and the public agenda", *American Journal of Political Science*, Vol. 39 No. 1, pp. 87-107.

Cunningham, G.M. and Harris, J.E. (2005), "Toward a theory of performance reporting to achieve public sector accountability: a field study", *Public Budgeting & Finance*, Vol. 25, pp. 15-42, doi: 10.1111/j.0275-1100.2005.00359.x.

Dunn, O.J. (1964), "Multiple comparisons using rank sums", *Technometrics*, Vol. 6, pp. 241-252. Edelman, M. (2013), *Political Language: Words that Succeed and Policies that Fail*, Academic Press, New York, NY.

Edwards, G. (2003), *On Deaf Ears: The Limits of the Bully Pulpit*, Yale University Press, New Haven, CT. Ellul, L. and Hodges, R. (2019), "Reforming the government budgeting system in Malta: pressures promoting or hindering the reform", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 518-538, available at: <https://doi.org/10.1108/JPBAFM-10-2018-0108>

Erikson, E. (1982), *The Life Cycle Completed*, Norton, New York, NY.

Ezzamel, M., Hyndman, N.S., Johnsen, Å., Lapsley, I. and Pallot, J. (2004), "Has devolution increased democratic accountability?", *Public Money and Management*, Vol. 24 No. 3, pp. 145-152.

Ferlie, E., Pollitt, C. and Lynn, L.E. (2005), *The Oxford Handbook of Public Management*, Oxford University Press, Oxford.

Gerring, J. (2007), "The case study: what it is and what it does", in Boix, C. and Stokes, S. (Eds), *Oxford Handbook of Comparative Politics*, Oxford University Press, New York, NY, pp. 90-122.

Giacomini, D., Sicilia, M. and Steccolini, I. (2016), "Contextualizing politicians' uses of accounting information: reassurance and ammunition", *Public Money and Management*, Vol. 36 No. 7, pp. 483-490.

Grossi, G., Reichard, C. and Ruggiero, P. (2016), "Appropriateness and use of performance information in the budgeting process: some experiences from German and Italian municipalities", *Public Performance and Management Review*, Vol. 39 No. 3, pp. 581-606.

Guarini, E. (2016), "The day after: newly-elected politicians and the use of accounting information", *Public Money and Management*, Vol. 36 No. 7, pp. 499-506. Hart, R. (1984), *Verbal Style and the American Presidency*, Academic Press, New York, NY.

Hatry, H. (2006), *Performance Measurement: Getting Results*, The Urban Institute, Washington, DC.

Haustein, E., Lorson, P.C., Caperchione, E. and Brusca, I. (2019), "The quest for users' needs in public sector budgeting and reporting", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 473-477, available at: <https://doi.org/10.1108/JPBAFM-08-2019-0131>

Hijal-Moghrabi, I. (2017), "The current practice of performance-based budgeting in the largest US cities: an innovation theory perspective", *Public Performance and Management Review*, Vol. 40 No. 4, pp. 652-675.

Ho, A. (2005), "Accounting for the value of performance measurement from the perspective of Midwestern mayors", *Journal of Public Administration Research and Theory*, Vol. 16 No. 2, pp. 217-237.

Johanson, U., Almqvist, R. and Skoog, M. (2019), "A conceptual framework for integrated performance management systems", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 3, pp. 309-324.

Johansson, T. and Siverbo, S. (2009), "Explaining the utilization of relative performance evaluation in local government: a multi-theoretical study using data from Sweden", *Financial Accountability and Management*, Vol. 25 No. 2, pp. 197-224.

Jorge, S., Jesus, M.A. and Nogueira, S. (2016), "Information brokers and the use of budgetary and financial information by politicians: the case of Portugal", *Public Money and Management*, Vol. 36 No. 8, pp. 515-520.

- Jorge, S., Jesus, M. and Nogueira, S. (2019), "The use of budgetary and financial information by politicians in parliament: a case study", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 539-557, available at: <https://doi.org/10.1108/JPBAFM-11-2018-0135>
- Joyce, P. (2011), "The Obama administration and PBB: building on the legacy of federal performance-informed budgeting?", *Public Administration Review*, Vol. 71 No. 3, pp. 356-367.
- Kenk, K. and Haldma, T. (2019), "The use of performance information in local government mergers", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 3, pp. 451-471, available at: <https://doi.org/10.1108/JPBAFM-03-2019-0056>
- Krippendorff, K. (2004), *Content Analysis: An Introduction to its Methodology*, Sage Publications, Thousand Oaks, CA.
- Kroll, A. and Proeller, I. (2013), "Controlling the control system: performance information in the German childcare administration", *International Journal of Public Sector Management*, Vol. 26 No. 1, pp. 74-85.
- Levin, H. and McEwan, P. (2001), *Cost-Effectiveness Analysis: Methods and Applications*, Sage Publications, Thousand Oaks, CA.
- Linz, J. (1990), "The perils of presidentialism", *Journal of Democracy*, Vol. 1 No. 1, pp. 51-69.
- Lu, E. and Willoughby, K. (2015), "Performance budgeting in American states: a framework of integrating performance with budgeting", *International Journal of Public Administration*, Vol. 38 No. 8, pp. 562-572.
- Lundgren, R. and McMakin, A. (2018), *Risk Communication: A Handbook for Communicating Environmental, Safety, and Health Risks*, John Wiley & Sons, Columbus, OH.
- MacKuen, M. (1983), "Political drama, economic conditions, and the dynamics of presidential popularity", *American Journal of Political Science*, Vol. 27 No. 2, pp. 165-192.
- Morley, E., Bryant, S. and Hatry, H. (2001), *Comparative Performance Measurement*, Urban Institute Press, Washington, DC.
- Moynihan, D. (2005), "Why and how do state governments adopt and implement 'managing for results' reforms?", *Journal of Public Administration Research and Theory*, Vol. 15 No. 2, pp. 219-243.
- Nakrošis, V. (2008), "Reforming performance management in Lithuania: towards results-based government", in Peters, B.G. (Ed.), *Mixes, Matches, and Mistakes: New 152 Public Management in Russia and the Former Soviet Republics*, Local Government and Public Service Reform Initiative, Open Society Institute, Budapest, pp. 53-115.
- OECD (2016), *Factbook 2015–2016*, Vol. 10, OECD Publishing, Paris.

Ospina, S., Cunill Grau, N. and Zaltsman, A. (2004), "Performance evaluation, public management improvement and democratic accountability: some lessons from Latin America", *Public Management Review*, Vol. 6 No. 2, pp. 229-251.

Ostrom, C. and Smith, R. (1993), "Error correction, attitude persistence, and executive rewards and punishments: a behavioral theory of presidential approval", *Political Analysis*, Vol. 4, pp. 127-184.

Ouda, H. and Klischewski, R. (2019), "Accounting and politicians: a theory of accounting information usefulness", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 496-517, available at: <https://doi.org/10.1108/JPBAFM-10-2018-0113>

Ozturk, A. and Swiss, J. (2008), "Implementing management tools in Turkish public hospitals: the impact of culture, politics and role status", *Public Administration and Development: The International Journal of Management Research and Practice*, Vol. 28 No. 2, pp. 138-148.

Pollitt, C. (2006), "Performance information for democracy: the missing link?", *Evaluation*, Vol. 12 No. 1, pp. 38-55.

Pollitt, C. and Bouckaert, G. (2004), *Public Management Reform. A Comparative Analysis*, Oxford University Press, Oxford.

Posner, P. and Fantone, D.M. (2007), "Assessing federal program performance: observations on the US office of management and budget's program assessment rating tool and its use in the budget process", *Public Performance and Management Review*, Vol. 30 No. 3, pp. 351-368.

President of Estonia (2018), "Powers and responsibilities", available at: www.president.ee/en/president/legal-authority/index.html

President of Lithuania (2018), "Legal framework", available at: <http://adamkus.president.lt/en/institutiontoday.phtml>

Raudla, R. (2012), "The use of performance information in budgetary decision-making by legislators: is Estonia any different?", *Public Administration*, Vol. 90 No. 4, pp. 1000-1015.

Raudla, R. and Savi, R. (2015), "The use of performance information in cutback budgeting", *Public Money and Management*, Vol. 35 No. 6, pp. 409-416. Rhee, D. (2014), "The impact of performance information on congressional appropriations", *Public Performance and Management Review*, Vol. 38 No. 1, pp. 100-124.

Rice, R. (2006), *Home Care Nursing Practice: Concepts and Application*, Mosby Elsevier, St Louis, MO.

Russell, R. (2016), "4000 most common English words", available at: www.rupert.id.au/resources/1000words.php (accessed June 17, 2019).

Saliterer, I. and Korac, S. (2013), "Performance information use by politicians and public managers for internal control and external accountability purposes", *Critical Perspectives on Accounting*, Vol. 24 Nos 7–8, pp. 502-517.

Saliterer, I. and Korac, S. (2014), "The discretionary use of performance information by different local government actors: analysing and comparing the predictive power of three factor sets", *International Review of Administrative Sciences*, Vol. 80 No. 3, pp. 637-658.

Sinervo, L. and Haapala, P. (2019), "Presence of financial information in local politicians' speech", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 558-577, available at: <https://doi.org/10.1108/JPBAFM-11-2018-0133>

Stec, F.J. (2014), *What's in a Number?: The Rhetoric of Numbers During Wartime*, The Pennsylvania State University, State College, PA.

Sterck, M. and Scheers, B. (2006), "Trends in performance budgeting in seven OECD countries", *Public Performance and Management Review*, Vol. 30 No. 1, pp. 47-72.

Sumanth, D. (1994), *Productivity Engineering and Management*, McGraw-Hill, New York, NY.

ter Bogt, H. (2001), "Politicians and output-oriented performance evaluation in municipalities", *European Accounting Review*, Vol. 10 No. 3, pp. 621-643.

ter Bogt, H. (2004), "Politicians in search of performance information? Survey research on Dutch aldermen's use of performance information", *Financial Accountability and Management*, Vol. 20 No. 3, pp. 221-252.

van Helden, J. (2016), "Literature review and challenging research agenda on politicians' use of accounting information", *Public Money and Management*, Vol. 36 No. 7, pp. 531-538.

van Helden, J. and Reichard, C. (2019), "Making sense of the users of public sector accounting information and their needs", *Journal of Public Budgeting, Accounting & Financial Management*, Vol. 31 No. 4, pp. 478-495, available at: <https://doi.org/10.1108/JPBAFM-10-2018-0124>

van Thiel, S. and Leeuw, F. (2002), "The performance paradox in the public sector", *Public Performance & Management Review*, Vol. 25 No. 3, pp. 267-281.

Wang, X. (2008), "Convincing legislators with performance measures", *International Journal of Public Administration*, Vol. 31 No. 6, pp. 654-667.

Wood, B. (2004), "Presidential rhetoric and economic leadership", *Presidential Studies Quarterly*, Vol. 34 No. 3, pp. 573-606.

Wood, B.D., Owens, C.T. and Durham, B.M. (2005), "Presidential rhetoric and the economy", *Journal of Politics*, Vol. 67, pp. 627-645.

Appendix 1

Words searched from the speeches were: 0–9, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, hundred, thousand, million, billion, trillion, dozen, couple, gross, quantity, half, double, first, second, third, fourth, fifth, sixth, seventh, eighth, ninth, tenth, eleventh, twelfth, thirteenth, fourteenth, fifteenth, sixteenth, seventeenth, eighteenth, nineteenth and twentieth.

Appendix 2.

In how many speeches was the measure used

<i>Measure</i>	<i>Kaljulaid</i>	<i>Grybauskaitė</i>
cost-efficiency	9	0
input	22	4
outcome	80	27
output	56	1
process	5	0
productivity	10	0
workload	1	0

