

Tallinn University of Technology Doctoral Theses
Series I: Social Sciences, No. 25

**The Attribution Problem in
Performance Measurement in the Public Sector:
Lessons from Performance Audits in Estonia**

KÜLLI TARO (NÕMM)

TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Social Sciences
Ragnar Nurkse School of Innovation and Governance

This dissertation was accepted for the defense of the degree of Doctor of Philosophy in Public Administration on November 23, 2015.

Supervisor: Professor Dr. Tiina Randma-Liiv, Tallinn University of Technology,
Ragnar Nurkse School of Innovation and Governance, Estonia

Opponents: Professor Dr. Jarmo Vakkuri, University of Tampere, Finland

Assistant Professor Dr. Bram Verschuere, Ghent University, the
Netherlands

Defense of the thesis: January 19, 2016

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for any academic degree.

/Küllli Taro/

Copyright: Külli Taro, 2016
ISSN 1406-4790
ISBN 978-9949-23-880-4 (publication)
ISBN 978-9949-23-881-1 (PDF)

CONTENTS

LIST OF ORIGINAL PUBLICATIONS	4
INTRODUCTION.....	5
SCOPE AND AIM OF THE THESIS.....	5
METHODOLOGY.....	10
PERFORMANCE MANAGEMENT AND MEASUREMENT.....	12
THE CONCEPT OF ATTRIBUTION IN SOCIAL SCIENCES.....	16
Roots in social psychology.....	16
Attribution theory in organizational studies.....	17
The attribution problem in evaluation studies and performance measurement literature.....	18
CONSEQUENCES OF THE ATTRIBUTION PROBLEM FOR PERFORMANCE MEASUREMENT.....	22
Descriptive, explanatory and advisory purposes of performance measurement ..	22
Evaluating perception.....	24
Developments in performance measurement and audit methodology.....	26
PERFORMANCE MEASUREMENT AND THE ATTRIBUTION PROBLEM IN THE ESTONIAN CONTEXT OF AN IMMATURE POLICY ENVIRONMENT.	28
IMPLICATIONS FOR FURTHER RESEARCH AND POLICY-MAKING	32
REFERENCES	35
SUMMARY IN ESTONIAN	42
ACKNOWLEDGEMENTS	50
PUBLICATIONS (Articles I-III)	51
CURRICULUM VITAE	75
ELULOOKIRJELDUS.....	125

LIST OF ORIGINAL PUBLICATIONS

The dissertation is based on the following original publications:

I Nõmm, Külli and Tiina Randma-Liiv. 2012. "Performance Measurement and Performance Information in New Democracies" *Public Management Review* 14(7), 2-21. (1.1)

II Raudla, Ringa, Külli Taro, Cherlin Agu and James W. Douglas. 2015. "Impact of Performance Audit on Public Sector Organizations: The Case of Estonia" *Public Organization Review*. Published online. DOI <http://dx.doi.org/10.1007/s11115-015-0308-0>. (1.1)

III Taro, Külli and Eerik Hanni. 2015. "Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite-Sports Policy: The Case of Estonia." *Administrative Culture* 16 (1), 24-46. (1.1)

INTRODUCTION

SCOPE AND AIM OF THE THESIS

Although collecting and utilizing performance information is as old as public administration itself (Hood 2007; Pollitt and Bouckaert 2011, 90), no other public sector reform movement has promoted performance management to the extent comparable to that advocated by New Public Management (NPM) (Van Dooren 2008). In recent studies, some authors say that NPM reforms have stalled or even reversed (Dunleavy et al. 2006; Christensen and Lægheid 2011), while others point out that ideas have just evolved or developed (Bouckaert and Halligan 2008; De Vries 2010; Put and Bouckaert 2011). In either case, the focus on performance in the public sector has remained.

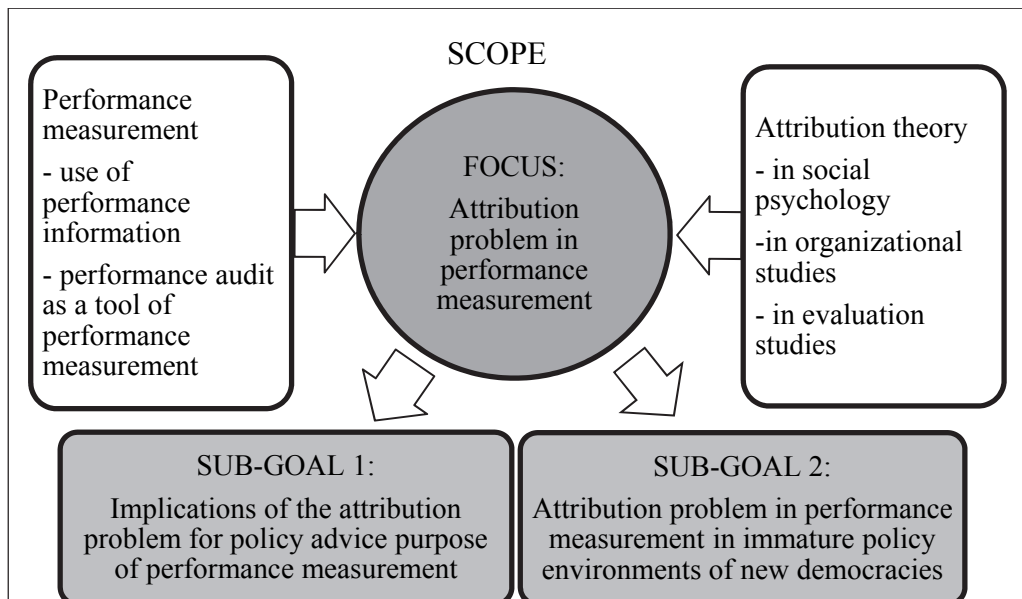
The performance measurement debate involves a wide array of issues. While much of the public sector performance literature deals with questions like “why?”, “what?” and “how?” are measured, the core of most current discussion forms around the question “whether and how this information is used?” Existing findings propose that efforts invested in measuring performance considerably surpass the eagerness shown in using the information for management or democratic purposes (see I; Lægheid et al. 2006; Pollitt and Bouckaert 2011). The reasons for limited use of performance measurement information stem from conceptual as well as technical or methodological obstacles. The main conceptual problem of performance measurement in the public sector is considered to be the failure to ineluctably define public sector aims (Van de Walle 2009). But the central technical challenge (Curristine 2005; Mayne 2007), analytical (United States General Accounting Office 1997; Bovaird 2014) or methodological (Davies 1999) weakness or even risk of performance measurement (Cuganesan et al. 2014) is called the “attribution problem”. The attribution problem denotes the difficulties in determining what contribution the specific policy or instrument in question has made to the outcome. How to differentiate between causality, correlation or just coincidence? How much success (or failure) can be attributed to program or other government activities? The attribution problem is an underlying obstacle in determining the causes of outcomes, as opposed to factors which are merely associated with outcomes. Attempts to bring outcomes into public sector decision-making have so far met with only limited success, partly due to the inherent difficulty of specifying outcomes, partly due to the poor measurement systems available to calibrate outcomes, and partly due to the difficulty of interpreting results based on outcomes, given that many factors other than public policy programs typically contribute to these outcomes (Pollitt and Bouckaert 2011; Bovaird 2014).

In performance measurement and audit literature, the attribution problem is increasingly being realized (see Davies 1999; Mayne 2001, 2007; Pollitt and Bouckaert 2011; Bovaird 2014, Hatry 2013; Cuganesan et al. 2014 for example), but still overlooked. There is a lack of in-depth analysis of the consequences of the attribution

problem for performance measurement and for use of performance information in practice. Many NPM texts and government practices do not pay attention to the cause-and-effect relationship of actions and results. As long as the results are proven, it does not matter how they came about (Van Dooren et al. 2010, 3). Performance measurement systems provide little information on factors that contributed to the success or failure of a given program (Bovaird 2014; Hatry 2013): causality is often just assumed (Nielsen and Ejler 2008). Therefore, the role of cause-and-effect analysis has been poorly understood in performance movements in recent decades.

The main focus of this thesis is on analyzing the challenges caused by the attribution problem for utilization of performance measurement results. The theoretical basis of the thesis rests upon performance measurement literature. Specific attention is paid to the use and different purposes of performance information. In order to contribute to bridging the gap in performance measurement literature concerning the attribution problem, the thesis integrates in the discussion another theoretical foundation of attribution theory which is known in other fields of social sciences, mainly in psychology and organizational studies. It explains the origins of the attribution problem by demonstrating that human psychology is naturally aimed at finding causal explanations for individual and social events. Different features of the attribution problem in performance measurement are illustrated by empirical examples from several Estonian performance audits that attempt to measure performance in the public sector, and a comparative empirical study about evaluating the impact of performance audits. (see Figure 1)

Figure 1. Scope and focus of the thesis



Source: author

The thesis has two sub-goals – firstly, to explore how the attribution problem affects the different purposes that performance measurement is intended to fulfil, especially the use of performance information for policy advice and change. The second goal is to examine the application of performance measurement initiatives in the Estonian context: that is, in the context of a new democracy with an immature policy environment. In fulfilling these purposes, performance audit is handled as a tool of performance measurement. As a synthesis of those two sub-goals, the thesis shows that the importance of the attribution problem has still been neglected in the performance literature and in application of performance measurement initiatives. This may lead to difficulties in implementing performance measurement, limited use of performance information or uninformed and inappropriate transfer of public sector reform initiatives (I; Dolowitz and Marsh 2000).

More specifically, the thesis addresses the following research questions:

- What are the implications of the attribution problem for performance measurement?
- How does the attribution problem particularly affect fulfillment of the purpose of giving policy advice in performance measurement?
- How does the attribution problem influence the development of performance measurement and the use of performance information in Estonia - within the context of a new democracy with an immature policy environment?

Drawing a distinction between three levels of performance measurement is helpful for structuring the discussion on performance in the public sector. These levels are: individual performance, organizational performance, and program performance (Talbot 2007, 494-496). This thesis is mainly focused on the third level - performance measurement of policy programs. In short, “program” means any activity, project, function, or policy that has an identifiable purpose or set of objectives (Newcomer 2007; Talbot 2010). It may (or may not) cross the boundaries of individual organizations. The author uses the terms “immature policy environment” (as an adjective) and “new democracy” (as a noun) to describe the Estonian policy-making context and administrative environment. It does not necessarily refer to a negative or under-developed, but to a different, policy-making and administrative context that stems from the still relatively short traditions of independent statehood, developing democracy, and a post-communist past.

The body of argument of the thesis is developed in three original articles. The overview article “Performance measurement and performance information in new democracies: A study of the Estonian central government” (I) (co-authored with Tiina Randma-Liiv) explores the development and implementation of performance measurement practices in Estonian central government since 1991. It highlights the difficulties in introducing performance measurement tools in immature policy

environments. Factors such as administrative instability, poor strategic planning and policy analysis, an implementation gap and uncritical, uninformed transfer of performance initiatives make the development of performance measurement in new democracies even more complicated than in the West. Although this thesis refrains from drawing conclusions about other Central and Eastern European (CEE) countries, these factors are brought out by researchers as characteristic of many CEE countries (see Verheijen 2003; Nakrošis 2008; Meyer-Sahling and Yesilkagit 2011). The second article “The Impact of Performance Audit on Public Sector Organizations: The Case of Estonia” (II) (co-authored with Ringa Raudla, Cherlin Agu and James W. Douglas) analyzes how performance audit as a performance measurement tool itself contributes to performance improvement. Civil servants perceive performance audit as useful even though it does not often lead to specific changes in policies or organizational practices. The third article “Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite Sports Policy: The Case of Estonia” (III) (co-authored with Eerik Hanni) demonstrates that Estonia’s relatively high ranking in elite sport medal tables does not correspond to the low scores given to its elite sport policy system by international assessment. Thus, sporting success has occurred despite shortcomings in public policy. Other contextual factors outside public policy and availability of resources may have led to success. The article points out that an unclear cause-and-effect relationship between sports policy and success of athletes creates difficulties in giving policy advice.

All three articles address the attribution problem by analyzing use of performance information, fulfilling different purposes of performance measurement and measuring performance in performance audits. However, all articles handle the issue from different perspectives. The first sub-goal is targeted principally in articles II and III (see also Figure 1). Article I is the main contribution to the second sub-goal of the thesis.

The author of this thesis was the lead author in article I. In article II the author was engaged in methodology development, data collection and in interpreting the results of quantitative analysis. Article III was constructed and written by the author of the thesis; the second author was mostly involved in data collection.

The introductory part of the dissertation is built up as follows. The first methodological section describes the sources of empirical data for the original articles and how the findings in the articles refer to the underlying problem of attribution in performance measurement. This is followed by a brief overview of the subject of performance measurement in public administration literature. Specific attention is paid to the different purposes of performance measurement. Thereafter, to explain the origins of the attribution problem, the theory of attribution in social psychology and its applications in other fields of social science are introduced. Next, the different consequences of the attribution problem to the purposes of performance measurement and its methodological approaches are discussed. A separate chapter is dedicated to

explaining how the importance of the contextual factors of immature policy environment in the Estonian case is amplifying the attribution problem. Finally, the conclusion explains the contribution of the thesis to existing research and policy implications of the attribution problem for performance measurement.

METHODOLOGY

The thesis is a combination of three independently written articles. Each article is founded on a separate empirical study but developed from the same theoretical basis. The theoretical framework is based on a synthesis of existing research on performance measurement and performance audit literature. The Estonian case is opened up within the context of new democracy and literature on policy transfer.

The empirical evidence in the thesis derives mainly from a qualitative document analysis (I) and two different international comparative studies (the performance audit effectiveness survey (II) and the Sports Policy Factors Leading to International Sporting Success (SPLISS) study on elite sport systems (III)):

- The document analysis carried out for the overview of different performance management initiatives applied in Estonia was based on various government documents and performance audit reports by the National Audit Office of Estonia. The study included over forty performance audits from the last decade. (I) The document analysis for article I was complemented with personal observations from the authors. For more than a decade, the authors of the article have closely followed performance measurement reforms from an academic as well as a practical perspective.
- Article II is based on the results of a survey of Estonian public officials who have experienced one or more performance audits conducted by the Estonian National Audit Office in the period 2005-2012. The survey was based on a research instrument developed by Reichborn-Kjennerud (2013) and is part of a Nordic cooperation project between Norway, Finland, Sweden, Denmark and Estonia entitled “The influence of performance audit on public administration and politics - A comparative study of the Nordic Supreme Audit Institutions”. The questionnaire for the survey was modified for the Estonian context by the author of the thesis. The survey looked at the perception of civil servants on the influence of performance audits.
- In article III, the effectiveness of Estonian elite sport policy, that is, the relationship between elite sport policy programs and international sporting success, is assessed according to a comparative international research methodology called SPLISS. The initial SPLISS study framework was worked out by a consortium of research groups from three nations (Belgium, the Netherlands and the United Kingdom) in 2002. Afterwards the research model was developed further and more countries joined to use the methodology. A comparative SPLISS 2.0 study involving 15 nations (including Estonia) was published in autumn 2015 (De Bosscher et al. 2015). The research model is based on nine “pillars”, i.e. factors that are believed to lead to international sporting success. These pillars or factors were derived from a review of

literature, pilot studies and interviews with athletes and high performance experts. Each pillar is made measurable by 6–21 critical success factors. Each critical success factor is measured by one to four different research instruments: 1) the overall sport policy inventory – a specific questionnaire on every pillar assembled by experts to be answered by researchers collecting data via semi-structured interviews or using secondary data sources (statistics, surveys) and document analysis (legislation, state budgets); 2) an elite athletes' survey; 3) an elite coaches' survey and 4) a sport federation survey.

The thesis uses these different empirical studies to demonstrate how the attribution problem becomes evident in performance measurement (**III**), hinders execution of the purpose of giving policy advice in performance measurement (**II**), and how these effects may be amplified in the context of a new democracy with an immature policy environment (**I**). All three original articles deal with performance measurement and use of performance information in the public sector. They all touch upon the causality issue and attribution problem by emphasizing the importance of contextual factors (**I, III**), the perception of actors in the performance measurement system, and highlighting that there is no clear link between government action and performance in a policy field (**II**). The common conclusion in these articles is that even if performance in the public sector is measured, use of this performance information still poses great challenges.

PERFORMANCE MANAGEMENT AND MEASUREMENT

In the literature, the terms ‘performance management’ and ‘performance measurement’ are used interchangeably with confusing definitions. This thesis is based on the definition of performance management as a type of management in which information in relation to performance is used for decision-making (Van Dooren et al. 2010, 30), whereas performance information is generated and utilized through performance measurement routines (Radnor 2008, 95) (I). Performance measurement, in turn, is understood as routine measurement of program inputs, outputs or outcomes undertaken in order to document program performance (Newcomer 2007). Inputs are defined as resources invested in a government program. While outputs are direct products and services delivered by the government, outcomes are the consequences of outputs and describe wider societal aims, often influenced by other externalities outside government activities. (III, Van Dooren et al. 2010, 16-36).

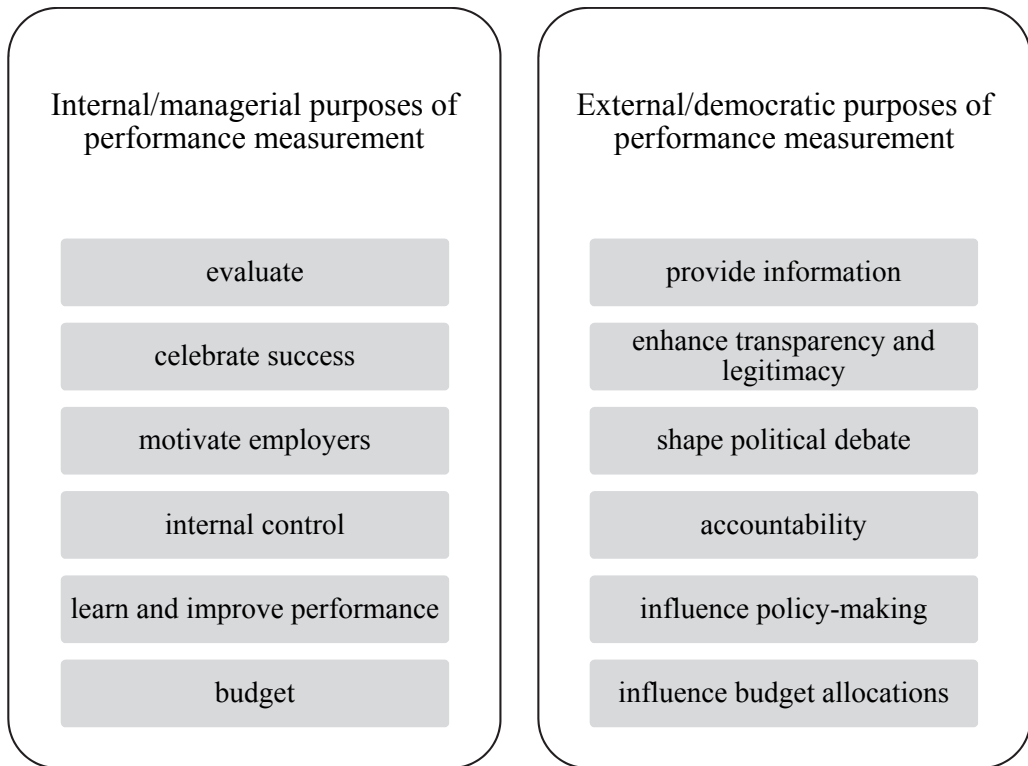
Integrating performance information into budgeting, managing and reporting has become a common component of public management. However, some authors cautiously advocate performance measurement (Wholey 1996; Perrin 1999; Holzer and Yang 2004), while many are skeptical towards its use and misuse, which may foster various kinds of performance paradoxes (e.g Halachmi 2005; Andrews et al. 2006; Bouckaert and Halligan 2008) or they dismiss performance measurement as a simplistic and crude form of knowledge production (Greene 1999; Nielsen and Ejler 2008). In practice, the positive impact of performance management systems on more effective and efficient policy-making and public sector performance is often taken for granted. However, there is no clear answer to the key question in performance management literature - whether performance management actually leads to better performance. Some studies have shown that performance management does not always lead to better performance and may even result in perverse effects (see Andrews et al. 2006; Andersen 2008; Radin 2009; Pollitt and Bouckaert 2011, for example), as performance measurement systems struggle to find a gaming-proof design (Bevan and Hood 2006; Radnor 2008). Although efforts to promote performance management have been under way for many years, progress is none the less usually seen as slow at best and few organizations would argue they have been completely successful (Mayne 2007; Christensen and Lægred 2011; Cuganesen et al. 2014). Despite heavy academic critique, performance measurement routines are deeply rooted in management practices in the public sector.

Performance management thus consists of performance measurement, incorporation and use of performance information (Van Dooren et al. 2010; Put and Bouckaert 2011). Recent research about performance management largely deals with the use of performance information: assertion of use or absence of use, different purposes of use, intended and unintended consequences of performance information utilization (see de Lancer Julnes and Holzer 2001; Pollitt 2006; Askim 2007, 2009; Hammerschmid et al. 2013 for example). The use of performance information is considered to be one of the

indicators of successes and failures of performance movements (Van Dooren 2008, 22). The thesis therefore also focuses on performance measurement practices and on utilization of information. Article **I** outlines the different performance measurement practices, while articles **II** and **III** focus more specifically on performance audit as one performance measurement tool. Performance measurement is the starting point and constituent part of any performance initiative.

The underlying question of any structured performance measurement debate is “Why measure performance?” Producing reports with reliable and valid information on government performance is no end in itself. Barely informative purpose is not enough. If a clear idea is lacking about how this performance information is used, the data provided and measurement methods used may not be appropriate for the intended purpose. To systematize the discussion, the purposes of performance measurement could be divided into internal and external purposes or managerial and democratic aims respectively (**I**; see Figure 2). This classification is based on the “end user” of the performance information (**I**; Pollitt 2006). The internal (managerial) purpose of performance information includes evaluating the progress of the program or organization, creating benchmarks, determining budget allocations, improving communication, giving feedback and serving as inputs for management decisions (Behn 2003). The external (democratic) use of performance information is aimed at informing the public, politicians, constituencies, or service users about the performance of public institutions. Publication of performance information is expected to deal with issues such as declining public confidence in government and growing demands for accountability by a more educated public (Halachmi 2005). This helps to achieve accountability and control, enhance democratic legitimacy and transparency in the public sector (Talbot 2007), and also serves as a symbolic public relations asset (Pollitt and Bouckaert 2011). There is considerable overlap between different purposes of performance measurement. Researchers use various lists to highlight the diversity of the purposes of performance measurement. Acknowledging that that these kinds of listings are provisional, Figure 2 provides a classification based on the aforementioned internal and external purposes of performance measurement.

Figure 2. Classification of purposes of performance measurement according to end users of performance information



Source: author; based on Behn 2003, Halachmi 2005, Pollitt 2006, Talbot 2007, Pollitt and Bouckaert 2011.

Different performance measurement tools are in place to fulfill these purposes (see overview of performance measurement tools in article I). Among a variety of performance measurement tools (e.g. performance budgeting, performance reports, service quality measures, client surveys), the rise of performance auditing during recent decades has shaped the way the public sector functions. Performance auditing has also gained a substantial position in the audit work of supreme audit institutions (II; III; Lonsdale 2000, 2011). For example, in the Estonian supreme audit institution, performance audit takes up nearly 60 percent of audit staff. In recent years, over one-third of published audit reports have been performance audits (III). In addition to being viewed as an instrument for enhancing accountability, performance audit is increasingly expected to contribute to wider policy debates, provide policy advice to governments and give guidance about how public administration can improve its work (Wilkins and Lonsdale 2007; Funkhouser 2011; Lonsdale and Bechberger 2011; Reichborn-Kjennerud 2013).

After decades of promoting performance management and measurement, it has turned out that, in practice, performance information is not so extensively utilized for management purposes or political decision-making (Lægneid et al. 2006; OECD 2007; Raudla 2012). Additionally, the majority of citizens do not benefit from the detailed performance information made available (Pollitt 2006). There are many reasons why performance information is not utilized to the extent expected and the improvement in actual performance has limited or even reverse evidence. Some authors point out that the rationality of policy-making and political decisions is overestimated and contextual factors underestimated (Pollitt 2006; Radin 2009; Pollitt and Bouckaert 2011; Virtanen 2013). Neither politicians nor citizens make decisions on considerations of economy, efficiency or effectiveness alone. They rely on their values, political views, personal or political gain (I). These reasons come, among other factors, from the conceptual shortcomings of performance measurement (Van de Walle 2009). Other reasons for low use of performance information arise from the technical (Currstine 2005, Mayne 2007), analytical (United States General Accounting Office 1997, Bovaird 2014) or methodological (Davies 1999) challenges of the attribution problem that is in the focus of the thesis at hand. Without assurance that performance information about the outcomes of public policy programs are causally linked to the government activities under observation, the opportunities to use this performance information are low.

THE CONCEPT OF ATTRIBUTION IN SOCIAL SCIENCES

Despite severe critique that performance measurement often fails to prove a causal relationship between measured outcomes and programs under observation (I-III, see also Mayne 2001, Bovaird 2014 ; Cuganesan et al. 2014), people rather tend to interpret that certain outputs and outcomes are results of a particular program or public policy. In political debate, it is understandable that politicians and lobby groups use somewhat questionable arguments to support their interests. There is also a tendency that causal link is generally recognized in the case of a positive relationship between a government program and its results (Askim 2007, 2009; Peters 2008). In the case of negative performance outcomes, the relationship is more easily neglected. The attribution problem therefore provides an opportunity to draw cause-and-effect relationships in a manner favorable to the interests of the participants in the debate. However, beliefs are often mistaken for knowledge. In practice, it is hard to accept the fact that there is no causal relationship or there are many other factors that contribute to the measured impact and the role of a particular program is often unclear.

Performance measurement literature does not provide an answer to this tendency to attribute results to measured activities, regardless of the existence of a causal relationship. Even evaluation studies that deal in more detail with the attribution problem of performance evaluation do not explain how people rationalize performance information. In order to contribute to relevant research, this thesis turns to other social sciences, particularly to organizational psychology, to explain the attribution phenomenon that seems to be deeply rooted in human nature. The attribution theory known from social psychology, but developed further in many other social science disciplines, handles the issue at the individual level, whereas this thesis elaborates it further to policy-analysis. As performance analysts and performance information users are humans, their way of interpreting social phenomena influences the results of performance evaluation. A broader view of the attribution phenomenon helps to understand the origin of the attribution problem and provide avenues for framing and potential alleviation of this analytical obstacle.

Roots in social psychology

The essence of attribution as a phenomenon in social sciences is explained with the help of attribution theory, which originates from social psychology. Attribution theory was developed in the works of Heider (1958), Kelley (1973), and Weiner (1986). It is a theory about how people make causal explanations, attempting to describe and explain how humans try to answer the question “why and how”. It deals with how the social perceiver uses information to arrive at everyday causal explanations for individual and social events. It examines what information is gathered and how it is combined to form a causal judgement.

Fritz Heider was the first to propose a coherent theory of attributions (Kelley 1973; Martinko et al. 2011). He claimed that individuals have a need to understand and explain the events around them. People are “naive psychologists” trying to make sense of the social world. Attributions are individuals’ explanations for the causes of their successes and failures. People are active interpreters and they use consistent and logical modes of sense-making in their interpretations. This preoccupation with causation is seen as a necessity for evolution and survival. People make sense of their surroundings on the basis of what they consider is the cause and what is the effect of a phenomenon. The underlying attempt to understand the world around us is considered universal and predictable. People tend to see cause-and-effect relationships even where none exist. (Heider 1958)

Kelley (1973) also worked on describing how people use information to make causal inferences, but his focus is not limited to interpersonal perception. Kelley aimed his attention at the subjective experience of attributional validity, at the social and self-perception. For example, he asked: “If a person fails on a test, does he have low ability, or is the test difficult?” (Kelley 1973, 107). Similarly in the case of competitive sport, questions arise whether the causes of success or failure of an athlete lie in the talent or effort of the athlete, the level of competition, the quality of the sporting system or other factors (see discussion on that issue in article **III**).

Kelley proposed the covariation principle and the configuration concept. The covariation principle (Kelley 1973, 108-113) means that events are attributed to causes with which, over time, they co-vary or co-occur. Causes are attributed to factors that are present when an event or effect is observed, and not present when the event or effect is absent. Likewise, in performance measurement, causes are often attributed to variables that correlate in a regression analysis. The configuration concept (Kelley 1973, 113) proposes that the role of a given cause in producing a given effect is discounted if other plausible causes are also present. Kelley used experimental methods to subtract other effects from the causal effect. But unlike Kelley’s research field, it would usually not be possible to carry out wide social experiments to extract different causes of program outcomes.

The third well-known scholar, Bernard Weiner (1986), believed that attributions reflect human thinking rather than what may actually have prompted the behavior or event. An individual's causal attributions of achievement affect subsequent behaviors and motivation. He focused on the consequences of attributional explanations and developed the attribution theory of motivation (Weiner 1986).

Attribution theory in organizational studies

Attribution theory emerged from psychology, but in its broadest context, it is concerned with ascribing characteristics to any entity. Therefore, as a general conception of the way people think about and analyze cause-and-effect it has been

applied in many other disciplines, like organizational sciences (see Staw 1975; Martinko et al. 2011), including leadership and motivation studies (see Green and Mitchell 1979; Weiner 1986; Martinez et al. 2012), communication research (see Baxter and Braithwaite 2008), program evaluation or evaluation studies (see Davies 1999; Mayne 2001, 2007; MacKenzie and Blamey 2005; Bovaird 2014). For example, Staw (1975) demonstrated in his experiments and interpersonal simulations that performance data may cause persons to assign an entire set of characteristics to individuals, groups, and organizations. The perception of the causes of a certain behavior may affect the judgement and actions of both managers and employees in organizations. This thesis contains an article (III) that examines the impact of performance audits via the perception of civil servants who have participated in a performance audit. As it is methodologically complicated to extract the influence of performance audits from other factors affecting the policy field (e.g. the political agenda of the government, interest groups), often the best way to evaluate impact is to ask for the perception of those involved in the process (see III; Johnsen et al. 2001; Reichborn-Kjennerud 2013). In accordance with attribution theory and particularly Staw's approach, it is still important to keep in mind that this way we obtain a subjective perception of impact, not an objective judgement as to whether an audit caused particular changes.

In organizational sciences it is stressed that the question whether the processes seen by managers as being associated with high performance actually contribute to performance, remains unresolved. Most organizational theories contain hypothesized independent variables which can either be the causes of performance, the effects of performance, co-variables of third variables, or the results of a network of reciprocal causation (Staw 1975).

Attribution theory can be seen as a set of general principles offered to explain certain observed phenomena of causal interpretations (Kelley 1973). In psychology and organizational sciences this is used to understand and explain. But attribution theory also has a somewhat different function. It helps not only to understand a phenomenon but also to recognize a problem. "Attribution theory" turns into "attribution problem" in evaluation studies and similarly in performance measurement (see also III; Talbot 2010, 48). It helps to identify and consider the human phenomenon of attributing causes to events without an actual causal relationship as a problem for performance evaluation. Therefore, in this thesis attribution is called a problem because it needs to be taken into account when measuring performance and especially when cause-and-effect validity is assumed in cases of giving policy advice.

The attribution problem in evaluation studies and performance measurement literature

The attribution problem has also been addressed within evaluation studies. This is a discipline closely related to, often considered as complementary to, performance

measurement. For many policymakers and managers, performance measurement is evaluation and they use “evaluation” and “performance measurement” as synonyms. Both pose the same question: “Is the program successful?” Program evaluation and performance measurement have essentially traveled on separate, but somewhat parallel paths. Both forms of knowledge production share similar processes and tools for structuring and planning, obtaining data, analyzing and inferring judgement. Differences lie in the scope, depth, multiplicity and frequency of the purposes to be carried out rather than in kind. (Wholey 1996; Mayne 2007; Nielsen and Ejler 2008; Hatry 2013)

However, performance management receives much more attention in public sector management practice. Evaluation reports, though equipped with knowledge of the cause-and-effect relationship between government policies and their impacts, often produce a body of knowledge that appears too late and is too long to be useful as a management tool (Wholey 1996, Davies 1999). To undertake an evaluation requires considerable skills and can be costly, with results not always guaranteed (Mayne 2007). Policymakers demand timely information on the results of the program delivered and they do not bother reading long detailed reports. This is a reason why management consultants (also from international organizations) embrace performance measurement producing a body of knowledge that is both timely and brief (Nielsen and Ejler 2008).

Wholey (1996) has stated that performance measurement can serve as a one-shot program evaluation, but this statement ignores one of the essential differences between evaluation studies and performance measurement literature, which is the approach to causality. A major purpose of program evaluation is to identify, to the extent possible, the causes of the outcomes being tracked (Hatry 2013). Evaluation pays great attention to determining causality; attribution of the outcomes to the program is often a key aim of evaluation studies, whereas in performance measurement causality is assumed (Nielsen and Ejler 2008). Performance measurement focuses on measuring program outputs and outcomes, without attempting to identify the extent to which the program has caused the outcomes. Outcomes, however, are often influenced by many other externalities outside government programs, and may not be caused or only caused by government activities (III; Van Dooren et al. 2010). As causality is assumed in performance management, the research literature rarely contains references to attribution theory or the attribution problem.

Differently from performance management research, program evaluation literature on attribution is vast (see Davies 1999; Rossi et al. 2004; Hatry 2013; Bovaird 2014 for example). Evaluation discipline mostly tries to find answers to the attribution question by using some form of controlled comparison to assess what happens with the program in place versus what would happen without it. Changes from the initial situation to the final outcome are captured so that explanations are based on what has actually been delivered (MacKenzie and Blamey 2005). With the help of attribution theory it is

explained how some important biases may arise in the performance evaluation process, including gender bias or subjective interpretations based on beliefs or cultural background (see Gedeon and Rubin 1999).

Extensive social science research methods have been designed to address the problem of attribution in evaluation studies. Methods like quasi-experimental pre- and post-intervention surveys, longitudinal studies, context analyses and a range of pre-specified qualitative investigations, often adapted from psychology and organizational studies, are used to reduce the problems associated with causal attribution that commonly plague evaluations of complex interventions. (Davies 1999; Mayne 2001; MacKenzie and Blamey 2005) However, not even sophisticated methods can avoid the attribution problem, although the research design systematically focusing on the issue in evaluation studies may alleviate it.

Compared to evaluation, performance measurement tends to rely on quantitative methods and far less on structured qualitative data collection techniques such as documentary research, observations or interviews (Nielsen and Ejler 2008). Even quantitative methods in use in performance measurement are generally quite limited. They are mostly after-the-fact evaluations, excluding the use of such evaluation designs as randomized controlled trials which professional program evaluators develop (Hatry 2013). In many quantitative performance analysis reports the need for more attribution studies is still recognized by the appearance of the final lines: “It is not possible to attribute these changes to the activities supported by the project.” (White 2010) However, this warning, often written in small letters, can be overlooked when performance information is used to explain the reasons behind measured outcomes, giving policy advice or allocating resources for further program initiatives.

To conclude, program evaluation provides in-depth but ad hoc information on major policy issues. It needs considerable time, money and expertise. Performance measurement usually provides regular, more timely information than evaluation and can cover a number of government programs. (Davies 1999; Mayne 2001; Newcomer 2007; Hatry 2013) Because of its ongoing nature, performance measurement can serve as an early warning system to management. As long as quality controls are in place, it can cover many public programs during each year (United States General Accounting Office 2011; Hatry 2013). But unlike in evaluation studies, performance information produced in performance management initiatives often does not provide sufficient analysis on the causes of revealed impact (**I-III**). For example, the impact of single performance audits is evaluated via regular feedback questionnaires from several supreme audit institutions, including the Estonian National Audit Office, though this routine feedback is not able to differentiate between changes made to public policy programs because of the audit, the government’s independent activities or lobbying by interest groups (**II**). Only rarely can causality be established with any confidence by data collected on a routine basis. The impact of government policy programs on wider societal outcomes (such as the popularity of sport among citizens, participation in

leisure sports and the international image of the country in the case of elite sport policy) or even outputs (success in international sporting events) **(III)**. The problem is that there are almost always competing explanations that cannot be ruled out with only outcome indicators **(I-III)**, see also De Lancer Julnes 2006).

CONSEQUENCES OF THE ATTRIBUTION PROBLEM FOR PERFORMANCE MEASUREMENT

An insight into the lessons from different disciplines of social sciences provides valuable hints on how to handle or overcome the attribution problem in measuring performance in the public sector. The tendency of humans to subjectively attribute results to specific activities needs to be addressed. The attribution problem affects methodological approaches towards measuring performance and also use of performance information.

Descriptive, explanatory and advisory purposes of performance measurement

Advocates of performance measurement point out that performance evaluations based on indicators without burdensome cause-and-effect analysis provide timely and easy-to-read information on public sector outcomes. Policymakers and the public will likely be considerably better informed than if no such information was available (Hatry 2013). Thus, any information is thought to be better than no information. Others point out that without information about the presumed causality between government activities and results, performance measurement information is insufficient for making decisions such as budget allocations (De Lancer Julnes 2006). Furthermore, the assumption of performance measurement that studying the past is a sure way to navigate into a better future is shaky. (I; Halachmi 2005)

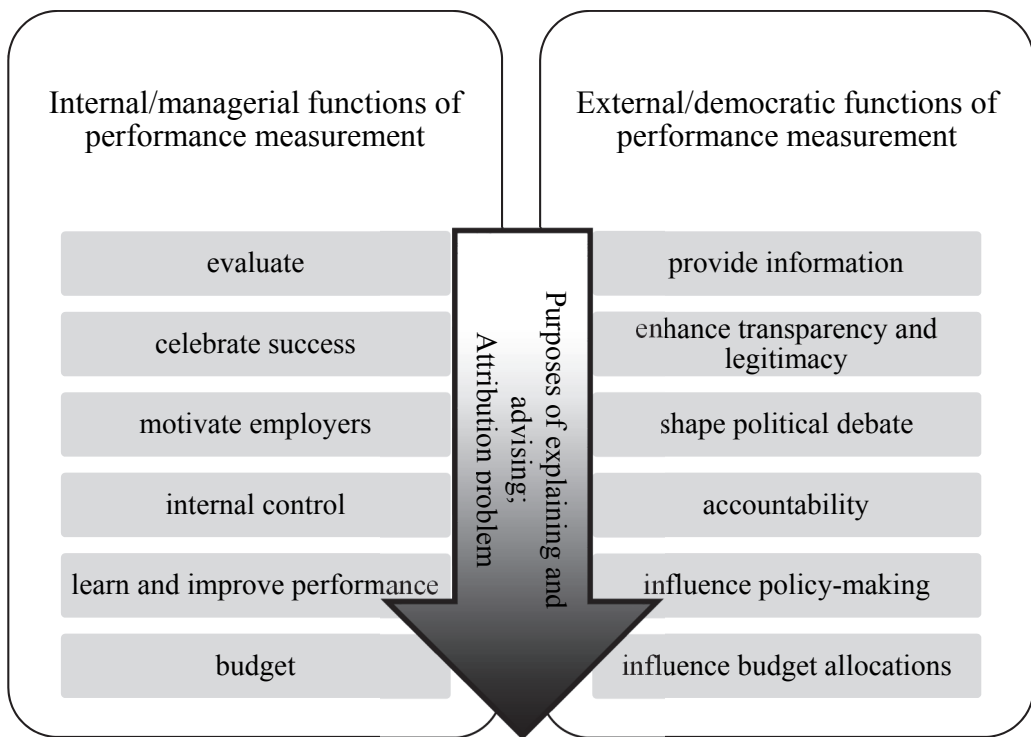
However, policymakers do not only want to be informed about public sector outcomes, but also ask for guidance on where to spend public money. Measuring outcomes, which means describing whether or not outcomes actually occur, is one challenge. Determining the extent to which a program contributed to those outcomes is quite another issue. Often a number of factors other than the program under scrutiny may have contributed to the observed outcomes. Indeed, the outcomes may have occurred without government involvement. To be able to make any assessment about the worth of spending public money on a program, some idea of how the program has affected the desired outcomes is needed. (Mayne 2001; 2007; Bovaird 2014; Cuganesan et al. 2014) The more focus of performance measurement goes to outcomes instead of outputs (see distinction in III), which is the recent trend in performance measurement, the more crucial the attribution problem becomes. (III; Mayne 2007; Lonsdale 2011; Pollitt and Bouckaert 2011; Put and Bouckaert 2011)

The attribution problem combines the topics of performance measurement and performance information use. This is a methodological problem, but becomes evident in utilization of performance information. Moreover, performance measurement may be expected to target several purposes at the same time. For example, analysis of the effectiveness of performance audit reveals that public servants expect performance

measurement to fulfill several purposes concurrently. There is no trade-off seen between accountability and improving the tasks of performance audit. (II)

Depending on what purpose the performance information is used for – descriptive, explanatory or advisory, the cause-and-effect relationship between actions and outcomes becomes more or less crucial for performance analysis. The more an explanatory or advisory purpose is prevalent, the more crucial the attribution problem becomes (see Figure 3 for illustration). Providing information on various socio-economic indicators for enhancing public awareness of the situation or celebrating success in the case of a positive change in measures does not necessarily require assurance on the causal relationship between government activities and measured outcomes. Governments may even be held politically accountable for impacts that are not directly connected to public policy programs. But as soon as the need arises to explain cause and effect, thus advising on further policy options or budget allocations, the attribution problem cannot be neglected. Otherwise, inadequate recommendations may lead to policy failure and waste of resources.

Figure 3. Relationship between different purposes of performance measurement and importance of attribution problem



Source: author; based on Figure 2 (Behn 2003, Halachmi 2005, Pollitt 2006, Talbot 2007, Pollitt and Bouckaert 2011).

Outcome-based or evidence-based policy-making has been at the center of recent performance movements in the public sector (Van Dooren 2008; Pollitt and Bouckaert 2011). But it only works if we can deal with the attribution problem, i.e. determining the causes of outcomes, as opposed to factors which are merely associated with outcomes. The extent to which an outcome and its observed variance can be attributed to the program remains a key methodological hurdle in performance measurement and indeed any goal-directed evaluation work (Bovaird 2014). Despite the measurement difficulties, attribution is a problem that cannot be ignored when trying to assess the performance of government policies. Without an answer to the attribution question, little can be said about the impact of government activities, nor can advice be provided about future directions. (III)

For example, Estonian elite sport policy analysis (III) demonstrates that success in the highly competitive international sporting arena may occur despite shortcomings in public policy programs. The comprehensive multidimensional international comparative study shows that although Estonia holds a relatively high position in international sporting indexes (especially in per capita results), most of the factors crucial to the sports system (e.g. talent identification, athletic and post-career support) are at the moderate or even limited level of development. This finding indicates that other contextual factors outside the public policy program (such as personal abilities, cultural background, natural resources or climate) and thus outside government activities may lead to success. Success may even occur randomly or unsystematically. According to attribution theory, the conclusion that good performance in elite sport may not be causally explained is difficult to accept. Policy analysts as well as policymakers tend to see a causal relationship even where there is none. It is in human nature to actively interpret success and failure, a necessity for evolution and survival. (Heider 1958; Kelley 1973)

Elite-sports policy is a distinctive example of the attribution problem as it is probably one of the few public-policy fields where measuring performance seems to be fairly straightforward. Methods vary, but in principle, success can be clearly measured in medals won at international competitions. Merely indicator-based descriptive performance measurement (performance audit in this particular case) would have concluded that Estonian elite sport policy programs provide good value for money. But based on this conclusion, it would have been misleading policy advice to make no changes to elite sport policy or recommend an Estonian model to others for policy transfer. Estonia's relative success in international competitions does not represent the effectiveness of public policy factors. (III)

Evaluating perception

Many researchers are skeptical about overcoming the analytical challenge of attributing outcomes to programs or policies (see Feinstein 2002; Mayne 2007; Miller and Fox 2007; Newcomer 2007; Bovaird 2014). Miller and Fox (2007, 36-37) have even stated

that the claim in performance measurement that outcomes, results, effects, and even non-events can be attributed to programs and policies amounts to “fanciful faith in what social science methodology can accomplish.”

One way to deal with the methodological problem of attributing results to policies or other activities is to turn to the roots of attribution theory and accept that performance data only reflect the respondents’ perception of performance rather than actual events (see Staw 1975; Gedeon and Rubin 1999; **II**). Indeed, performance may not be an objective reality waiting to be measured, but is, rather, a “socially constructed reality” instead (Newcomer 2007).

For example, measuring the impact of performance audit is complicated, if possible at all. One can find that after the audit, things will be done in a way consistent with the audit’s recommendations. Is this evidence of audit effectiveness? The fact of consistency between audit recommendations and changes implemented in practice is not necessarily an indication of use of audit results. Other explanatory factors may be involved. (Lonsdale 2000; Feinstein 2002; **II**) But the expansive exploitation of performance auditing during recent decades (**II**; **III**; Lonsdale 2011) should encourage analysis of the impact of this massive auditing work. So far different studies have provided contradictory evidence whether performance audit actually leads to changes within audited organizations (see **III**; Van Looke and Put 2011).

Article **II** in this thesis explores the impact of performance audits of the Estonian Supreme Audit Institution by looking at the perceived usefulness of audits in the eyes of the auditees and by examining to what extent audits (as perceived by the auditees) led to changes in the audited organizations (see also Johnsen et al. 2001; Reichborn-Kjennerud 2013). In the Estonian case, while 40 % of the auditees found performance audits to have been useful, only 21 % agreed that they led to the adoption of changes in the audited organizations. However, the analysis suggested that although there is a correlation between the perceived usefulness of the audit and the extent of changes it brought about, the auditees can perceive performance audit to be useful even if it does not lead to specific changes in policies or organizational practices. By examining the perceived usefulness of performance audits it is possible to gauge the impacts of the audit more broadly so that it could be used to provide a more holistic picture of the role of performance audits in the public sector. In addition to positive or negative perceptions of audit impacts and usefulness, the analysis explored in more detail what internal factors (e.g. willingness of the auditors to engage in dialogue during the audit process, more specific recommendations) and external factors (e.g. attention from parliamentarians or the media) contribute to the perceived usefulness of audits and are believed to actually lead to changes. Therefore, this analysis fulfilled the explanatory and advisory purposes of performance measurement and allowed a step beyond the merely descriptive or informative approach. (**II**)

Developments in performance measurement and audit methodology

Traditional performance measurement systems that are based on measuring quantitative indicators provide little information on factors that contributed to the success or failure of a given program. It is difficult to establish a causal link between government effort and the eventual outcome for the individual or society served. And in some cases it is simply impossible to measure program outcomes in the tight timeframe that performance measurement process requires. (Davies 1999; Mayne 2007; Newcomer 2007)

The shortcomings in performance measurement methodology that do not allow the challenges of taking into account the attribution problem to be met have led to adoption of more complex research techniques similar to evaluation studies. While evaluation reports may be burdened by too much information, performance measurement may tend to be burdened by too little information. Evaluation may help to explain performance by adding contextual evidence to the reporting that will provide a balanced assessment of government performance (Newcomer 2007; Nielsen and Ejler 2008).

Therefore, the current distinction from performance measurement is not a tenable one (Nielsen and Ejler 2008, 188). Evaluators need to recognize performance measurement as a complementary form of knowledge production if they want to be engaged in more than academic discussions and provide relevant reports for policy decision-makers. And vice versa, the performance measurement discipline should take advantage of the wide array of analytical tools that social scientists have long advocated for use in evaluating the effectiveness of public programs, taking into account the contextual or mediating factors external to the program that may have caused documented results (Newcomer 2007). With the rapid development of ICT, data on outcomes become ever more accessible and up-to-date. Possibly, this has been an enabler for the trend of management disciplines such as auditing, monitoring and evaluation overlapping more than ever before (Nielsen and Ejler 2008).

Indeed, even the definition of performance audit has moved from the measurement of economy, efficiency and effectiveness towards more “diagnostic” performance audits (Put 2011). A recent definition of performance audit by Furubo (2011) emphasizes that performance audit is also an evaluative activity. It has gone beyond concerns with regularity and compliance, just describing and informing. Performance audit has developed from being a mere accountability mechanism towards giving guidance on how public administration can improve and contributing to the wider policy debate.

Performance audit is currently understood as “a hybrid activity, making use of whatever means it can find to generate sufficiently robust evidence for its purposes” (Lonsdale 2011, 15). Although the role of a performance auditor may be crucially different from that of an evaluator by being part of the authoritative system of control

(see Pollitt and Summa 1999), performance auditors are turning more and more to the wide array of evaluation and scientific methods, making use of the best available qualitative as well as quantitative methods. It may even be claimed that in case performance audit takes an in-depth approach to determining the causality of success or failure of government programs, it represents more the tradition of evaluation studies than performance management and measurement discipline (see Kelly and Swindell 2002).

These developments are supposed to alleviate the explanatory deficiency of performance management. This thesis demonstrates how advanced research methodologies are used in a performance audit in order to include explanatory analysis of a policy process to provide appropriate policy advice (III). It also shows how lessons from attribution theory are used to overcome the challenge of evaluating the impact of performance audit as such by studying the perception of the impact (II).

Nevertheless, when attribution is an important issue, a well-designed evaluation is the best way to go. Less sophisticated approaches can only be useful in reducing at least to some extent the uncertainty surrounding attribution (Mayne 2007). It is unrealistic, however, to think that just incorporating some evaluation techniques will, in itself, resolve the methodological difficulties of outcome attribution, especially in the usual performance management implementation context (Davies 1999).

PERFORMANCE MEASUREMENT AND THE ATTRIBUTION PROBLEM IN THE ESTONIAN CONTEXT OF AN IMMATURE POLICY ENVIRONMENT

In the previous discussion, Estonian cases have been dealt with as typical examples of performance measurement regardless of the administrative context of the country. However, the political context and the administrative capacity of the public administration are critical factors for successful performance measurement and use of performance information (I). Article I in this thesis highlights a set of specific factors characteristic of the Estonian political and administrative context: an unstable political and administrative environment, lack of policy analysis competences, poor strategic planning, fragmentary and discontinued reforms, and willingness to please external actors. Because of these characteristics, the Estonian context can be described as an immature policy environment which seriously impedes implementation of a functional performance measurement system.

The underlying pre-condition for performance measurement is a well-established quality control system and routine measurement of inputs, outputs and outcomes (Newcomer 2007; United States General Accounting Office 2011; Hatry 2013). Performance measurement assumes that the test of causality between government activities and measurable results, thus addressing the attribution problem, is undertaken prior to the introduction of a measurement system (Nielsen and Ejler 2008). This confirms that implementing and using performance measurement systems requires much more from managers than mere technical measurement skills. Experience from countries across the world has shown that effective use of performance measurement presents managers with complex communication, analytical, political, and measurement challenges (Newcomer 2007). If there is a shortage of competent local policy-makers and quality controls are not in place, performance measurement poses severe methodological challenges including the attribution problem. For these reasons, the specific context of the country requires special attention in introducing public management initiatives, including performance measurement practices. (I)

Instability in the political and administrative arena, deficient analytical expertise and administrative capacity contribute to difficulties in implementing performance measurement initiatives in immature policy environments. Performance measurement when applied in the context of an immature policy environment is not suited to routinely cover many public programs during each year. It also cannot guarantee an adequate quality control system of performance information. But routine measurement and quality controls are the basic assumptions for introducing a functioning performance measurement system. Therefore, it is particularly complicated to deal with the attribution problem in immature policy environments. (I)

At the beginning of the 1990s Estonia, like many other newly independent post-communist states, inherited a highly politicized administrative apparatus with no

experience in policy design, policy evaluation and policy co-ordination. This was followed by frequent changes in administrative arrangements, short average lifespan of cabinets, and high turnover in political and administrative leadership. Any performance measurement system presumes a certain stability of performance targets and measurement routines to perform its evaluative and steering purposes. In Estonia, several reforms have been discontinued halfway through and new reforms started before the results of previous initiatives are evaluated. Unsustainability of initiatives and deficient performance indicators are often the results of poor strategic planning and analytical skills. Inexperienced managers are not able to make use of performance information and avoid gaming performance information. Inadequacy of accountability and control mechanisms may also prove risky for implementing a performance based administrative framework. (see I; Verheijen 2003, World Bank 2006; Nõmm and Randma-Liiv 2011)

Despite this unfavorable political and administrative context, Estonia, like most post-communist countries, has introduced performance measurement tools during the last couple of decades. The first reforms began in the early 1990s and were supported by many international donors, including OECD/SIGMA, and the PHARE program, the World Bank, the International Monetary Fund and bilateral donors (I; Peters 2008; Bouckaert et al. 2011; Nõmm and Randma-Liiv 2011). Later, the prospect of becoming a member of the EU also supported adoption of performance initiatives. Although performance management was not part of the *Acquis Communautaire*, several targets were set by the EU regarding management reforms in the public sector. This external pressure was accompanied by internal motivation to have a modern, efficient and effective public administration. (Nakrošis 2008; Peters 2008; Bouckaert et al. 2011)

Even after the EU-accession period, the role of exogenous actors remained important in directing reforms in the public administration of the new democracies. Despite severe academic criticism of performance management, (see e.g. Pollitt and Bouckaert 2004; Talbot 2007; Radnor 2008; Van de Walle 2009), numerous international organizations and experts continue promoting performance management and its tools as international “best practice”. For example, in its Public Governance Review report released in 2011, the OECD strongly encouraged the Estonian government to introduce several performance management tools without any reservations as to specific sectors or organizations, or attention to pre-conditions for successful implementation and potential misuse of the performance measurement system. In addition, conditions for the use of EU structural funds include setting and controlling output and outcome performance targets. (OECD 2011; Nõmm and Randma-Liiv 2011; Raudla 2013)

Many researchers studying new democracies agree that introducing NPM reforms, of which performance management and measurement reforms form part, provide particular challenges. For successful implementation, NPM in general, and performance measurement more specifically, require an environment of a well-functioning democratic administrative tradition (see Drechsler 2005; Nakrošis 2008;

Peters 2008; Bouckaert et al. 2011). Others represent a somewhat more optimistic view generally emphasizing that performance orientation in public management is commonly essential for successful government and has led to improvements in public service organization (see Curristine 2005; Dan and Pollitt 2015). The basic assumption of many reform advocates has been that there is a universal model of good public management that includes performance measurement. And if a good idea could be made to work in one place that would suit another place just as well (I; Peters 2008).

The process whereby knowledge about policies or administrative arrangements in one time and place is used in the development of another political or administrative system in another time or place is called policy transfer (Dolowitz and Marsh 2000). It is a common practice to draw lessons from the experience of other countries and to attempt to define good practice. But uninformed, incomplete or inappropriate policy transfer represents a potential for failure (ibid). Administrative capacity that is lacking in immature policy environments is a key determinant in the shift from erratic policy copying to responsible and competent policy learning (I). Therefore, when performance measurement is suggested and implemented as a technical tool without ensuring that the preconditions for adequate use are met, may lead to policy failure.

In fact, there is little evidence that introducing performance measurement in new democracies has improved the efficiency of public management or led to more effective public management. And the eagerness of new democracies to implement performance measurement practices often exceeds the practice in the role-model (mostly OECD) countries (I; II; World Bank 2003; Nakrošis 2008). The quality of performance information is poor and its use in the decision-making process is limited apart from external reporting (Nakrošis 2008; Raudla 2012; Raudla and Savi 2015). Although collecting data on performance may be a useful tool, it is no guarantee of improvement in management or in program performance (Newcomer 2007).

The Estonian government keeps experimenting with performance measurement tools because of the underlying belief or perception of their effectiveness. This is accompanied by a willingness to earn credits from international experts and organizations. Acceptance by international organizations and credibility for foreign partners has been important for Estonian society. With performance measurement being widely promoted as “best practice” of public administration, this foreign influence stimulates implementation of performance initiatives without the necessary administrative capacity and political preparedness to make good use of performance practices. (I)

Although the country strategies and starting positions in CEE countries were different, because of its positive attitude towards and eagerness to implement NPM ideas, Estonia frequently serves as a benchmark for evaluating the appropriateness and suitability of NPM principles and tools for new democracies with an immature policy environment. (Bouckaert et al. 2011) Therefore, the experience of Estonia in performance measurement may provide beneficial lessons for other countries.

However, it is ultimately not possible to distinguish between the impact of the transitional background and other contextual factors such as political culture. And any application of the Estonian experience to other CEE countries would need to be executed with great caution. (I)

To conclude, the context of an immature policy environment amplifies the attribution problem in performance measurement. Although international policy advice tends to attribute effectiveness in public administration to the implementation of performance measurement system, there is little evidence that performance measurement has actually caused improvements in the impact of public policy programs in Estonia. Development of a modern performance management system does not automatically imply an improvement in the efficiency of public spending or effectiveness of public policies.

IMPLICATIONS FOR FURTHER RESEARCH AND POLICY-MAKING

This thesis follows the cutting edge research of performance measurement that focuses on different purposes of measurement and use of performance information in the public sector. In contrast to the great majority of studies, this thesis shows that the attribution problem exaggerates or even causes many malfunctions of performance measurement. The attribution problem describes difficulties in determining what contribution the specific program or instrument in question has made to the outcome and thus has a far more complex impact on performance measurement systems than acknowledged in the existing performance literature. The thesis therefore focuses on the attribution problem as the underlying reason for challenges and risks to performance measurement.

By illustrating different perspectives of the attribution problem, the thesis discusses core methodological and analytical questions of this particular research field. First, it demonstrates that the importance of the attribution problem depends on the purpose that performance measurement is intended to fulfil in any particular case. Depending on the aim of performance measurement and how this information is later used, different requirements arise as to evidence concerning the cause-and-effect relationship between policy processes, government activities and documented results. As long as performance measurement is used merely for descriptive or informative purposes, reporting or monitoring, the attribution problem is not an essential obstacle to measurement. But if performance information is expected to give explanations for policy processes and their impacts, in short, to offer policy advice, the need to thoroughly investigate causal relationships and overcome the attribution problem becomes crucial. The second perspective takes into account the specific political and administrative context where performance measurement takes place. The thesis shows that the attribution problem is amplified in an immature policy-making environment because performance measurement presumes stable administrative routines with quality controls in place and experienced management with high levels of analytical expertise. The Estonian experience demonstrates that new democracies with immature policy environments may lack the necessary prerequisites for implementing an effective performance measurement system.

By combining the current debate in the performance measurement literature, studies on attribution in other fields of social sciences, and empirical findings of the articles, the thesis contributes to the academic discussion, provides new avenues for further research, as well as identifying practical implications.

In the literature, the attribution problem is mainly approached as the crucial analytical, methodological or technical challenge of performance measurement that hinders research-based policy development. But a conceptual challenge also arises as to how to take into account human nature to give causal reasons for the outcomes of activities and processes, and attribute changes to programs and policies even in a situation where

there is no cause-and-effect relationship. For theoretical contribution in performance measurement literature, the thesis synthesizes the performance measurement debate with research on the attribution phenomenon that is better known in other fields of social sciences, mainly psychology, organizational studies and program evaluation literature. It contributes to the existing line of research by incorporating lessons from attribution theory in other social sciences and analyzing the consequences of the attribution phenomenon for performance measurement. While attribution theory analyses the attribution phenomenon at the individual level, this thesis brings the findings of the theory to the policy-analysis level.

In a way, attribution theory is concerned with common sense. But Kelley (1973, 108) stressed that it is precisely the role of social scientists to explicate and systematize what at first seems obvious, not “to confound common sense but rather to analyze, refine, and enlarge on it”. The aim of performance measurement research, therefore, is to elaborate on how to overcome the attribution phenomenon in measuring performance that is yet so inherent to human nature. How to use the perception of the impact of public policy programs and government activities to evaluate performance? What are the advantages and disadvantages of the subjective interpretation of results and its causes?

Another challenge for future researchers is to explore how to combine the need for timely performance information and in-depth analysis of the cause-and-effect relationship between government activities and policy program performance. Performance measurement may benefit from integrating program evaluation tools into performance measurement systems. At least, it would be useful to triangulate quantitative indicator information with more detailed qualitative case studies. However, even the most sophisticated quantitative or qualitative methods may not be able to differentiate between attribution of activities to outcomes and causal relationships. And complex analysis needs time, money and expertise. Therefore, researchers need to ask in what circumstances the need arises to differentiate between occurrence of the outcome and the causal impact of government programs on the measured outcome? In what cases are we willing to overcome the intrinsic human urge for causal explanations and accept that measured outcomes may or may not be the results of government programs?

For practitioners, this discussion helps to find ways to improved policy analysis and suitable use of performance information. Taking into account the attribution problem in fulfilling different purposes of performance measurement is likely to lead to more appropriate use of performance information and adequate policy advice. In designing a performance measurement system, the purpose of the information and its potential use has to be kept in mind. Practitioners should ask: what are the necessary prerequisites for using performance information for descriptive, explanatory or advisory purposes? This may mean that managers and the general public need different indicators.

To overcome the methodological obstacles in determining a cause-and-effect relationship in performance measurement, evaluating the perceptions of the actors involved in a public policy program or government activities is one suitable option in practice. Performance data interpreters and users are also individuals and an understanding of human nature gives hints to explaining and overcoming the attribution challenge.

Finally, in implementing a well-functioning performance measurement system in a country, the political and administrative context plays a crucial role. Although several international organizations and policy experts actively promote introduction of performance measurement as the “best practice” of public administration reforms, this thesis brings out that there is no guarantee of success. A wide array of contextual factors, including the level of development of the administrative and political system, need to be taken into account in implementing a successful performance measurement system and relevant performance information use. Transfer of a performance measurement system from one country to another is a risky step and policy advice based on performance information has to be given with great caution.

REFERENCES

- Andersen, S. C. 2008. "The Impact of Public Management Reforms on Student Performance in Danish Schools." *Public Administration* 86(2), 541-58.
- Andrews, R., Boyne, G. and G. Enticott. 2006. "Performance Failure in the Public Sector." *Public Management Review* 8(2), 273-96.
- 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.
- 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.
- Baxter, L. A. and D. O. Braithwaite. 2008. *Engaging Theories in Interpersonal Communication Multiple Perspectives*. California: SAGE Publications, Inc.
- Behn, R. D. 2003. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63(5), 586-606.
- Bevan, G. and C. Hood. 2006. "What's Measured is What Matters: Targets and Gaming in the English Public Health Care System." *Public Administration* 84(3), 517-538.
- Bouckaert, G. and J. Halligan. 2008. *Managing Performance: International Comparisons*. London, New York: Routledge.
- Bouckaert, G., Nakrošis, V. and J. Nemec. 2011. "Public Administration and Management Reforms in CEE: Main Trajectories and Results." *The NISPAcee Journal of Public Administration and Policy* 4(1), 9-29.
- Bovaird, T. 2014. "Attributing Outcomes to Social Policy Interventions – 'Gold Standard' or 'Fool's Gold' in Public Policy and Management?" *Social Policy & Administration* 48(1), 1-23.
- Christensen, T. and P. Lægveid. 2011. "Beyond NPM? Some Development Features." In Christensen, T. and P. Lægveid (eds) *The Ashgate Research Companion to New Public Management*. Surrey, Burlington: Ashgate Publishing, 391-404.
- Cuganesan, S., Guthrie, J., and V. Vranic. 2014. "The Riskiness of Public Sector Performance Measurement: A Review and Research Agenda." *Financial Accountability & Management* 30(3), 279-302.

Curristine, T. 2005. "Government Performance: Lessons and Challenges." *OECD Journal on Budgeting* 5(1), 127-51.

Dan, S. and C. Pollitt. 2015. "NPM Can Work: An Optimistic Review of the Impact of New Public Management Reforms in Central and Eastern Europe." *Public Management Review*, 17(9), 1305-1332.

Davies, I. C. 1999. "Evaluation and Performance Management in Government." *Evaluation* 5(2), 150-159.

De Bosscher, V., Shibli, S., Westerbeek, H. and M. van Bottenburg. 2015. *Successful Elite Sport Policies. An International Comparison of the Sports Policy Factors Leading to International Sporting Success (SPLISS 2.0) in 15 Nations*. Aachen: Meyer & Meyer.

De Lancer Julnes, P. 2006. "Performance Measurement. An Effective Tool for Government Accountability? The Debate Goes On." *Evaluation* 12(2), 219-235.

De Lancer Julnes, P. and M. Holzer. 2001. "Promoting the Utilization of Performance Measures in Public Organizations: an Empirical Study of Factors Affecting Adoption and Implementation." *Public Administration Review* 61, 693-708.

De Vries, J. 2010. "Is New Public Management Really Dead?" *OECD Journal on Budgeting* 2010/1.

Dolowitz, D. and D. Marsh. 2000. "Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making." *Governance: An International Journal of Policy and Administration* 13(1), 5-24.

Drechsler, W. 2005. "The Re-Emergence of 'Weberian' Public Administration after the Fall of New Public Management: The Central and Eastern European Perspective." *Halduskultuur-Administrative Culture* 6, 94-108.

Dunleavy, P., Margetts, H., Bastow, S. and J. Tinkler. 2006. "New Public Management Is Dead – Long Live Digital-Era Governance." *Journal of Public Administration Research and Theory* 16 (3), 467-494.

Feinstein, O. N. 2002. "Use of Evaluations and the Evaluation of their Use." *Evaluation* 8(4), 433-439.

Gedeon, J. A. and R. E. Rubin. 1999. "Attribution Theory and Academic Library Performance Evaluation." *The Journal of Academic Librarianship* 25(1), 18-25.

- Green, S. G., and T. R. Mitchell. 1979. "Attributional Processes of Leaders in Leader-Member Interactions." *Organizational Behavior and Human Performance* 23, 429-458.
- Greene, J. C. 1999. "The Inequality of Performance Measurements." *Evaluation* 5(3), 160-172.
- Halachmi, A. 2005. "Performance Measurement: Test the Water Before You Dive In." *International Review of Administrative Sciences* 71(2), 255-266.
- Hammerschmid, G., Van de Walle, S. and V. Stimac. 2013. "Internal and External Use of Performance Information in Public Organizations: Results from an International Survey." *Public Money & Management* 33(4), 261-268.
- Hatry, H. P. 2013. "Sorting the Relationships Among Performance Measurement, Program Evaluation, and Performance Management." In Nielsen, S. B. and D. E. K. Hunter (eds) *Performance Management and Evaluation. New Directions for Evaluation* 137, 19-32.
- Holzer, M. and K. Yang. 2004. "Performance Measurement and Improvement: An Assessment of the State of the Art." *International Review of Administrative Sciences* 70(1), 15-31.
- Heider, F. 1958. *The Psychology of Interpersonal Relations*. New York: Wiley.
- Hood, C. 2007. "Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles?" *Public Money & Management* 27(2), 95-102.
- Johnsen, Å., Meklin, P., Oulasvirta, L. and J. Vakkuri. 2001. "Performance Auditing in Local Government: An Exploratory Study of Perceived Efficiency of Municipal Value for Money Auditing in Finland and Norway." *European Accounting Review* 10(3), 583-599.
- Kelley, H. H. 1973. "The Processes of Causal Attribution." *American Psychologist* 28(2), 107-128.
- Kelly, J. M. and D. Swindell. 2002. "A Multiple-Indicator Approach to Municipal Service Evaluation: Correlating Performance Measurement and Citizen Satisfaction across Jurisdictions." *Public Administration Review* 62(5), 610-621.
- Lægreid, P., Roness, P. G. and K. Rubecksen. 2006. "Performance Management in Practice: The Norwegian Way." *Financial Accountability and Management* 22(3), 251-70.

- Lonsdale, J. 2000. "Developments in Value-For-Money Audits: Impacts and Implications." *International Review of Administrative Sciences* 66, 73-89.
- Lonsdale, J. 2011. "Introduction." In J. Lonsdale, P. Wilkins and T. Ling (eds) *Performance Auditing: Contribution to Accountability in Democratic Government*. Cheltenham: Edward Elgar, 1-21.
- MacKenzie, M. and A. Blamey. 2005. "The Practice and the Theory: Lessons from the Application of a Theories of Change Approach." *Evaluation* 11(2), 151-168.
- Martinko, M. J., Harvey, P. and M. T. Dasborough. 2011. "Attribution Theory in the Organizational Sciences: A Case of Unrealized Potential." *Journal of Organizational Behavior* 32, 144-149.
- Martinez, A. D., Martinko, M. J. and G. R. Ferris. 2012. "Fuzzy Attribution Styles." *Journal of Leadership & Organizational Studies* 19(1), 17-24.
- Mayne, J. 2001. "Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly." *Canadian Journal of Program Evaluation* 16(1), 1-24.
- Mayne, J. 2007. "Challenges and Lessons in Implementing Results-Based Management." *Evaluation* 13(1), 87-109.
- Meyer-Sahling, J.-H. and K. Yesilkagit. 2011. "Differential Legacy Effects: Three Propositions on the Impact of Administrative Traditions on Public Administration in Europe East and West." *Journal of European Public Policy* 18(2), 311-22.
- Miller, H. T. and C. J. Fox. 2007. *Postmodern Public Administration*. New York: M.E. Sharpe.
- Nakrošis, V. 2008. "Reforming Performance Management in Lithuania: Towards Results-Based Management." In Peters, G. (ed.) *Mixes, Matches, and Mistakes: New Public Management in Russia and the Former Soviet Republic*. Budapest: OSI/LGI, 53-116.
- Newcomer, K. E. 2007. "Measuring Government Performance." *International Journal of Public Administration* 30, 307-329.
- Nielsen, S. B. and N. Ejler. 2008. "Improving Performance? Exploring the Complementarities between Evaluation and Performance Management." *Evaluation* 14(2), 171-192.

Nõmm, K. and T. Randma-Liiv. 2011. "International Organizations As Performance Management Promoters: the Case of the OECD Recommendations for Estonia." Paper presented at the EGPA conference in Bucharest, 7-9 September 2011.

OECD. 2007. *Performance Budgeting in OECD Countries*. Paris: OECD.

OECD. 2011. "Estonia. Towards a Single Government Approach." Public Governance Reviews, OECD, Paris.

Perrin, B. 1999. "Performance Measurement: Does the Reality Match the Rhetoric? A Rejoinder to Bernstein and Winston." *American Journal of Evaluation* 20(1), 101-11.

Peters, G. 2008. *Mixes, Matches, and Mistakes: New Public Management in Russia and the Former Soviet Republic*. Budapest: OSI/LGI.

Pollitt, C. 2006. "Performance Information for Democracy: The Missing Link?" *Evaluation* 12(1), 38-55.

Pollitt, C. and G. Bouckaert. 2011. *Public Management Reform: A Comparative Analysis - New Public Management, Governance, and the Neo-Weberian State*. New York: Oxford University Press.

Pollitt, C. and H. Summa. 1999. "Performance Audit and Public Management Reform." In C. Pollitt, X. Girre, J. Lonsdale, R. Mul, H. Summa and M. Waerness (eds) *Performance or Compliance? Performance Audit and Public Management in Five Countries*. New York: Oxford University Press, 1-10.

Put, V. 2011. "Norms Used: Some Strategic Considerations from the Netherlands and the UK." In J. Lonsdale, P. Wilkins and T. Ling (eds) *Performance Auditing: Contributing to Accountability in Democratic Government*. Cheltenham: Edward Elgar Publishing, 75-94.

Put, V. and G. Bouckaert. 2011. "Managing Performance and Auditing Performance." In Christensen, T. and P. Lægveid (eds) *The Ashgate Research Companion to New Public Management*. Surrey, Burlington: Ashgate Publishing, 223-236.

Radin, B. A. 2009. *Challenging the Performance Movement: Accountability, Complexity, and, Democratic Values*. Washington, D.C.: Georgetown University Press.

Radnor, Z. 2008. "Hitting the Target and Missing the Point? Developing an Understanding of Organizational Gaming." In W. Van Dooren and S. Van de Walle (eds) *Performance Information in the Public Sector: How It Is Used*. Basingstoke: Palgrave Macmillan, 94-105.

Raudla, R. 2012. "The Use of Performance Information in Budgetary Decision-Making by Legislators: Is Estonia Any Different?" *Public Administration* 90 (4), 1000-1015.

Raudla, R. 2013. "Pitfalls of Contracting for Policy Advice: Preparing Performance Budgeting Reform in Estonia." *Governance: An International Journal of Policy and Administration* 26(4), 605-629.

Raudla, R. and R. Savi. 2015. "The Use of Performance Information in Cutback Budgeting." *Public Money and Management* 36(6), 409-416.

Reichborn-Kjennerud, K. 2013. "Political Accountability and Performance Audit: the Case of the Auditor General in Norway." *Public Administration* 91(3), 680-695.

Rossi, P., Lipsey, M. W. and H. E., Freeman. 2004. *Evaluation: A Systemic Approach*. California: Sage Publications, Inc.

Staw, B. M. 1975. "Attribution of the "Causes" of Performance: A General Alternative Interpretation of Cross-Sectional Research on Organizations." *Organizational Behavior and Human Performance* 13, 414-432.

Talbot, C. 2007. "Performance Management." In E. Ferlie, L. Lynn and C. Pollitt (eds) *The Oxford Handbook of Public Management*. New York: Oxford University Press Inc., 491-517.

Talbot, C. 2010. *Theories of Performance: Organizational and Service Improvement in the Public Domain*. Oxford, New York: Oxford University Press.

United States General Accounting Office. 1997. *Managing for Results: Analytic Challenges in Measuring Performance*. GAO/HEHS/GGD-97-138. Washington, DC: United States General Accounting Office.

United States General Accounting Office. 2011. *Performance Measurement and Evaluation: Definitions and Relationships*. Washington, DC: United States General Accounting Office.

Van de Walle, S. 2009. "International Comparisons of Public Sector Performance. How to Move Ahead?" *Public Management Review* 11(1), 39-56.

Van Dooren, W. 2008. "Nothing New Under the Sun? Change and Continuity in the Twentieth-Century Performance Movements." In W. Van Dooren and S. Van de Walle (eds) *Performance Information in the Public Sector: How It Is Used*. Basingstoke: Palgrave Macmillan, 11-23.

Van Dooren, W., Bouckaert, G. and J. Halligan. 2010. *Performance Management in the Public Sector*. London: Routledge.

Van Looche, E., and V. Put. 2011. "The Impact of Performance Audits: A Review of Existing Evidence." In J. Lonsdale, P. Wilkins and T. Ling (eds) *Performance Auditing: Contribution to Accountability in Democratic Government*. Cheltenham: Edward Elgar, 175-208.

Verheijen, T. 2003. "Public Administration in Post-communist States." In G.B. Peters and J. Pierre (eds) *Handbook of Public Administration*. London: Sage, 489-99.

Weiner, B. 1986. *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag.

White, H. 2010. "A Contribution to Current Debates in Impact Evaluation." *Evaluation* 16(2), 153-164.

Wholey, J. 1996. "Formative and Summative Evaluation: Related Issues in Performance Measurement" *Evaluation Practice* 17(2), 145-9.

World Bank. 2003. "Understanding Public Sector Performance in Transition Countries: An Empirical Contribution." Document of the World Bank.

World Bank. 2006. "EU-8 Administrative Capacity in the New Member States: The Limits of Innovation?" Report Number: 36930-GLB.

SUMMARY IN ESTONIAN

Omistamise probleem tulemuslikkuse hindamisel avalikus sektoris: õppetunnid tulemusaudititest Eestis

Kuigi avaliku sektori tegevuse tulemuslikkuse kohta info kogumine on sama vana kui avalik haldus (Hood 2007; Pollitt ja Bouckaert 2011), tõsis tulemuste hindamine enneolematult keskseks teemaks seoses Uue haldusjuhtimise (*New Public Management*) reformide ja tulemusjuhtimise propageerimisega avalikus sektoris (Van Dooren 2008). Kaasaegne erialane debatt on keskendunud eelkõige küsimustele, mis eesmärgil tulemusi mõõdetakse ning kas ja kuidas tulemusinfo üldse kasutatakse. Selgub, et tegelikult panustatakse tulemusinfo kasutamisse hoopis vähem kui selle tootmisse ehk tulemuste mõõtmisse (loe lähemalt I; Lægread et al. 2006; Pollitt and Bouckaert 2011).

Üks oluline põhjus, miks tulemusinfo ei leia oodatud määral kasutamist, tuleneb tulemuste hindamise olemuslikust probleemist – suutmatusest selgelt ja üheselt arusaadavalt defineerida avaliku sektori tegevuse eesmärgid ehk tulemusi, mille täitmist peaks mõõtma. Teine probleem, miks tulemusinfo leiab oodatust vähem kasutamist, seondub tulemuste mõõtmise tehnilise (Currstine 2005; Mayne 2007), analüütilise (United States General Accounting Office 1997; Bovaird 2014) või metodoloogilise (Davies 1999) väljakutsega tõestada põhjuslikku seost valitsuse tegevuse ja mõõdetud tulemuste vahel. Isegi kui avaliku sektori tegevuse tulemuste hindamiseks suudetakse välja mõelda sobilikud indikaatorid ning neid mõõta, siis määratleda, milline panus oli just konkreetsel hinnataval tegevusel või programmil saavutatud tulemustesse, on tihtipeale väga keeruline. Teha vahet põhjuslikkusel, korrelatsioonil ja lihtsalt juhuslikkusel on meetodiliselt raske, vahel ka võimatu. Seda nimetataksegi omistamise probleemiks tulemuste hindamisel. Omistamise probleemi on peetud peamiseks riskiks või nõrkuseks tulemuste hindamisel avalikus sektoris (Cuganesan et al. 2014). Senises tulemusjuhtimise ja tulemuste mõõtmise kirjanduses tunnistatakse omistamise probleemi küll oluliseks, kuid sellegipoolest ei leia see enamasti põhjalikumat käsitlemist. Veel vähem pööratakse tähelepanu omistamise probleemi laiematele tagajärgedele.

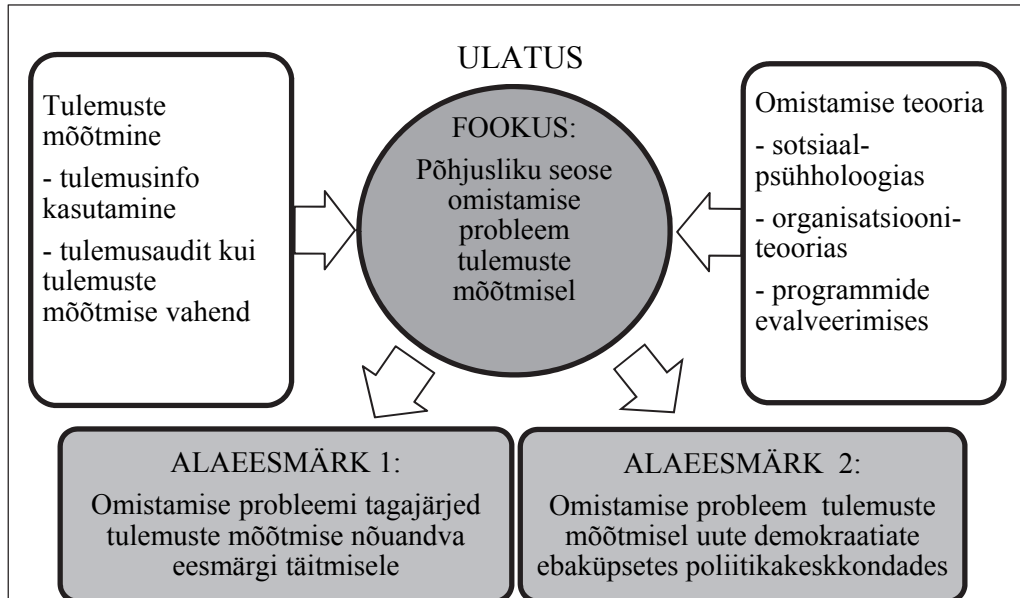
Käesolev väitekiri põhineb definitsioonil, mille kohaselt tulemusjuhtimine on selline juhtimisviis, kus infot tulemuslikkuse kohta kasutatakse juhtimisotsuste tegemiseks, kusjuures tulemusinfo loomine ja selle kasutamine toimub kindlaks kujunenud tulemuste mõõtmise protsessi raames (Radnor 2008, 95; I). See on rutiinne protsess sisendite, väljundite, mõju ja nende vaheliste seoste hindamisel (Newcomer 2007). Kuigi tulemusi on võimalik mõõta ka üksikisiku ja organisatsiooni tasandil, analüüsib käesolev väitekiri tulemuste mõõtmisega seotud küsimusi poliitika või programmi tasandil.

Väitekiri keskendub peamiselt omistamise probleemi tagajärgedele tulemusinfo kasutamisel ja sellel on kaks alaeesmärki. Esiteks näidata, kuidas omistamise probleem takistab eelkõige tulemuste hindamise nõuandva funktsiooni täitmist, sest sellisel juhul on põhjusliku seose olemasolu tegevuse ja tulemuse vahel kõige kriitilisem. Teine alaeesmärk uurib tulemuste mõõtmise algatuste rakendamist spetsiifiliselt Eesti kontekstis ja selgitab, miks omistamise probleem võimendub vähem arenenud poliitikakujundamise oskuste ja traditsioonidega uutes demokraatias (vt joonis 1).

Teoreetiline osa hõlmab avaliku halduse kirjandust, mis keskendub eeskätt tulemuste mõõtmise eesmärkide ning tulemusinfo kasutamisele. Selleks, et täita tühimikke senises diskussioonis just omistamise probleemi alge ja tagajärgede seletamisel, kombineeritakse avaliku halduse distsipliinis põhjusliku seose otsimise käsitlust teistes sotsiaalteadustes enam analüüsimist leidnud omistamise teooriat puudutava kirjandusega. Eeskätt on omistamise teooria tuntud sotsiaalsühholoogiast ning leidnud rakendamist näiteks organisatsiooniteoorias juhtimise uurimisel ning programmide evalveerimise kirjanduses, mis on tulemuste hindamise uurimissuunale päris lähedane. Sotsiaalsühholoogiast pärit omistamise teooria aitab seletada, et inimpsühholoogiale on loomult omane otsida põhjuslikke seoseid ja omistada tulemustele seletatavaid põhjuseid, ka juhtudel, kui need seosed puuduvad (Heider 1958, Kelley 1973, Weiner 1986). Inimesed kasutavad tulemusinfot nii omaenda edu ja ebaedu kui ka sotsiaalsete nähtuste seletamisel subjektiivselt. Selline põhjusliku seose tõlgendamine mõjutab inimeste käitumist ja motivatsiooni. Teiste sotsiaalteaduste distsipliinide tähelepanekuid kasutades on võimalik saada kasulikke õppetunde ka avalikus halduses omistamise probleemiga tegelemiseks.

Väitekirjas näitlikustatakse omistamise probleemi erinevaid tahke empiirilise materjaliga, mis pärineb erinevatest valitsuse dokumentidest ja aruannetest, Riigikontrolli tulemusaudititest (**I**, **III**) ning doktoritöö raames läbi viidud rahvusvahelisel võrdleval meetodikal põhineval tulemusauditite mõju analüüsist (**II**).

Joonis 1: Väitekirja ulatus ja fookus



Allikas: autor

Kokkuvõttes otsitakse väitekirjas vastust kolmele uurimisküsimusele:

- Mis on põhjusliku seose omistamise probleemi tagajärjed tulemuste mõõtmisele avalikus sektoris?
- Kuidas takistab omistamise probleem eelkõige nõuandva funktsiooni täitmist tulemuste hindamisel?
- Kuidas mõjutab omistamise probleem tulemuste mõõtmist ja tulemusinfo kasutamist just ebaküpse poliitikakujundamise keskkonnaga uutes demokraatiates?

Väitekiri põhineb kolmele algupärasele teadusartiklile:

- 1) “Performance measurement and performance information in new democracies: A study of the Estonian central government” (I) (kaasautor Tiina Randma-Liiv) on ülevaateartikkel, mis uurib avaliku sektori tulemuslikkuse mõõtmise praktikaid Eesti keskvalitsuses alates 1991. aastast. Artikkel toob välja, et ebaküpse poliitikakujundamise keskkonna tunnused nagu administratiivne ebastabiilsus, nõrk strateegilise planeerimise ja poliitikaanalüüsi võimekus, vajakajäämised poliitikate rakendamisel ning mitteteadlik poliitikate ülevõtmine, muudab avaliku sektori tulemuslikkuse mõõtmise uutes demokraatiates raskemaks ülesandeks kui väga kõrgelt arenenud riikides.

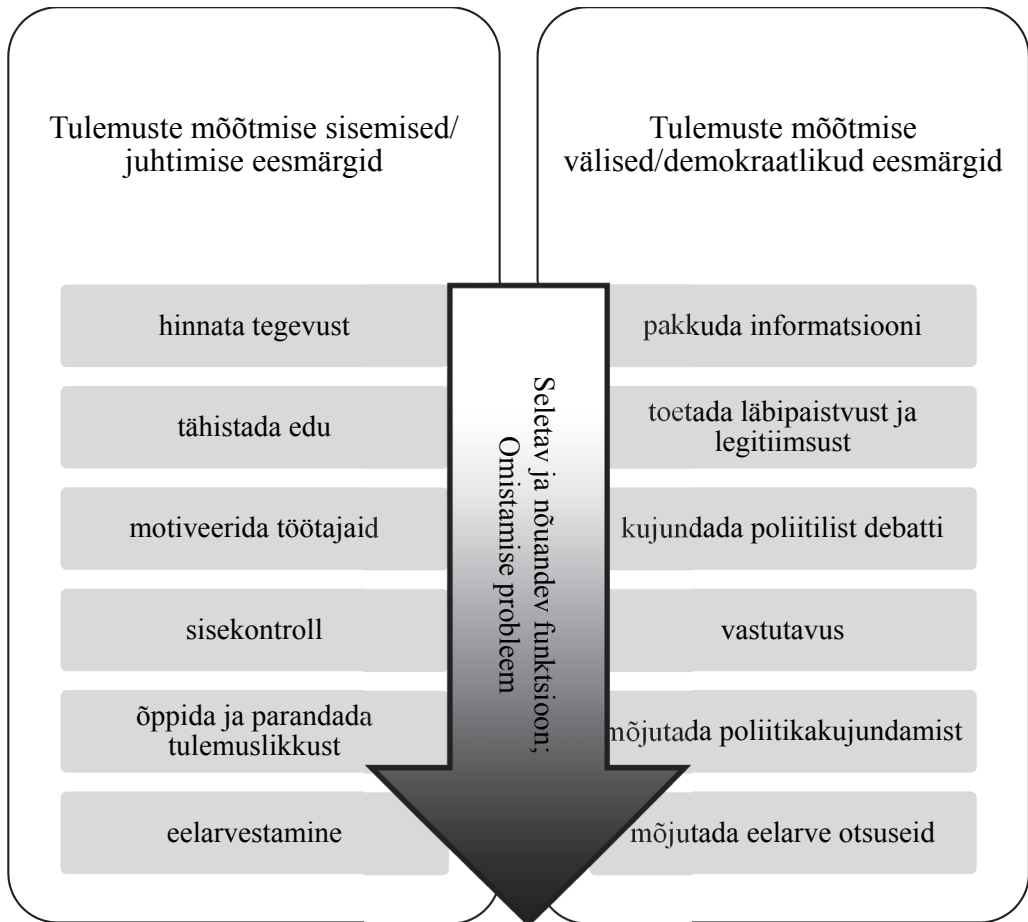
- 2) “The Impact of Performance Audit on Public Sector Organizations: The Case of Estonia” (II) (kaasautorid Ringa Raudla, Cherlin Agu and James W. Douglas) lahkab, kuidas tulemusaudit kui üks tulemuslikkuse mõõtmise vahend ise panustab avaliku sektori tulemuslikkuse parandamisse. Artikkel põhineb koostöös Põhjamaade teadlastega välja töötatud võrdleval meetodikal ja originaalsel analüüsil, mis hindab, kas tulemusauditiga kokku puutunud avaliku sektori töötajad tajuvad, et auditid olid kasulikud ja viisid muutusteni poliitikakujundamises või organisatsiooni praktikates. Eestis läbi viidud küsitlus näitas, et avaliku sektori töötajad peavad tulemusauditeid küll kasulikuks, kuid need ei vii tihti muutusteni senistes poliitikates või tegevustes.
- 3) “Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite Sports Policy: The Case of Estonia” (III) (kaasautor Eerik Hanni) demonstreerib, et Eesti suhteliselt kõrge positsioon rahvusvahelistes medalite edetabelites ei ole kooskõlas rahvusvahelises poliitikauuringus antud kesise hinnanguga Eesti tippspordi süsteemile. Seega kõrged sportlaste tulemused on saavutatud vaatamata vajakajäämistele spordipoliitikas ja edu aluseks võivad olla hoopis muud tegurid kui riiklik spordipoliitika. Artikkel viitab, et ebaselge põhjuslik seos Eesti spordipoliitika ja sportlaste tulemuste vahel muudab õppimise Eesti spordisüsteemist ja tulemuste hindamise nõuandva eesmärgi täitmise keeruliseks.

Avalikus sektoris võib tulemuste mõõtmise eesmärgid jagada laias laastus kaheks: sisemised ehk juhtimise funktsioonid ning välimised ehk demokraatlikud funktsioonid (vt joonis 2). Eesmärgiks võib olla lihtsalt kirjeldada olukorda, pakkuda informatsiooni saavutatud tulemuste kohta, ja sellisel juhul ei olegi põhjusliku seose olemasolu ja selle tõestamine väga oluline. Selline info avaldamine aitab kaasa avaliku sektori avatusele, toetab valitsuse tegevuse legitiimsust, võimaldab tähistada edu või tõsta teatud teemasid poliitilisse agendasse. Ent juhtudel, kui muutub tähtsaks edu või ebaedu põhjuste selgitamine ja oodatakse nõu andmist, kuidas paremate tulemuste saavutamiseks edasi toimida või milliseid eelarvelisi otsuseid teha, muutub põhjusliku seose olemasolu valitsuse tegevuse ja saavutatud tulemuste vahel kriitiliseks. Näiteks ei ole võimalik Eesti tippsportlaste edu pinnalt soovitada Eesti riiklikku spordipoliitikat teistes riikides üle võtta, sest tõenäoliselt on senise edu aluseks hoopis muud või lisaks muud tegurid kui riiklik poliitika (nt traditsioonid, eraalgatuslikud tiimid jne) (III).

Väitekiri näitab, et vastavalt sellele, mis eesmärgil tulemusi mõõdetakse, tuleb tähelepanu pöörata ka põhjusliku seose olemasolule ja selle tõestamisele tulemuste hindamisel. Omistamise probleem on küll eelkõige tehniline või meetodiline väljakutse, kuid see ilmneb just tulemusinfo kasutamisel. Ootused, et tulemuste mõõtmine täidab üheaegselt mitut eesmärki, näiteks võtta valitsus vastutusele ning mõjutada eelarve otsuseid, seavad meetodiliselt täiesti uued väljakutsed. Ja mida enam keskendutakse mitte üksnes väljundi, vaid ka laiema mõju hindamisele, seda olulisemaks muutub omistamise probleem. Ilma omistamise probleemi lahendamata ei

ole seega võimalik anda adekvaatset hinnangut valitsuse tegevuse tulemuslikkusele või jagada nõu teadmistepõhiseks poliitikakujundamiseks. (I; III)

Joonis 2. Tulemuste mõõtmise erinevate eesmärkide ja omistamise probleemi seose olulisus



Allikas: autor; Behn 2003, Halachmi 2005, Pollitt 2006, Talbot 2007, Pollitt ja Bouckaert 2011 põhjal

Vajadus selgitada, kas tulemusindikaatorite muutuse on põhjustanud valitsuse tegevus või teised olulised tegurid, on mõjutanud ka tulemuste mõõtmise meetodika arengut. Kui algselt oli näiteks tulemusauditi peamiseks ülesandeks üksnes objektiivselt kirjeldada, kas seatud tulemuseesmärgid on saavutatud või mitte, siis üha enam oodatakse, et tulemusaudit selgitab välja muutuste toimumise või mitte ilmumise põhjused ning annab hinnangu poliitikamuutuse seosele valitsuse tegevusega (Furubo 2011). Isegi kaasaegne tulemusauditi definitsioon on seetõttu teisenenud. Pelgalt säästlikkuse, tõhususe ja mõjususe hindamisest on liigutud enam „diagnoosivate“

auditite poole, mis seletaks ka valitsuse tegevuse ja poliitikamuutuste ning poliitikate mõju vahelisi põhjuslikke seoseid (Put 2011). Tagajärjena on tulemusauditid läinud metoodiliselt järjest keerulisemaks ning põhjalikumaks. Kui algselt põhines tulemuste hindamine suuresti kvantitatiivsete indikaatorite mõõtmisele, siis üha enam võetakse lisaks appi kvalitatiivseid sotsiaalteaduslikke meetodeid (Lonsdale 2011). (II, III)

Kui programmide evalveerimise raportitele heidetakse ette, et need sisaldavad liiga palju detailset infot ning tihtipeale lõpevad siiski tõdemusega, et „tuvastatud muutuseid pole võimalik omistada antud projekti tegevustele“ (White 2010), siis tavapärase mõõdikutel põhinev tulemuste hindamine ei ole olnudki suunatud selgitamisele, mis täpselt on põhjustanud indikaatorite muutuseid. Eesmärgiks on olnud kiire ja ajakohase info pakkumine. Piiratud ajakava tingimustes ei olegi võimalik põhjuslikke seoseid hinnata. (Davies 1999; Mayne 2007; Newcomer 2007) Akadeemilises kirjanduses juhitakse tähelepanu, et omistamise probleemile reageerimise tõttu on kaasajal tegelikkuses vähenenud ka erinevused programmide hindamise ja tulemuste mõõtmise distsipliinide vahel (Nielsen ja Ejler 2008).

Tulemuste mõõtmine võidab ühelt poolt arenenud IT süsteemidest, andmete paremast kättesaadavusest ja analüütilisest töödeldavusest, kuid suuremad ootused põhjuslike seose väljaselgitamisele sunnib üha enam pöörduma aega ja spetsiifilist ekspertiisi nõudvate analüüside poole. Ent mõnikord ei olegi võimalik keeruliste ja aeganõudvate meetodite kasutamisele vaatamata põhjuslikku seost kindlaks määrata. Saab küll informatsiooni olukorra kohta, kuid pole võimalik tõsikindlalt väita, et näiteks senine poliitika on olnud edukas või valitsuse tegevus viinud tuvastatud muutusteni. Muutuste põhjused võivad peituda hoopis laiemates sotsiaalmajanduslikes protsessides, makromajanduslikes seaduspärasustes või demograafilistes muutustes, mis omakorda on tulenevad näiteks elustiili muutustest, tehnika arengust jne. Sellisel juhul tuleb hoiduda järelduste tegemisest eesmärgil, mis eeldaks põhjusliku seose olemasolu (vt joonis 2). (I; II; III)

Üks võimalus saada üle analüütilisest väljakutsest omistada tulemusi konkreetsele tegevusele või poliitikale, on pöörduda tagasi sotsiaalpsühholoogia teadmiste ehk omistamise teooria juurte juurde ning leppida, et tulemusinfo kajastab üksnes hindajate taju või arvamust, mitte objektiivset reaalsust (vt Staw 1975; Gedeon ja Rubin 1999; Newcomer 2007). Näiteks tulemusauditite enda mõju hindamine on problemaatiline. Saab küll kindlaks teha auditi lõppedes toimunud muutused valitsuse tegevustes, kuid pole võimalik lõpuni eristada, kas muutused toimusid just auditi või muude tegurite tagajärjena. Kuid viimastel aastakümnetel järjest enam levinud tulemusauditi aktiivne läbiviimine avalikus sektoris eeldab, et jõutakse veendumusele ka selle auditeerimise ehk tulemuste mõõtmise enda mõjukuses (II; III; Lonsdale 2010). Väitekiri sisaldab analüüsi, mis hindas just auditeeritute hinnangut küsimustele, kas tulemusauditid olid kasulikud ning kas need viisid muutustele. Kui 40% auditeeritavatest arvasid, et tulemusauditid on olnud kasulikud, siis üksnes 21% leidsid, et auditid on viinud ka muutusteni auditeeritud organisatsioonis või valitsuse tegevustes. Ent samas põhjused,

miks auditi soovitusi viiakse ellu või mitte, ei tulene ainult auditist, vaid on seotud ka laiema poliitilise keskkonnaga ja teiste osapoolte (nt parlamendi ja meedia) tegevusega. (III)

Lisaks tulemusinfo kasutamise eesmärgile mõjutab omistamise probleemi ilmumise tõenäosust või olulisust ka keskkond, kus tulemuste mõõtmine toimub. Edukas tulemuste mõõtmise süsteem avaliku sektoris eeldab hästi toimivat andmete ja kasutatavate meetodite kvaliteedi kontrolli ning sisendite, väljundite ja mõju hindamist regulaarselt ning pika aja jooksul (Newcomer 2007; United States General Accounting Office 2011; Hatry 2013). Samuti eeldab indikaatorite põhiselt tulemuslikkuse hindamine, et põhjuslik seos valitsuse tegevuse ja mõõdetud tulemuse või poliitika ja selle mõju vahel on selgeks tehtud enne hindamissüsteemi juurutamist. Teaduskirjanduses rõhutatakse, et tulemuslikkuse mõõtmisega saavad tegeleda need riigid, kellel on palju aega, raha ja ekspertiisi ning kelle poliitiline ja administratiivne keskkond on stabiilne. Juhtidele seab tulemustel põhineb hindamissüsteem suuri analüütilisi, poliitilisi ja kommunikatiivseid väljakutseid. See on midagi palju enam kui üksnes tehniline tegevus. (vt näiteks Davies 1999; Newcomer 2007; Nilesen ja Ejler 2008; Hatry 2013; Cuganesan et al 2014) Kui poliitikakujundajate ettevalmistuses ja kogemuses on puudujäägid ning metoodiline kvaliteedikontroll ei ole tugev, õõnestab see võimalusi edukaks tulemuste mõõtmiseks ja sellel infol põhinevaks juhtimiseks avalikus sektoris. Seetõttu vajab poliitiline ja administratiivne kontekst, kus tulemuste hindamist juurutatakse, erilist tähelepanu. (I)

1990. aastate algul päris Eesti, nagu paljud teised postkommunistlikud riigid, haldusorganisatsiooni, millel puudus oskus ja kogemus iseseisvaks poliitikakujundamiseks, koordineerimiseks või hindamiseks. Sellele lisandusid sagedased muutused poliitilises ja administratiivses juhtkonnas. Paljud reformid ja arendustegevused rakendati ilma põhjaliku eelneva analüüsita ning järelanalüüsile ja kogemustest õppimisele pole ka piisavalt tähelepanu pööratud. Erinevad lühiajaliseks initsiatiivid lõppesid enne kui oleks võimalik üldse adekvaatselt nende tulemusi hinnata. Analüütiline võimekus ja oskus saadud tulemusinfot sihipäraselt kasutada oli kesine. (I; Verheijen 2003; World Bank 2006; Nõmm ja Randma-Liiv 2011; Raudla 2012; Raudla ja Savi 2015)

Vaatamata ebasobivatele keskkonna eeldustele ehk faktoritele (administratiivne ebastabiilsus, nõrk strateegilise planeerimise ja poliitikaanalüüsi võimekus, vajakajäämised poliitikate rakendamisel ning mitteteadlik poliitikate ülevõtmine) on Eesti sarnaselt teistele postkommunistlikele riikidele paari viimase aastakümne jooksul aktiivselt rakendanud erinevaid tulemusjuhtimise ja sealhulgas tulemuste mõõtmise algatusi. Kuigi väitekirjas hoidutakse järelduste tegemisest teiste sarnase ajaloolise taustaga riikide kohta, sest iga administratiivne süsteem ja poliitikakujundamise traditsioonid on omanäolised, on välja toodud faktorid iseloomulikud ka teistele Kesk- ja Ida-Euroopa riikidele (see Verheijen 2003; Nakrošis 2008; Meyer-Sahling and Yesilkagit 2011).

Oluline roll avaliku sektori tulemuslikkuse mõõtmise algatuste sisseviimisel on olnud ka välisel mõjul, rahvusvahelistel organisatsioonidel ja konsultantidel (nt OECD, Maailmapank, Rahvusvaheline Valuutafond, Euroopa Liit), kes aktiivselt, kuid tihtipeale konkreetset konteksti silma pidamata, propageerivad tulemusjuhtimise rakendamist. (I; Peters 2008; Bouckaert et al. 2011; Nõmm ja Randma-Liiv 2011) Seejuures unustatakse, et ka kõrgelt arenenud ja tugeva haldusorganisatsiooniga riikides pole tulemuste mõõtmise rakendamine alati viinud tulemusinfo eduka kasutamiseni või veelgi enam, riigi tegevuse tulemuslikkuse paranemiseni. Tulemusinfo võib olla küll olukorrast ülevaate saamiseks vajalik ning erinevatele poliitikaprotsessi osapooltele enda argumentide toetamiseks kasulik, ent see ei garanteeri iseenesest tulemuslikumat juhtimist või mõjusamaid poliitikaprogramme. (I; II; III; Pollit and Bouckaert 2004; Newcomer 2007; Talbot 2007; Radnor 2008; Van de Walle 2009)

Kokkuvõttes seab edukas tulemuste mõõtmine ja adekvaatne tulemusinfo kasutamine teatud tingimused riigi poliitilisele süsteemile ja administratiivsele võimekusele. Kui aga rakendada tulemuslikkuse hindamist avalikus sektoris, kus eeltingimused analüütilistele ja rakenduslikele oskustele, kvaliteedikontrollile ja stabiilsele poliitikakujundamisele pole täidetud, siis võimendab see nii tulemuslikkuse mõõtmise olemuslikke kui ka analüütilisi puudujääke, nagu üheste ja selgelt defineeritavate eesmärkide puudumine ning omistamise probleem. Seega oht omistamise probleemi tõttu tulemuslikkuse mõõtmisel põhimõttelisi meetodilisi vigu teha ning hiljem saadud infot mittesihipäraselt kasutada on eriti suur just ebaküpse poliitikakujundamise keskkondadega uutes demokraatiates.

ACKNOWLEDGEMENTS

This dissertation has greatly benefited from the support of many people. First of all, my supervisor, Prof. Dr. Tiina Randma-Liiv, motivated me to start PhD studies and has since continuously encouraged my academic endeavors. I am deeply grateful for her backing on this long journey. Also, all other members of the Ragnar Nurkse School of Innovation and Governance have deeply influenced my academic development and provided an inspiring retreat in the academic world. My dear friend, Prof. Dr. Ringa Raudla, has been a role model and encouraging companion since the beginning of our university studies at the University of Tartu. She has always been there for me when I needed some uplifting words.

I have had excellent opportunity to simultaneously benefit from academic and practical work. For this I owe thanks to many people whom I have met during my professional career. These unique discussions have considerably contributed to my knowledge and undertakings. Colleagues from the National Audit Office of Estonia and Estonian Cooperation Assembly have kindly tolerated my academic engagements and let me take advantage of both academic and practical experience. I have also had the pleasure of working with many international colleagues from other supreme audit offices and universities. These numerous international working projects, seminars and conferences around the world helped to broaden my research perspective and practical experience. I am greatly thankful to the co-researchers from the latest cooperative research project “The influence of performance audit on public administration and politics: A comparative study of the Nordic Supreme Audit Institutions” for their fruitful contribution.

This thesis has benefitted from financial support from many sources; the Estonian Ministry of Education and Research mobility schemes Kristjan Jaak and DoRa, the Nordic Research Council, Tallinn University of Technology among them.

As sometimes the best ideas come up while being out of the office, I am deeply grateful to my friends who introduced me to the equestrian world or accompanied me on the skiing tracks. However, my greatest gratitude goes to my husband who has helped me become a better person.

PUBLICATIONS (Articles I-III)

I Nõmm, Külli and Tiina Randma-Liiv. 2012. "Performance Measurement and Performance Information in New Democracies" *Public Management Review* 14(7), 2-21. (1.1)

Abstract

This article explores the introduction of performance measurement tools in new democracies by presenting a qualitative study based on Estonian governmental documents and performance audits. A set of specific factors help to explain difficulties in introducing performance measurement tools in immature policy environments: instability, poor strategic planning and policy analysis, an implementation gap and uncritical, uninformed transfer of Western performance management initiatives. Although the authors recognize the limits of generalizations based on the Estonian example, the presence of such 'transitional' factors is likely to make the development of performance measurement in new democracies even more complicated than in the West.

Key words

Performance measurement, performance information, new democracies, Estonia, Central and Eastern Europe

PERFORMANCE MEASUREMENT AND PERFORMANCE INFORMATION IN NEW DEMOCRACIES

A study of the Estonian central government

Küllli Nõmm and Tiina Randma-Liiv

Küllli Nõmm

Department of Public Administration
Tallinn University of Technology
Tallinn
Estonia
E-mail: kylli.nommm@riigikontroll.ee

Tiina Randma-Liiv

Department of Public Administration
Tallinn University of Technology
Tallinn
Estonia
E-mail: tiina.randma-liiv@ttu.ee



INTRODUCTION

Since the 1980s, the public sector worldwide has been under constant pressure to improve its performance in pursuit of more efficiency and effectiveness, and in order to revive the citizens' trust in public institutions. Although performance measurement is as old as public administration itself and played a central role in many other models (Pollitt and Bouckaert, 2004: 90), no other public sector reform movement has promoted the use of performance information on a scale comparable to that advocated by 'New Public Management' (NPM) reforms (Van Dooren, 2008). Using private sector management instruments in the public sector as well as the fostering of 'performance culture' or 'performance orientedness' has led to an increase in the use of various performance management tools (Van Dooren *et al.*, 2010). In the literature, the terms 'performance management' and 'performance measurement' are used interchangeably with confusing definitions. This article is based on the definition of performance management as a type of management in which information in relation to performance is used for decision-making (Van Dooren *et al.*, 2010: 30), whereas performance information is generated and utilized through performance measurement routines (Radnor, 2008: 95). We focus on both the supply and the demand sides of performance measurement (Askim, 2007, 2009) by studying why and how government makes performance information available and how this information is actually used in the decision-making.

A number of studies (e.g. Bouckaert *et al.*, 2008; Dunn *et al.*, 2006; Goetz, 2001; Meyer-Sahling and Yesilkagit, 2011; Randma-Liiv *et al.*, 2011; SIGMA, 2009; Verheijen, 2003; World Bank, 2006; Zubek and Goetz, 2010) have already considered various aspects of public management reforms in Central and Eastern European (CEE) countries. Yet, the performance management – and more particularly, performance measurement – component of these reform attempts has not received sufficient academic attention so far. The few studies (e.g. Condrey *et al.*, 2001; Nakrošis, 2008; Nemeč *et al.*, 2008; Randma-Liiv, 2005; Raudla, 2011; Schick, 1998; Tonnisson and Wilson, 2007; Verheijen and Dobrolyubova, 2007; World Bank, 2006) that focus on the introduction of particular performance based management tools are rather descriptive and have ended up with contradictory conclusions. Furthermore, not much attention has been paid to the actual implementation or use of performance information. The article at hand proposes to make up for that deficit.

The aim of this article is not only to describe the introduction of various performance measurement tools in a post-communist country but also to explain how the transitional context influences the adoption of performance measurement instruments and the use of performance information in new democracies. Through exploring the introduction of performance management tools in Estonia, it will be demonstrated that although problems common in performance measurement in Western countries are also present in the Estonian context, they are compounded by specific contextual factors

encountered in post-communist administrations: instability, poor strategic planning and policy analysis, the implementation gap and foreign influence.

This study is based on qualitative exploratory research leading to the elaboration of explanatory variables that contribute to further understanding of performance measurement in new democracies. The theoretical foundation is based on a synthesis of existing literature of performance measurement and of post-communist transition. Taking into account very recent trends in the research of performance management (see Andrews *et al.*, 2006; Askim, 2009; Van Dooren *et al.*, 2010, for example), this study focuses on the purposes of performance measurement, the choice of performance indicators and the practice of performance information use.

Empirical evidence is drawn from the analysis of various governmental documents, the findings of the *Riigikontroll* (National Audit Office of Estonia, the NAO) of Estonia and from personal observations. For more than a decade, the authors of this study have closely followed performance measurement reforms from an academic as well as a practical perspective (the latter through their work for the NAO (as performance audit manager and expert, respectively)). The NAO started to carry out performance audits in accordance with international auditing standards in the late 1990s. This study is based on the analysis of forty performance audits. The World Bank (2005) classification of performance management tools is used for the general framework.

This study has a potential for policy learning. Some of our findings may apply to other CEE countries because of the similarity of post-communist contexts. Although Estonia has been claimed to be 'one of the CEE countries closest to NPM models' (Drechsler, 2004: 391), most post-communist countries have also introduced performance management tools during the last two decades (Bouckaert *et al.*, 2008). Yet, the generalization of Estonian experience should be done with caution because of different historical, political and cultural environments within CEE. It could, however, be argued that the presence of proposed 'transitional' factors is likely to make the development of performance measurement in new democracies even more complicated than in the West.

PERFORMANCE: A CONTROVERSIAL CONCEPT

Alongside the rise of NPM reforms, performance measurement has become more extensive, intensive and external during the last few decades (Pollitt and Bouckaert, 2004). However, public administration scholars tend to approach performance measurement in the public sector as a conceptual problem (Van de Walle, 2009). Efforts to measure government ineluctably meet with the essential problem of defining public sector aims (Van de Walle, 2009). Ironically, performance management does not always lead to better performance and may even result in pervasive effects (see Andersen, 2008; Andrews *et al.*, 2006; Pollitt and Bouckaert, 2004, for example), as performance measurement systems struggle to find a gaming-proof design (Bevan and

Hood, 2006; Radnor, 2008), and there is little evidence on the actual use of performance information (Pollitt, 2006).

To explain this conceptual controversy, recent scientific literature often goes back to the roots of the issue by analysing the aims of performance measurement in the public sector and the actual use of the information, also called ‘demand side’ of the performance management (Askim, 2007). This section provides a brief overview of the state-of-the-art research on performance measurement and performance information use by asking the questions ‘why?’, ‘how?’ and ‘what for?’ The following analysis of the Estonian experience in introducing performance management tools is based on this framework.

Why measure performance?

Performance measurement has an internal and an external function, which are termed, respectively, the managerial and the democratic aspect of performance information, depending on who is regarded as the ‘end user’ of that information (Pollitt, 2006). In many cases, however, the line between the two is vague. Most performance measurement literature deals with the managerial aspects of performance information – with the aim to inform specific decisions, to create benchmarks, to determine budget allocations, to improve communication and feedback and to serve as input for motivation systems, career development and the promotion of individuals. Management decisions are justified in the context of performance because legal or internal standards are not sufficient to guide public sector operations (Behn, 2003).

The external use of performance information is aimed at informing the public (the legislature, other politicians, constituencies, taxpayers and service users) about the performance of public institutions. Publication of performance information is expected to help achieve accountability and control, and to enhance democratic legitimacy and transparency in the public sector (Talbot, 2007). In addition, performance measurement possesses an important symbolic function because of the importance of performance management as a public relations asset (Pollitt and Bouckaert, 2004: 6). For the public and its representatives, performance measurement symbolizes the fact that public managers care about the performance of their staff, which contributes to improving the image and buttressing the legitimacy of the public sector.

The concept of performance measurement looks intuitively so appealing to policymakers that it seems almost ridiculous to question. However, performance measurement in the public sector has received severe academic critique and raised a number of practical issues (e.g. Andersen, 2008; Bevan and Hood, 2006; Halachmi, 1993; Pollitt and Bouckaert, 2004; Radnor, 2008; Talbot, 2007; Van de Walle, 2009; Van Dooren, 2008). These analyses are useful in understanding the problems that result from inaccurate assumptions and implementation difficulties (e.g. over-quantification, oversimplification, emphasizing short-term goals over broader ones, bureaucracy

reinforcement and high transaction costs, problems of attribution and causality and organizational gaming). Contrary to the belief that performance measurement can help to make better and more rational decisions, the aforementioned problems may eventually lead to a drop in the quality of the public service concerned or even to a decrease in public sector efficiency in general.

How is performance measured?

The development of performance indicators is the foundation of any performance measurement system. Several authors (e.g. Bevan and Hood, 2006; Talbot, 2007) have drawn attention to the problem of the great diversity and near intangibility of objectives and performance criteria in the public sector – the goals are often vague, difficult to follow and hard to evaluate quantitatively. Yet, performance measurement favours quantitative indicators that are easily measured and simple to check (Talbot, 2007; Van Dooren, 2008). Numerical targets have an air of objectivity and are thus more appealing to decision-makers than qualitative assessments.

The choice of performance indicators is shaped by social and political mechanisms. These include domestic political and institutional forces, and international exchange and diffusion of experience (Van Stolk and Wegrich, 2008). Setting performance goals and indicators is a process that is intimately linked to specifying and implementing strategic political choices. The choice of indicators is thus dependent on political leadership and overall (strategic) planning considerations. Since strategy planning is also strategy formation (Mintzberg, 1994: 24–5), planners exert a significant impact on how those goals and indicators are implemented. However, it is not only domestic actors and the corresponding political environment but also international institutions and trends that determine how and what is measured (Van Stolk and Wegrich, 2008).

Because of the above-mentioned difficulties and dilemmas, performance information may sometimes prove incomprehensive or unobjective, shown to favour selective measurement, presentation and interpretation, thus serving the interests of the actors involved (Bevan and Hood, 2006).

How is performance information used?

If we want to study the successes and failures of performance movements, we have to study the use of performance information (Van Dooren, 2008: 22). Recent public sector reforms have contributed to collecting enormous amounts of performance information (Behn, 2003; Pollitt, 2006). The underlying assumption to the use of performance information for internal, management purposes is that performance information can be integrated into the decision-making process, leading to better-quality decisions and, ultimately, improved performance. However, existing findings

suggest that the efforts invested in measuring performance considerably outstrip the eagerness shown in using the information gathered (e.g. Lægreid *et al.*, 2006; Pollitt and Bouckaert, 2004). It is rather rare that even the ministries of finance use performance information for action. In OECD countries only 4 per cent of ministries of finance use performance measures to eliminate programmes, 10 per cent to cut expenditure and 11 per cent to determine pay (OECD, 2007). Evidence of active and systematic external use of performance information for democratic purposes is even more limited. Researchers tend to be sceptical about the use of performance information by politicians (Lægreid *et al.*, 2006; Raudla, 2011), although some (e.g. Askim, 2007) have found positive evidence as well.

Evaluating the impact of management reforms has not been a priority of politicians, because there is more popularity to be won by announcing new initiatives than by evaluating previous ones. No performance movement can depoliticize policymaking (Pollitt, 2006; Pollitt and Bouckaert, 2004). Politicians do not make decisions on considerations of economy, efficiency or effectiveness alone – to a great extent, they rely on their values and political views.

Evidence about citizens' interest in performance measurement is even more difficult to pin down. Some authors are positively disposed towards the concept of citizen-driven government performance (Holzer and Yang, 2004), while others argue that there is no direct evidence of reforms that have been undertaken as a response to public opinion and point out that the vast majority of citizens are unlikely to benefit from detailed performance information (Pollitt, 2006; Pollitt and Bouckaert, 2004). If not used, performance information, no matter how valid and detailed, cannot fulfil its function of fostering goal setting, accountability, public sector transparency and legitimacy.

PERFORMANCE MEASUREMENT IN ESTONIA

Estonia regained its independence in 1991 after the collapse of the Soviet Union. Thus, its democratization and institution building process started at the beginning of the 1990s, when the NPM movement in the West was at its peak. The NPM ideology sat well with countries that were abolishing their one-sector economies and carrying out large-scale privatizations. The development of the Estonian public administration has been influenced by the consecutive neo-liberal governments and by the values and ideas of NPM, as NPM ideas have prevailed in various public administration reform concepts and strategies originating in the second half of the 1990s (see, for example, Drechsler, 2004; Tõnnisson and Randma-Liiv, 2008). Anti-statist views within Estonian society, partly influenced by the Soviet legacy (Drechsler, 2004), and the elite's belief in the 'lean' state have paved the way to the development of a decentralized administrative system.

According to The Constitution of the Republic of Estonia (1992), the executive power in Estonia rests with the government, which exercises that power directly through

ministries and government agencies. Executive government institutions include the ministries, the State Chancellery and county governments, as well as executive agencies and inspectorates with the authority to exercise executive power. A ministry is a superior body ranking above executive agencies, inspectorates and other state agencies. Government agencies are accountable to the respective minister, who directs and coordinates their activities. Executive government institutions covered by the civil service employ around 20,000 employees altogether, whereas the workforce of the eleven Estonian ministries counted 2,448 employees in 2009. A typical ministry employs a staff of fewer than 200. Although ministries are small, they represent strong administrative actors that have considerable leverage over the issues belonging to their areas of governance (Sarapuu, 2011). In addition, the European Union (EU) accession process further cemented a decentralized setup and functioning of the Estonian public administration (see more in Sarapuu, 2011). A high level of decentralization has been linked to a great interest in performance management tools as means of steering autonomous organizations.

The following analysis focuses on the development and implementation of performance measurement practices in Estonian central government since 1991. The Estonian government has experimented with most of the internationally known performance management tools (see Table 1). Some tools have already been abandoned by 2011, while others have been integrated into a set of central and regularly employed measures. The application of performance management tools has involved a few horizontal and even compulsory initiatives (e.g. the introduction of detailed action plans with performance targets and pay-for-performance). However, as Estonian public administration is decentralized to a large degree, it is difficult to make generalized assessments about the application of other performance management initiatives. Motives behind such exercises are highly dependent on individual and organizational factors, and tend to be haphazard rather than systematic. On the organizational level, the extent and way of using performance information depend on the internal regulations of the organization, and the corresponding practices are highly diverse.

Why is performance measured?

Performance measurement in the Estonian central government is considered to help make better decisions, to lead to more economic, efficient and effective public administration, to result in improving the quality of public service provision, to make the public sector more transparent and to enhance its legitimacy (*Rahandusministeerium*, 2008).

Estonian public administration has experienced different performance management tools (see overview in Table 1), but there are two remarkable cases in which performance initiatives were implemented throughout the central government – the introduction of pay-for-performance and the strategic planning reform. In 2001, the Ministry of Finance introduced a central pay-for-performance initiative. It was believed

Table 1: Overview of the application of key performance management tools in Estonia

<i>Performance management tool (World Bank 2005)</i>	<i>Examples of application in Estonia</i>
Annual performance reports of ministries	Ministries are required to prepare detailed action plans with annual and 4-year performance targets during the budgeting process and to report on progress towards targets in their annual reports.
Performance measurement methodology	There is no integrated central performance measurement methodology; the Ministry of Finance only provides standard document forms with brief guidelines.
Performance budgeting methodology	Budgeting documents contain performance information. Expenditures are loosely and/or occasionally linked to output or outcome targets. In some cases, performance contracts for public funding are entered into (e.g. public universities are financed according to the number of graduates).
Macroeconomic indicators	These provide information about and projections for a series of key indicators: GDP, budget deficit, inflation, unemployment, exports, productivity, foreign direct investment, etc.
Societal indicators	Government performance measures include a wide range of targets connected to societal indicators such as poverty level, life expectancy, gender equality and societal integration.
International assessments	Some international evaluations and rankings are integrated into domestic targets and reports: e.g. World Values Survey; Freedom House measures for Freedom in the World and Freedom of the Press; Doing Business rankings; Economic freedom indicators; Transparency International's corruption perceptions index.
Service quality measures	In most cases, the indicator measured is the accessibility of public services (length of waiting lists, percentage of population covered by the service, speed of the service).
Outcome/output measures	Performance targets are mostly output measures, usually expressed in units of work or product or completed projects. Several ministries have also attempted to measure outcomes (e.g. employment rate among young people who have completed a vocational education, delinquency rate, repeat imprisonment rate, fatality rate in traffic accidents and fatality rate in fire accidents).
Client surveys	Client surveys are carried out on the initiative of individual organizations; for example, the Estonian Health Insurance Fund conducts annual client surveys among its insured, employees and contractors. A few other agencies, such as the Estonian Tax and Customs Board, have introduced client satisfaction indicators into their performance targets. In most cases, client surveys are not conducted on a regular basis, or their results are intended for internal use only.

(continued)

Table 1: (Continued)

<i>Performance management tool (World Bank 2005)</i>	<i>Examples of application in Estonia</i>
Public opinion polls	Results of public opinion polls are included in the detailed strategic goals of the ministries (e.g. trustworthiness, commitment and competence of civil servants, and trust in the government institutions).
Third party validation	Quality awards for public sector institutions and public services, international assessment rankings (see above).
Individual performance contracts/targets	A central pay-for-performance initiative was launched in 2001 that included setting individual performance targets. Organizational practices differ; individual performance contracts have been applied only in some rather rare cases.
Audit reports	External performance audits by the NAO, occasional performance audits by the ministries' internal audit departments.
Programme evaluations	Programme evaluations are carried out mostly for programmes supported by EU funds. Evaluations for domestically funded programmes are performed on a random basis.

to help achieve better results at no extra cost, to increase public sector transparency and attractiveness, to improve public sector planning and to cut the red tape. Despite the absence of evidence of an improved performance, most of the funds set aside for pay-for-performance were paid out in 2002 – bonuses were even allocated for unfulfilled tasks or released before evaluation reports for the corresponding projects were drawn up. For example, the Ministry of Finance allocated pay-for-performance to 100 per cent of its employees, and performance bonuses were paid to 80–90 per cent of civil servants in the Ministry of Defence and in the Ministry of Highways and Communications (Randma-Liiv, 2005; *Riigikontroll*, 2002). The initiative required a considerable amount of additional paperwork from all government units. After this experiment, the centrally regulated pay-for-performance scheme was discontinued a few years later as it did not achieve its intended aims.

In 2005, the Estonian government reached an agreement regarding the overhaul of its strategic planning system. The new system was founded on performance measurement and thus requires the setting of targets. Before 2005, planning decisions were made pursuant to a series of uncoordinated strategies, announced without legal or financial commitment and often inconsistent in their content. However, many analytical reports, international as well as domestic, still refer to deficiencies in strategic planning in Estonia (e.g. OECD, 2011; *Riigikontroll*, 2008b; World Bank, 2006). Partly in response to these, the Ministry of Finance initiated a project to reform the strategic planning system in 2010. The main aim is to restructure the strategic planning, budgeting and reporting system. The current division of planning responsibilities, in

which each ministry is accountable for its sphere of government, is expected to be replaced by a division into integrated performance areas, i.e. a system consisting of comprehensive fields of activities and responsibilities, each of which is likely to include input from several ministries, resembling a programme budgeting model. Among other aims, this initiative is intended to improve performance measurement and reporting in central government (PRAXIS, 2010). At the time that the current study was conducted, the design of the new system was still in process.

There is wide recognition in Estonia's public institutions of the need to use performance information for public sector legitimization purposes, and for improving the overall image of the public sector. In 2010, only slightly more than half (56 per cent) of poll respondents thought public servants were competent, and only 45 per cent regarded government officials as trustworthy. In 2005, the results were correspondingly 67 per cent and 60 per cent (*Rahandusministeerium*). In spite of public institutions' belief in performance measurement as a means of improving the image of the public sector, there is no evidence to suggest that this has exerted a positive impact on public perceptions of the state. On the contrary, large-scale failures of performance management initiatives might result in a negative impact on public sector legitimacy. For example, in 2005, the so-called 'performance target scandal' sparked a government crisis, causing the Minister of Justice to step down, and later the entire cabinet to resign. The case concerned a set of performance targets that the Ministry of Justice had set to the police, including among other things a target for misdemeanour convictions. This caused a public outcry, being condemned as an instance of Soviet-style 'planned economy' and 'repression' (*Riigikogu*, 2005). The scandal resulted in the public discourse on performance targeting being cast in a negative light for some time.

From the perspective of external aims of performance measurement, international insistence on the need for performance measurement has played a crucial role in Estonia. Having a good international image is essential for a new democracy to win recognition as a trustworthy partner to other members of the international community. When a government realizes this, it accords considerable weight to foreign opinions. In the Estonian case, references to the need to comply with the 'soft' recommendations of international organizations (such as the OECD recommendations to reform higher education financing, for example) are included in the relevant performance targets and in explanatory memoranda laid before the parliament together with new legislative proposals. In addition, various EU targets are incorporated into the domestic targets. There are even cases where the Estonian government has set itself higher targets, or has resolved to achieve a goal in a shorter timeframe than that established in European policy documents. For example, the Estonian government has adopted an ambitious 13.5 per cent as the 2013 target rate of participation in lifelong learning, while the corresponding European target, set in the Lisbon strategy, is 12.5 per cent. No explanations have been given for setting these higher targets (*Riigikontroll*, 2010).

How is performance measured?

The development of a public sector performance measurement system requires a presence of good analytical skills. However, the Estonian NAO found in ten out of fourteen performance audit reports published in 2009 that performance indicators and targets were not based on a sufficient analytical foundation (by that confirming previously found concerns over the quality of performance data, see Pollitt and Bouckaert, 2004: 115). More often than not, the methodology for measuring the targets or impact indicators had not been agreed upon before the start of the policy planning process (see PRAXIS, 2007). The previously mentioned pay-for-performance initiative was also criticized because of the poor analytical basis (Randma-Liiv, 2005).

Although Estonia is sometimes recognized as one of the forerunners in the development of e-governance (e.g. OECD, 2011; World Bank, 2006), and its electronic registers and information systems provide the government with a plethora of information, these possibilities are often underused for performance measurement purposes. In the ministries responsible for setting and measuring performance targets, analytical departments are understaffed or in some cases missing altogether (see also Drechsler, 2004 on Estonian administrative capacity). No central training or methodological support in performance measurement techniques is provided. Thus, it can happen that performance indicators are calculated incorrectly or do not adequately reflect the performance aspect they purport to measure. For example, to calculate the dropout rate of the country's vocational schools, the Ministry of Education and Research applied a formula that resulted in rates which for some groups exceeded 100 per cent (*Riigikontroll*, 2009). Therefore, official statistics of dropout levels in vocational schools, which represent an important input for policy decisions, remained misleading for years.

The political and administrative culture in Estonia strongly supports quantifiable targets, which is also reflected in the manifestos of political parties that include increasingly detailed targets. For example, the leading Reform Party's run-up campaign to the 2007 general parliamentary elections was constructed around the ambitious slogan 'We will take Estonia into the top five wealthiest European countries in 15 years'. Each year, strategic documents contain more and more targets that are quantitative. For example, a strategic plan of the Ministry of Finance (which is the most influential promoter of quantification of targets throughout the central government) for the years 2006–9 included 54 per cent quantitative indicators while the strategic plan for the years 2011–4 has 83 per cent quantitative indicators.

However, targets and reporting standards change, the time series of various indicators remain short, and the accuracy of performance information provided by the public institutions is not evaluated systematically. Without the appropriate analytical, technical and strategic capacity, performance measurement risks to succumb to the pitfalls of oversimplification, neglect of broad perspective, bureaucracy reinforcement and gaming (see Andersen, 2008; Bevan and Hood, 2006; Radnor, 2008; Talbot, 2007;

Van de Walle, 2009). Numerous streamlining efforts as a rule aim to perfect the technical procedures of the Estonian system rather than take up the politically risky task of incorporating qualitative goals into strategic action plans. The impact of excessive quantification of targets on democratic values and the cost of implementing the system has not been properly considered.

How is performance information used?

The development of a performance measurement system is not merely a managerial task – it involves considerable strategic and political considerations. Much depends on whether and how individual policymakers decide to employ the information provided. In Estonia, performance targets are integrated into the annual action plans of the ministries. The targets and information about their achievement (or non-achievement) are expected to constitute an input to the budgeting process. However, the role of performance information in the actual budget negotiations is insignificant as budgetary decisions largely depend on political negotiations (*Riigikontroll*, 2008b).

Western experience with performance targeting shows that because of the perceived need to avoid possibly negative consequences resulting from not meeting a target, a host of problems related to gaming with the performance information appear (e.g. Bevan and Hood, 2006; Radnor, 2008). In Estonia, with some rare exceptions, performance measurement tends to be indirectly related to specific rewards or punishments. A public manager will not lose his or her job because of not meeting a performance target – at best, such failure may be a secondary reason or simply a pretext. Even in cases where, under the terms of a performance contract, future funding could be reduced because a target established in a previous contract was not met – such is often the case in government-funded provision of higher education – the government prefers not to exercise this option (*Riigikontroll*, 2008a). In the case of the central pay-for-performance initiative, most bonus payments were made without any reference to an achievement that they supposedly remunerated (Randma-Liiv, 2005).

If meeting or not meeting a performance target does not entail any tangible consequences, the purpose of performance measurement is understandably called into question. Nevertheless, it would be an ill recommendation to suggest that performance targets should be rigidly linked to specific rewards or punishments, because the failure of policy implementation all too often reflects a failure of goal setting. Because of poor *ex ante* policy analysis, Estonian policy targets have often been found unrealistic, or their achievement has been influenced by actions of political parties rather than public sector institutions (*Riigikontroll*, 2008a, 2009, 2010). In such cases, performance targets run the risk of creating a lock-in effect – rewards and punishments could lead to undesirable consequences instead of helping to achieve the actual policy goal.

The interest of Estonian legislators and politicians in performance information is hectic, mostly driven by short-term political considerations. A few previous studies show that although politicians like to report the use of performance information (see Tonnisson and Wilson, 2007 on benchmarking), they tend to overestimate the actual use of performance information (Nemec *et al.*, 2011), and in practice, legislators make only limited use of the formal documents containing performance information (Raudla, 2011).

The government prefers to release positive information only (*Riigikontroll*, 2011), while the purpose of the opposition is to emphasize failures, and various media editions hunt for headlines. For example, every performance audit report published by the NAO is discussed in the Select Committee for State Budget Control. The personal experience of the author of this article confirms that members of the committee only manage to read the summary of the report, and very rarely the report itself. Opposition politicians are usually much more active in the discussions of NAO reports than representatives of the ruling coalition. And last but not least, drawing up a national budget is a political process informed by considerations extending beyond mere efficiency and effectiveness (Raudla, 2011). For this very reason, it is difficult to integrate performance measurement into the budgeting process.

There are precious few examples of performance information targeted to citizen use by the government. For example, information about waiting lists in hospitals is provided publicly. However, there is no systematic evidence about the extent of use of this kind of information. The performance information that is most widely used by the citizens – school rankings – is compiled by journalists and published in the newspapers, instead of being publicly provided by the government. Similarly to Western experience, evidence about the use that legislators, politicians and citizens make of performance information remains limited in Estonia.

The introduction of performance management tools possesses a strong symbolic value – it allows a government to pose as progressive to its constituents. However, the Estonian case study demonstrates that the implementation gap between declared aims of performance measurement and its actual practice may not only reduce the practical effects of such tools but also have a negative impact on the legitimacy of government institutions.

DISCUSSION

The findings of the exploratory study of Estonia suggest three principal conclusions. Firstly, they provide support to critical conclusions reached by previous studies of performance measurement in Western countries. Secondly, this study reported here allows to outline a set of *specific* factors that help to explain the problems of performance measurement in immature policy environments (instability, poor strategic planning and policy analysis, the implementation gap and foreign influence). Thirdly, it

will be demonstrated that such contextual factors are very influential in other CEE countries which, in turn, allows it to generalize some of the findings to other post-communist countries.

Instability

Very few measurement systems are free from design flaws and do not need any adjustment or fine-tuning to function properly. The policy evaluation process itself can lead to questioning the validity of policy goals and performance indicators. Yet, any performance measurement system presumes a certain stability of performance targets to make it possible for the system to perform its evaluative and steering functions. These, in turn, presuppose valid and fixed indicators if they are to result in a reliable and accurate assessment of performance. Any performance measurement system requires a stable baseline for comparison (Behn, 2003: 598–9) – otherwise, no improvement or deterioration over time can be registered.

In Estonia, fourteen different cabinets have held office during the period of 1991–2011. The average lifespan of a cabinet has been 1.38 years (Conrad and Golder, 2010). Turbulent environment represents a serious challenge for performance measurement systems in terms of discharging their accountability and control functions that are highly dependent on valid indicators. As performance indicators are dependent on the policymaking process in a democratic society, one might expect that with changes in the government, performance targets will change as well. In Estonia, however, a change of cabinet has not resulted in instant and radical overhauls of performance targets because detailed action plans are not perceived to have a high political significance. Public reporting by politicians is based more on party programmes than government plans.

However, instability is not limited to political changes. The staff leaving rate in the Estonian public service has been between 10 per cent and 15 per cent over the last 5 years (*Rahandusministeerium*). Although the Estonian civil service has been characterized as one of the least politicized in CEE (SIGMA, 2009), during the eighth cabinet (1999–2002) to hold office after Estonia regained its re-independence, only one out of twelve permanent secretaries was able to stay in office, and the tenth cabinet (2003–5) changed one half of the permanent secretaries. Although the reasons for such replacements have ranged from politically motivated to practically justified, such turmoil in high public offices makes trust-building and serious commitment very difficult, especially considering the fact that top officials usually play a key role in developing performance indicators and leading the implementation of measurement systems.

In addition to frequent turnover in political and administrative leadership, the legal environment in Estonia has also undergone rapid development. Therefore, definitions of indicators are subject to change as well. Consequently, the comparison of performance indicators between different measurement periods is very complicated and at times impossible (*Riigikontroll*, 2008b).

Previous studies on other CEE countries demonstrate that instability in the political and administrative arena as well as in the legislative context is a typical phenomenon in the entire region. The average lifespan of government cabinets in CEE countries is 1.4 years, compared to 1.9 in Western Europe (Blondel and Müller-Rommel, 2001; Conrad and Golder, 2010). Although there is fluctuation among the average figures for different CEE countries, it may still be concluded that cabinets in CEE are replaced more often than in the West. Several authors researching public management in CEE (e.g. Dunn *et al.*, 2006; Bouckaert *et al.*, 2008; Meyer-Sahling and Yesilkagit, 2011; Nakrošis, 2008; Randma-Liiv *et al.*, 2011) point in their analyses to instability as one of the major factors differentiating Eastern Europe from Western democracies.

Poor strategic planning and analytical skills: Unsustainability of initiatives and deficient performance indicators

The Estonian experience demonstrates how poor strategic planning and policy analysis may affect the outcome of performance measurement initiatives. Although necessary democratic institutions have been created in Estonia, politicians still have difficulties in setting strategic goals, which, in turn, engenders conflicting policies, symbolic objectives, frequently changing goals and unsustainable initiatives. As both politicians and civil servants only have two decades of experience in elaborating policy proposals, there is as yet no general culture which would require an *ex ante* analysis to precede the adoption of a new regulation or policy. In addition, since the entire society has grown accustomed to rapid and radical changes of the immediate past, it has been relatively easy for politicians, civil servants as well as citizens to accept new initiatives (including those related to performance measurement) without any major criticism. Moreover, when public organizations do not have a clear idea about the actual costs incurred in connection with performance measurement (such as the expenditure for designing the system; the productivity loss inherent in diverting the time and effort necessary to compile the assessment from other tasks; the expenditures required to update, develop, test and keep inventories of the instrument(s) used for performance measurement as well as the expenditure for keeping records of performance results and dealing with complaints – see Halachmi, 1993), performance-oriented reforms are easy to be reversed when these costs must be borne.

A key challenge faced by governments in new democracies is one of maintaining a strategic view in a rapidly changing political framework. However, the history of public administration in CEE tells us a mixed story of piecemeal attempts to reform existing systems (Verheijen, 2003). Inconsistency of ideas, institutions and practices, deriving not only from political instability but also from deficient strategic planning and poor analytical skills, have been highlighted by several studies (see, for example, Dunn *et al.*, 2006; Bouckaert *et al.*, 2008; Meyer-Sahling and Yesilkagit, 2011; Nakrošis, 2008;

Nemec *et al.*, 2008; Randma-Liiv *et al.*, 2011; SIGMA, 2009; Verheijen, 2003; World Bank, 2006).

Poor analytical capacity in CEE (see, for example, Verheijen, 2003; World Bank, 2006) is likely to lead to the development of inadequate performance indicators. On the one hand, limited analytical skills pose the risk of applying inappropriate measurement methods. On the other hand, performance measurement is not a valueless exercise. Several 'democratic' goals such as transparency, equal opportunities, access to public services, fair procedures and citizen participation in decision-making may conflict with more 'technocratic' goals such as efficiency, effectiveness, value-for-money or fast decision-making. Contradictions of this type are especially hard to solve in new democracies, where the principles informing the aforementioned democratic goals are not as deeply held and broadly accepted as in countries with long democratic traditions. As a rule, 'technocratic' goals are more easily quantifiable than 'democratic' goals. Decision-makers with relatively little (democratic) management experience and poor analytical capacities may easily gravitate down the slippery path of over-quantifying performance indicators because, as aptly summed up by Mintzberg (1994), these are easier to measure and look more objective, transparent and understandable than qualitative data. The situation will be even more complicated if limited resources bring pressure to prefer 'technocratic' goals. As a result, the concept of accountability to the general public with its multiple dimensions may easily change to single-track accountability for financial outcomes only.

Implementation gap

Political and administrative instability contributes to the fact that nationwide initiatives of performance measurement are often met with implementation difficulties. Several reforms (e.g. the introduction of pay-for-performance in Estonia) have been discontinued halfway through and new initiatives can be started before previous steps in the field could even be evaluated (see also Dunn *et al.*, 2006; Meyer-Sahling and Yesilkagit, 2011; Verheijen, 2003 for broader CEE experience). Initiators of important reform plans rarely remain in office long enough to oversee their implementation and resolve any glitches in a manner coherent with the purposes of the reforms.

Performance management is one of the most sophisticated areas of public management. Therefore, public sector managers in immature policy environments who often have insufficient managerial experience are likely to pay too little attention to its conceptual and implementation problems. In addition, decentralization and a high degree of discretion – which usually go hand in hand with the introduction of performance management tools – may create opportunities for abuse by individual public organizations and their leaders. The introduction of performance measurement is susceptible not only to gaming with performance indicators and information but further to increasing politicization and to facilitating corruption – problems that are already

pressing in new democracies (Nemec *et al.*, 2008; SIGMA, 2009). The implementation of performance measurement may also prove risky because of an insufficiently developed legislative framework and the general inadequacy of accountability and control mechanisms in CEE (World Bank, 2006).

The Estonian experience indicates that the ability of managers to make use of performance information is rather limited, largely because of insufficient management experience and weak links between strategic goals and evaluation. However, in terms of external use, performance information is even more neglected. In large part, the role and responsibility of the parliament have been limited to the preparation of large amounts of legislation required by post-communist transition and Europeanization due to the pressures of the immediate past. As a consequence, the parliament has proved unable to responsibly exercise its steering role of the executive, including setting performance targets and analysing performance information. The lack of experience and analytical skills of Members of the Parliament (MPs), coupled with the lack of resources and support staff, represent a serious handicap in terms of the scope and depth of parliamentary scrutiny. Additional problems specific to new democracies such as inadequate trust between politicians and civil servants (SIGMA, 2009) are likely to compound the problem by turning the performance measurement process into an exercise of control.

Foreign influence

Why do new democracies keep experimenting with performance measurement tools although their contextual factors are not supportive for such endeavours? Besides a general fashion of performance measurement in the framework of NPM reforms, international organizations (primarily the IMF, the World Bank and the OECD) have played a major role in the introduction of performance management tools in CEE. A number of such organizations have directly promoted and recommended performance management tools (see OECD, 2011; OECD/SIGMA, 2007; World Bank, 2006, for example). Similarly to the senior public officials who have made a career out of 'modernizing' and 'streamlining' activities (Pollitt and Bouckaert, 2004: 6), there are also many international consultants making a living out of spreading 'best practice'. The fact that the last few decades have resounded with 'performance talk' has paved the way for international consultants advocating performance management across the globe, including in CEE, despite the different context of post-communist countries. New democracies do not enjoy the history of democratic traditions and institutions, resulting, for instance, in a lack of administrative capacity and control mechanisms. Therefore, reform ideologies common to Western practices and advocated by Western public and private organizations may not be appropriate when applied to these countries (see also Condrey *et al.*, 2001; Randma-Liiv, 2007; Schick, 1998; Sobis and de Vries, 2009). 'One-size-fits-all' types of recommendations ignore the danger of policy failure caused by uninformed, incomplete or inappropriate policy transfer (Dolowitz and

Marsh, 2000). For new democracies, positive feedback from international evaluations is vital as it may determine the country's acceptance in international organizations and its credibility for foreign partners. Such a predisposition creates an environment in which even general guidelines, 'best practices' and soft policy recommendations on the introduction of performance measurement tools can be taken very seriously by newly democratic states.

CONCLUSIONS

This study confirms the results of previous research in that it highlights contradictory aspects of the concept of performance measurement and the challenges inherent in the corresponding implementation process. The analysis of the Estonian central government shows that many of the assumptions which have led to widespread public sector performance measurement as valid *a priori* turned out to stand on feet of clay. The Estonian governments of the last two decades have tried out most of the performance management tools known from Western experience. Many attempts have remained half-hearted or sporadic, yet performance measurement initiatives have become the subject of widespread and regular practice. This study at hand confirms Western experience; it shows that because of a variety of implementation difficulties and conceptual contradictions, management by targets and indicators does not *guarantee* better performance, or fulfilment of the managerial or democratic purposes of performance initiatives.

This study demonstrates that although problems common to Western countries are also present in the Estonian context, they are compounded in the latter by specific contextual factors characteristic of new democracies. These include instability, poor analytical skills, implementation gap and foreign influence in introducing performance measurement initiatives. These factors explain implementation difficulties of performance measurement and limited use of performance information in the context of post-communist countries.

This study has several implications for policy learning. Performance-orientation has been the centrepiece of the rhetoric of public management reforms in most new democracies in Europe due to the general appeal of NPM ideas and values during the last two decades. Although there are important differences among post-communist countries, their shared history of post-communist transition as well as Europeanization still appears to be taking them through a number of similar developments, offering them similar opportunities and exposing them to similar risks. The presence of 'transitional' factors analysed in this paper (e.g. political and administrative instability, unsustainability of reforms, poor analytical skills in parliaments as well as in civil services, implementation gap and foreign influence) is likely to make the development of performance measurement in new democracies even more complicated than in the West.

ACKNOWLEDGEMENTS

This study has received funding from the European Community's Seventh Framework Programme under Grant Agreement No. 266887 (Project COCOPS), Socio-economic Sciences & Humanities, and from the Estonian Science Foundation Grant No. 7441.

REFERENCES

- Andersen, S. C. (2008) The Impact of Public Management Reforms on Student Performance in Danish Schools. *Public Administration*, 86:2 pp541–58.
- Andrews, R., Boyne, G. and Enticott, G. (2006) Performance Failure in the Public Sector. *Public Management Review*, 8:2 pp273–96.
- Askim, J. (2007) How Do Politicians Use Performance Information? An Analysis of the Norwegian Local Government Experience. *International Review of Administrative Sciences*, 73:3 pp453–72.
- (2009) The Demand Side of Performance Measurement. Explaining Councillors' Utilization of Performance Information in Policymaking. *International Public Management Journal*, 12:1 pp24–47.
- Behn, R. D. (2003) Why Measure Performance? Different Purposes Require Different Measures. *Public Administration Review*, 63:5 pp586–606.
- Bevan, G. and Hood, C. (2006) What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System. *Public Administration*, 84:3 pp517–38.
- Blondel, J. and Müller-Rommel F. (2001) *Cabinets in Eastern Europe*, 193–201, Houndmills: Palgrave.
- Bouckaert, G., Nemeč, J., Nakrošis, V., Hajnal, G. and Tönnisson, K. (eds) (2008) *Public Management Reforms in Central and Eastern Europe*, Bratislava: NISPAcee Press.
- Condrey, S. E., Purvis, K. and Slava, S.S. (2001) Public Management Reform Under Stress – The Ukrainian Civil Service Experience. *Public Management Review*, 3:2 pp271–80.
- Conrad, C. R. and Golder, S. N. (2010) Measuring Government Duration and Stability in Central Eastern European Democracies. *European Journal of Political Research*, 49:1 pp119–50.
- Dolowitz, D. and Marsh, D. (2000) Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making. *Governance: An International Journal of Policy and Administration*, 13:1 pp5–24.
- Drechsler, W. (2004) Governance, Good Governance, and Government: The Case for Estonian Administrative Capacity. *Trames*, 8(50/53):4 pp388–96.
- Dunn, W., Staronova, K. and Pushkarev, S. (eds) (2006) *Implementation. The Missing Link in Public Administration Reform in Central and Eastern Europe*, Bratislava: NISPAcee Press.
- Goetz, K. H. (2001) Making Sense of Post-Communist Central Administration: Modernization, Europeanization or Latinization. *Journal of European Policy*, 8:6 pp1032–51.
- Halachmi, A. (1993) From Performance Appraisal to Performance Targeting. *Public Personnel Management*, 22:2 pp323–44.
- Holzer, M. and Yang, K. (2004) Performance Measurement and Improvement: An Assessment of the State of the Art. *International Review of Administrative Sciences*, 70:1 pp15–31.
- Lægreid, P., Roness, P. G. and Rubecksen, K. (2006) Performance Management in Practice: The Norwegian Way. *Financial Accountability and Management*, 22:3 pp251–70.
- Meyer-Sahling, J.-H. and Yesilkagit, K. (2011) Differential Legacy Effects: Three Propositions on the Impact of Administrative Traditions on Public Administration in Europe East and West. *Journal of European Public Policy*, 18:2 pp311–22.
- Mintzberg, H. (1994) *The Rise and Fall of Strategic Planning: Reconceiving Roles for Planning, Plans, Planners*, New York: The Free Press.

- Nakrošis, V. (2008) 'Reforming Performance Management in Lithuania: towards Results-based Government' in B. G. Peters (ed.) *Mixes, Matches, and Mistakes – New Public Management in Russia and the Former Soviet Republics*, 53–116. Budapest: OSI/LGI.
- Nemec, J., Merickova, B. and Ochrana, F. (2008) Introducing Benchmarking in the Czech Republic and Slovakia. *Public Management Review*, 10:5 pp673–84.
- (2011) Is the Estonian Municipal Benchmarking Really Better? *Public Management Review*, 13:4 pp539–49.
- OECD (2007) *Performance Budgeting in OECD Countries*, Paris: OECD.
- (2011) Estonia: Towards a Single Government Approach. *Public Governance Reviews*, Paris: OECD.
- OECD/SIGMA (2007) *Organising the Central State Administration: Policies and Instruments*, Sigma Paper No. 43, Paris: OECD.
- Pollitt, C. (2006) Performance Information for Democracy: The Missing Link? *Evaluation*, 12:1 pp38–55.
- Pollitt, C. and Bouckaert, G. (2004) *Public Management Reform: A Comparative Analysis* (2nd edn), New York: Oxford University Press.
- PRAXIS Centre for Policy Studies. (2007) 'Riikliku Arengukava 2004–2006 mõjuindikaatorite saavutusmääradev äljasegitamine ja analüüs' (Identification and analysis of the targets of the impact indicators). *PRAXISe Toimetised 31/2007*. Available in Estonian at <http://www.praxis.ee> (accessed 1 May 2011).
- (2010) *Riigi tulemusvaldkondade määratlemine (The Definition of Government Performance Areas)*. Final report of a project financed by the Ministry of Finance of Estonia. Available in Estonian at <http://www.praxis.ee> (accessed 1 May 2011).
- Radnor, Z. (2008) 'Hitting the Target and Missing the Point? Developing an Understanding of Organizational Gaming' in W. Van Dooren and S. Van de Walle (eds) *Performance Information in the Public Sector: How It Is Used*, 94–105. Basingstoke: Palgrave Macmillan.
- Rahandusministeerium (Ministry of Finance of Estonia). Avaliku teenistuse uuringud (Public Service in Figures). Available in Estonian at <http://www.avalikteenistus.ee> (accessed 1 May 2011).
- (2008) Riigi finantsjuhtimise arendamise kontseptsioon (The Conception of Public Financial Management). Available in Estonian at <http://www.avalikteenistus.ee>
- Randma-Liiv, T. (2005) Performance Management in Transitional Administration: Introduction of Pay-for-Performance in the Estonian Civil Service. *Journal of Comparative Policy Analysis*, 7:1 pp95–115.
- (2007) 'From Policy Transfer to Policy Learning in Central and Eastern Europe' in D. Coombes and L. Vass (eds) *Post-Communist Public Administration: Restoring Professionalism and Accountability*, 27–36. Bratislava: NISPAcee.
- Randma-Liiv, T., Nakrošis, V. and Hajnal, G. (2011) Public Sector Organization in Central and Eastern Europe: From Agencification to De-Agencification. *Transylvanian Review of Administrative Sciences*, November 160–75.
- Raudla, R. (2011) 'The Use of Performance Information in Budgetary Decision-Making by Legislators: Is Estonia Any Different?' Paper prepared for the EGPA Annual Conference in Bucharest, Romania, 7–10 September 2011.
- Riigikogu (The Parliament of Estonia) (2005) Transcript of the 21.03.2005 session. Available in Estonian at <http://www.riigikogu.ee/?op=steno> (accessed 1 May 2011).
- Riigikontroll (National Audit Office of Estonia) (2002) *Tulemusjuhtimiseks eraldatud töötasu vahendite kasutamine (The Use of Pay-for-PERFORMANCE Funds)*. Audit report. Summary in English available at <http://www.riigikontroll.ee> (accessed 1 May 2011).
- (2008a) *Kõrghariduse riiklik koolitustellimus (Government-commissioned Higher Education)*. Audit report. Summary in English available at <http://www.riigikontroll.ee> (accessed 1 May 2011).
- (2008b) *Rahandusministeeriumi tegevus eelarvestamise põhimõtete kujundamisel (Activities of the Ministry of Finance in Developing Budgeting Principles)*. Audit report. Summary in English available at <http://www.riigikontroll.ee> (accessed 1 May 2011).

- (2009) *Riigikutsõppeasutuste ümberkorraldamise tulemuslikkus (Effectiveness of Reorganization of Public Vocational Educational Institutions)*. Audit report. Summary in English available at <http://www.riigikontroll.ee> (accessed 1 May 2011).
- (2010) *Täiskasvanute täiendus- ja ümberõpe (In-service and Re-training of Adults)*. Audit report. Summary in English available at <http://www.riigikontroll.ee> (accessed 1 May 2011).
- (2011) *Overview of Use and Preservation of State Assets in 2010*. Annual report by the National Audit Office to the Parliament.
- Sarapuu, K. (2011) Post-Communist Development of Administrative Structure in Estonia: From Fragmentation to Segmentation. *Transylvanian Review of Administrative Sciences*, November pp54–73.
- Schick, A. (1998) Why Most Developing Countries Should Not Try New Zealand's Reforms. *World Bank Research Observer*, 13:1 pp123–31.
- SIGMA (2009) *Sustainability of Civil Service Reforms in Central and Eastern Europe Five Years after EU Accession*. Paper No. 44, Paris: SIGMA.
- Sobis, I. and de Vries, M. S. (2009) *The Story behind Western Advice to Central Europe during Its Transition Period*, Bratislava: NISPAcee Press.
- Talbot, C. (2007) 'Performance Management' in E. Ferlie, L. Lynn and C. Pollitt (eds) *The Oxford Handbook of Public Management*, 491–517. Oxford/New York: Oxford University Press.
- Tõnnisson, K. and Randma-Liiv, T. (2008) 'Public Management Reforms: Estonia' in G. Bouckaert, J. Nemeč, V. Nakrosis, G. Hajnal and K. Tõnnisson (eds) *Public Management Reforms in Central and Eastern Europe*, 93–120. Bratislava: NISPAcee Press.
- Tõnnisson, K. and Wilson, J. (2007) 'Best Value in Transitional Countries?' *Public Management Review*, 9:1 pp87–106.
- Van de Walle, S. (2009) International Comparisons of Public Sector Performance. How to Move Ahead? *Public Management Review*, 11:1 pp39–56.
- Van Dooren, W. (2008). 'Nothing New Under the Sun? Change and Continuity in the Twentieth-Century Performance Movements' in W. Van Dooren and S. Van de Walle (eds) *Performance Information in the Public Sector: How It Is Used*, 11–23. Hampshire: Palgrave Macmillan.
- Van Dooren, W., Bouckaert, G. and Halligan, J. (2010) *Performance Management in the Public Sector*, London: Routledge.
- Van Stolk, C. and Wegrich, K. (2008) Convergence without Diffusion? A Comparative Analysis of the Choice of Performance Indicators in Tax Administration and Social Security *International Review of Administrative Sciences*, 74:4 pp589–614.
- Verheijen, T. (2003). 'Public Administration in Post-communist States' in G.B. Peters and J. Pierre (eds) *Handbook of Public Administration*, 489–99. London: Sage.
- Verheijen, T. and Dobrolyubova, Y. (2007) Performance Management in the Baltic States and Russia: Success against the Odds? *International Review of Administrative Sciences*, 73:2 pp205–15.
- World Bank. (2005) *Increasing Government Effectiveness: Approaches to Administrative Reform in the Russian Federation*. Report Number: 36142.
- (2006) *EU-8 Administrative Capacity in the New Member States: The Limits of Innovation?* Report Number: 36930-GLB.
- Zubek, R. and Goetz, K. H. (2010) Performing to Type? How State Institutions Matter in East Central Europe. *Journal of Public Policy*, 30:1 pp1–22.

II Raudla, Ringa, **Külli Taro**, Cherlin Agu and James W. Douglas. 2015. “Impact of Performance Audit on Public Sector Organizations: The Case of Estonia” *Public Organization Review*. Published online. DOI <http://dx.doi.org/10.1007/s11115-015-0308-0>. (1.1)

The Impact of Performance Audit on Public Sector Organizations: The Case of Estonia

Ringa Raudla · Külli Taro · Cherlin Agu ·
James W. Douglas

© Springer Science+Business Media New York 2015

Abstract The results of a survey of 118 Estonian public officials show that auditees can perceive performance audit to be useful even if it does not lead to specific changes in policies or organizational practices. No trade-off between the accountability and improvement functions of performance audit could be observed, though a surprisingly low percentage of the respondents considered performance audit to have been used to hold the audited organization accountable for their actions. While the factors internal to the audit process – including the perceived expertise of the auditors, their openness to dialogue with the auditees and the quality of audit report – influenced the perceived usefulness of the audit, they had less bearing on the adoption of changes by the audited organizations. The study indicates that when parliamentarians pay attention to performance audit and when media attention leads to political debate, the adoption of changes recommended by the performance audit report is more likely.

Keywords Performance audit · Performance management · Accountability · State audit institution

R. Raudla (✉) · K. Taro · C. Agu

Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Akadeemia 3,
Tallinn 12618, Estonia
e-mail: ringa.raudla@ttu.ee

K. Taro
e-mail: kylli@kogu.ee

C. Agu
e-mail: cherlin.agu@ttu.ee

J. W. Douglas
Department of Political Science and Public Administration, University of North Carolina at Charlotte,
9201 University Blvd, Charlotte, NC 28223, USA
e-mail: jwdougla@uncc.edu

Introduction

Public sector performance has been at the very center of the public management debate over the past 15–20 years (Pollitt and Bouckaert 2004; Talbot 2010). Alongside the rise of New Public Management reforms and more recent economic and fiscal crises, the public sector worldwide has been under constant pressure to improve its performance in pursuit of more efficiency and effectiveness and in order to revive citizens' trust in public institutions. Following the rise of performance movements, Supreme Audit Institutions (SAIs) in many countries have introduced the practice of performance auditing. There are many definitions of performance audit (see INTOSAI; Pollitt et al. 1999; Furubo 2011, for example). Probably the most well-known definition by the International Organization of Supreme Audit Institutions (INTOSAI) is based on the concept of the three Es (economy, efficiency and effectiveness): “Performance auditing is an independent and objective examination of government undertakings, systems, programs or organizations, with regard to one or more of the three aspects of economy, efficiency and effectiveness, aiming to lead to improvements.” By now, performance audit (PA) is considered to be an essential element of “an audit portfolio” for a SAI (Lonsdale 2011; Pollitt et al. 1999). By holding substantial, if not predominant position, and by gaining much public attention, it has even shifted the overall balance of the audit work in SAIs (Lonsdale 2011; Pollitt et al. 1999).

Increasingly, the academic discussions on performance audit have started to focus on the role and tasks of the audits. It can be observed that in addition to being viewed as an instrument for enhancing accountability, performance audit is increasingly expected to contribute to the wider policy debates, provide policy advice to governments and give guidance about how public administration can improve its work (Funkhouser 2011; Lonsdale and Bechberger 2011; Reichborn-Kjennerud 2013b; Wilkins and Lonsdale 2007). Despite the growing use of PA by SAIs, the volume of academic, empirical studies on performance auditing is surprisingly small – especially if compared with studies on other topics of performance management and evaluation (Funkhouser 2011; Lonsdale 2011; Pollitt et al. 1999). There are even fewer studies that have looked at the *impacts* of performance audits on public sector organizations (Kells 2011; Leeuw 2011; Morin 2014; Reichborn-Kjennerud 2013b; Van Loocke and Put 2011). Given its rising importance and growing expectations about what performance audit is supposed to achieve, it is justified to ask how well performance audit has actually “performed” and what have been its impacts.

With our paper we seek to address this gap in the existing literature. In particular, we address the following research questions: 1) Are performance audits perceived to be about holding ministries and agencies accountable or are they viewed to be primarily about improving the activities by the officials in the audited entities? Can tradeoffs between these two functions be observed or are they mutually compatible? 2) Do the auditees perceive the performance audits to be useful? 3) Does the performance audit lead to changes in the audited organizations? If yes, what kind of changes? 4) Which factors increase the perceived usefulness of performance audit and which factors facilitate the undertaking of changes in the audited organizations? In order to answer these research questions, we use the results from a survey (conducted in 2013–2014) of 118 civil servants in Estonia who have experienced performance audits.

The contribution of our study is to enhance the understanding of the impacts of performance audit on public sector organizations. Some of the existing empirical studies have already explored the perceived usefulness of PA and the changes it has brought about (see Lonsdale 1999; Morin 2004, 2008), investigated the factors that influence the perceived usefulness of PA (Reichborn-Kjennerud 2013a), and identified variables conducive to the implementation of changes after PA (Morin 2004, 2008, 2014). The new angle provided by our study is to look at *both* types of impacts of PA – its perceived *usefulness* and induced *changes* – and to analyze whether the same factors that increase the perceived usefulness of PA also help to facilitate organizational changes. Further, unlike existing studies we explicitly distinguish between *internal* and *external* factors that can influence whether PA leads to the implementation of changes recommended in the audit report and examine which type of factors – external or internal – play a more significant role. Finally, while none of the existing empirical studies have explicitly examined whether there is a trade-off between the accountability and improvement functions of performance audit, we seek to test whether such a trade-off exists. The paper is structured as follows. [Section 2](#) outlines the analytical framework. [Section 3](#) presents the results of data analysis and [section 4](#) concludes with a discussion.

Analytical Framework

The Role of Performance Audit: Accountability Versus Learning?

In the discussions over the rationale, functions and role of performance audit, two main “functions” or “goals” have emerged. First, performance audit is viewed as a mechanism to enhance accountability in the public sector. Second, performance audit is expected to contribute to learning, improvement, and change. Regarding the compatibility of these two functions, theoretical discussions on the functions and impacts of performance audit again point to two diverging lines of argumentation. On the one hand, from the perspective of principal-agent theory (see, e.g. Moe 1984; Bendor et al. 1987), it can be maintained that the two main functions of performance audit – ensuring accountability and improving organizational performance – are congruous (or at least not entirely incompatible) in that the auditors’ focus on accountability can facilitate learning by the audited actors and hence promote organizational improvement and change (Funkhouser 2011; Weiss 1998; Wilkins and Lonsdale 2007). New information and insights revealed by the audit can point to areas where improvements are needed and, given the accountability context in which the audit takes place, the actors would have incentives to undertake these changes (Justesen and Skaerbek 2010; Funkhouser 2011; Lonsdale and Bechberger 2011).

On the other hand, based on the cultural approaches to organizational learning (e.g. Popper and Lipshitz 1998), it can be conjectured that the two main functions of performance audit – ensuring accountability and improving (actual) performance (via learning and change) – can be difficult to reconcile or are even incompatible (e.g. Behn 2001; Dubnick 2005; Lonsdale and Bechberger 2011; van der Meer and Edelenbos 2006). In this line of argumentation it is often noted that the auditors’ focus on the organizations’ compliance with rules and set targets in order to increase accountability

can make it difficult to use the audit process for facilitating learning (Lonsdale and Bechberger 2011). A focus on accountability entails emphasis on rigour, independence, even some social distance, whereas a focus on learning would imply stake-holder “buy-in”, trust, social proximity, and space for discussion (Lonsdale and Bechberger 2011; Morin 2003; Van Looke and Put 2011). The public nature of performance audits, especially if they emphasize control and attribute blame, can generate defensiveness (or even gaming) by the auditees and limit the opportunities to learn (e.g. Lonsdale and Bechberger 2011; van der Meer and Edelenbos 2006). Furthermore, the auditors’ focus on compliance, control and formal procedures may impede innovation, risk-taking and experimentation in the audited organizations (Bemelmans-Videc et al. 2007; Kells 2011; Leeuw 2009).

Whether these two functions – ensuring accountability and improving performance – are compatible or incongruent still needs empirical testing: none of the existing empirical studies has explicitly done it. Thus, in the empirical part of the paper we will test the following hypothesis:

H 1: There is a trade-off between the two functions of performance audit: if the performance audit is perceived to be about accountability, it is less likely to be perceived as being used for improving operations in the audited organizations.

What Factors Increase the Perceived Usefulness of the Performance Audit and Facilitate the Undertaking of Changes?

In conceptualizing the possible *impacts* of performance audit (or the “performance of the performance audit”) a number of taxonomies have been put forth (e.g. Van Looke and Put 2011; Morin 2001, 2004, 2008). The most prevalently discussed impacts of PA have been the *changes* implemented as a result of PA and the perceived *usefulness* of PA by the auditees (e.g. Alwardat 2010; Brown and Craft 1980; Morin 2001, 2004, 2008, 2014; Reichborn-Kjennerud 2013b; Van Looke and Put 2011).¹ More specifically, the following changes have been mentioned as potentially resulting from PA: modifications of laws and regulations, changes in human resource management, shifts in strategic planning and performance and risk management, the adoption of best practices, and changes in personnel (e.g., Lapsley and Pong 2000; Morin 2004; 2008). In addition to the *changes* resulting from PA, the *usefulness* of performance audit (as perceived by the auditees) has been regarded as a possible impact (Morin 2001, 2004, 2008, 2014; Reichborn-Kjennerud 2013a; Van Looke and Put 2011). As noted by Reichborn-Kjennerud (2013a), the perceived usefulness may in fact be a more valid indicator of the general impact of the performance audit than the induced changes (especially if a survey method is used), given that the perception of improvements and actual change would have to be verified with other sources of data.

In addition to looking at the impacts of PA as such, several existing studies have attempted to uncover different factors that facilitate or aggravate these impacts. A number of variables have been discussed in the existing studies (e.g. Alwardat 2010; Reichborn-Kjennerud 2013a; Morin 2001, 2004, 2008; 2014). For the purposes of

¹ For an overview of the existing studies on the impacts of PA, see Van Looke and Put (2011).

analytical clarity – and in order to uncover the different impacts such factors can have – we propose to divide them into “internal” and “external” factors.

Internal factors (i.e. factors “internal” to the audit) refer to the auditees’ perception of the *auditors* (including their perceived competence and expertise), the characteristics of the *audit process* (e.g. openness to dialogue by the auditors) and the (perceived) quality of the *audit report* (Alwardat 2010; Van Loocke and Put 2011; Morin 2001; Reichborn-Kjennerud 2013a, b). First, it can be expected – drawing on theories of social influence (e.g. Tedeschi 1972) and expert power (e.g. Raven 1992) – that if the auditees consider the auditors to have expertise and competence, they are more likely to consider the audit as useful and also more likely to implement changes proposed (Alwardat 2010; Morin 2004, 2008). Second, it has been argued that if the *audit report* is viewed as being of *high quality*, it would be more likely that the auditees use it for undertaking recommended changes and view the audit itself as more useful (Hatherly and Parker 1988; Lonsdale and Bechberger 2011; Reichborn-Kjennerud 2013a, b). Third, extending the insights of interactive-dialogue theory of performance information use (see Moynihan 2006, 2008a,b) to the context of performance audit, it can be hypothesized that the perceived usefulness of the audit and the implementation of changes can be facilitated by dialogic and reflective approach in the audit process, in which the auditors involve the stakeholders in the evaluation process, are open to dialogue and take into account the observations of the auditees (Lonsdale and Bechberger 2011; Morin 2001, 2004; Preskill and Torres 2000; Reichborn-Kjennerud and Johnsen 2011; Weets 2011). Communication between the auditor and the auditee can contribute to clearer understanding of the issues involved and hence increase the perceived usefulness of the audit (and also increase the likelihood of the implementation of recommended changes) (Van Loocke and Put 2011; Van Der Meer 1999). Finally, based on sociological institutionalism or organizational institutionalism (e.g. March and Olsen 1989; Meyer and Rowan 1977), we would expect the auditees to filter the proposed recommendations through the “logic of appropriateness” of their own organizations (Reichborn-Kjennerud 2013b). Thus, the more the auditees agree with the audit criteria and audit conclusions, the more likely they would view the audit as useful (Reichborn-Kjennerud 2013a) and, therefore, the more likely the PA would lead to changes in the audited organizations. The existing empirical studies show that the auditees in Norway considered the performance audit to be more useful when they perceived the audit report to be of high quality, agreed with the audit criteria, and were allowed to influence the audit process (Reichborn-Kjennerud 2013a), that the auditees in the UK were more willing to implement changes when they perceived the auditors to have high competence (Alwardat 2010) and that PA in Canada had stronger impacts when the auditees perceived the auditors to be open to communication (Morin 2004).

The main *external* factors that have been discussed in the existing literature are the involvement of members of parliament and the media in the discussion of the audit results. In particular, it can be argued, in correspondence with the principal-agent theory, that if the audit report receives wider attention – especially by the media and the parliament – the publicness of the report (and also formal accountability relationships between the parliament and the executive in parliamentary systems) can facilitate adaptive action (e.g. Lonsdale 2008; Morin 2001, 2004, 2008; Pollitt and Summa 1996; Weiss 1998). The existing empirical studies undertaken in Canada show that the intervention of the parliamentarians has been unambiguously important factor in

facilitating the changes undertaken in response to the PA (Morin 2004, 2008, 2014), whereas the effects media coverage of the audit have been found to be weak in one study (Morin 2004) but more significant in another (Morin 2008). In a study of the US states Brown and Craft (1980) found that both the actions of the legislators and media attention increased the impacts of PA.

Based on these theoretical discussions, we can formulate the following hypotheses concerning the *perceived usefulness* of the performance audit:

The perceived *usefulness* of the performance audit by auditees is likely to be higher,

- H 2.1.: the higher their perceived expertise of the members of the audit team
- H 2.2: the higher their perceived quality of the audit report
- H 2.3: the higher their perceived willingness of the auditors to engage in a dialogue during the audit process
- H 2.4: the more extensive the explanation of the shortcomings in the audit report
- H 2.5: the higher their agreement with audit criteria and the audit conclusions by the auditees
- H 2.6: the more concrete the recommendations made in the audit report
- H 2.7: the more changes made as a result of the performance audit

With regard to the changes in the audited organizations adopted in response to performance audit, we can put forth the following set of hypotheses:

The changes in the audited organizations adopted in response to performance audit, as perceived by the auditees, are likely to be more extensive,

- H 3.1.: the higher their perceived expertise of the members of the audit team
- H 3.2: the higher their perceived quality of the audit report
- H 3.3: the higher their perceived willingness of the auditors to engage in a dialogue during the audit process
- H 3.4: the more extensive the explanation of the shortcomings in the audit report
- H 3.5: the higher their agreement with audit criteria and conclusions by the auditees
- H 3.6: the more concrete the explanation of the shortcomings in the audit report
- H 3.7: the more attention performance audit received from the media.
- H 3.8: the more attention the parliamentarians paid to the performance audit

The Empirical Study

In answering the research questions, we use a survey of Estonian public officials who have experienced one or more performance audits conducted by the National Audit Office (NAO). The NAO in Estonia is modeled according to modern European practices. It is headed by the Auditor General who is appointed for a renewable 5-year term by the Parliament (named *Riigikogu*) on the proposal of the President. The independence of the NAO is secured by the Constitution and the National Audit Office Act. The role of the Audit Office is to assure the public and the Parliament that public sector assets are used legally and effectively. It decides

independently what, how, and when to audit. The NAO has the right to make (non-binding) proposals to the Government, ministers and local authorities to draft or amend legislation. The NAO has conducted performance audits since 2002 (usually there are 10–20 performance audits per year). Audit reports are published and sent to the Parliament’s Select Committee on the Control of State Budget and respective Standing Committees for review.

The survey was conducted in 2013–2014 and covered all performance audits that have taken place between 2005 and 2012. The survey was based on a research instrument developed by Reichborn-Kjennerud (2013a,b). The Estonian research team translated the survey questions into Estonian and administered the survey. The survey consisted of 35 questions about the performance audit(s) the auditees had experienced; most of the questions required answers on a 5-point Likert scale. The survey was sent out to the whole census of civil servants who had been involved in the performance audits since between 2005 and 2012. The survey was sent to 398 officials and the number of respondents was 118 (resulting in a response rate of 30 %).² The relatively low response rate can be explained by high turnover in the Estonian public sector (i.e. many of the potential respondents were not working in the audited organizations anymore), high workload of the officials, and a large number of questions in the survey.

General Impacts of Performance Audit in Estonia: Perceived Usefulness and Resulting Changes

In the survey, the auditees were asked to what extent the audited organizations had implemented changes recommended in the PA and whether they had found the performance audit to be useful.

As can be seen from Table 1, 40 % of the auditees found PA to have been useful, but there was significantly less agreement with the statement that PA led to the adoption of changes (21 %).

As can be seen from Table 2, among the types of change that were adopted in response to the performance audit, the three most prevalent ones were: changes in strategies, planning and performance management; changes in coordination process; and changes in relevant procedures.

Is There a Trade-Off Between the Functions of the Performance Audit?

As discussed in *Analytical Framework* section, one of the central questions of performance audit is whether its two main functions – ensuring accountability and contribution to change – are incompatible or compatible. According to the performance paradox, we expected in H1 that there is a trade-off: the more PA is used for accountability purposes, the less likely it is perceived to lead to changes and improvements (e.g., Behn 2001; Dubnick 2005; Reichborn-Kjennerud 2013a). One of the survey questions asked the respondents to what extent performance audit was perceived to be about holding ministries and agencies *accountable*, and to what extent it was viewed to be about *improving* operations in the audited entities?

² Among the 118 respondents, 75 completed the full survey and 43 completed it partially.

Table 1 The perception of the extent to which the performance audit had led to changes in the audited organization and the auditees' perception of the usefulness of the audit

Question:	1 To a very small extent	2	3	4	5 To a very large extent	Number	Mean
To what extent did the audited agency make changes as a consequence of the assessments in the performance audit report?	10 14.1 %	19 26.8 %	27 38.0 %	13 18.3 %	2 2.8 %	71	2.69
To what extent did you find the performance audit useful?	3 4.3 %	15 21.4 %	24 34.3 %	20 28.6 %	8 11.4 %	70	3.21

First, answers to the two questions in Table 3 allow us to check whether the respondents perceived there to be a tradeoff between the functions of performance audit. A simple correlation analysis showed that there is *no* statistically significant negative correlation between the answers reported in Table 3 (about whether the PA was used to hold ministers accountable and whether PA was used to improve processes within the audited organizations) [-0.054 ; $p=0.701$). Thus, Hypothesis 1 could not be confirmed. Second, in light of the theoretical discussions of performance audit in the existing literature (covered in [Analytical Framework](#) section), it is certainly surprising that a large majority of the respondents did *not* think that the PA was used to hold ministers accountable. It is especially puzzling given that in their audit reports, the

Table 2 Types of change adopted by audited organizations in response to the performance audit

What changes have taken place at your place of work after the performance audit? (you can choose several options)		
Answer options	Response percent	Response count
Changes in laws and regulations	27.9 %	17
Changes in significant procedures pinpointed by the NAO	32.8 %	20
Changes in the strategies, planning or performance management	39.3 %	24
Changes in internal control and risk management	14.8 %	9
Increased budgetary allocation to the audited policy area	9.8 %	6
Changes in procedures of coordination	32.8 %	20
Changes in management or organization	13.1 %	8
Changes in the training of the staff	6.6 %	4
Increased documentation and reporting	23.0 %	14
Changes in personnel	8.2 %	5
Layoffs and changes of workplace	8.2 %	5
Other		10
Answered question		61
Skipped question		57

Table 3 Accountability versus improving operations

To what extent did you agree with the statements about the performance audit?	1 To a very small extent	2	3	4	5 To a very large extent	Number	Mean
Performance audit was used to hold the ministry and agencies accountable for their actions	23 42.6 %	19 35.2 %	7 13.0 %	5 9.3 %	0 0.0 %	54	1.89
Performance audit was used in improving operations in the audited entities	6 8.5 %	6 8.5 %	18 25.4 %	26 36.6 %	15 21.5 %	71	3.54

NAO auditors in Estonia specifically address the responsible ministers (and also expect the ministers to answer to the recommendations outlined in the audit reports).

One of the possible reasons behind the perception that performance audits are not used for holding ministers accountable is that the Estonian parliament in general and the parliamentary committee dealing with performance audits, in particular, has been relatively passive in utilizing the PA reports (Nõmm and Randma-Liiv 2012). In the survey, the respondents were also asked about the extent to which the parliamentary committee required the audited organizations to undertake changes pointed out in the PA. Only 11 % of the respondents thought that the committee did so to some significant degree (i.e. only 11 % chose 4 and 5 on the 5-point scale). Correlation analysis also indicated that there is a positive correlation between the answers to the two questions (PA used for holding the minister accountable and the parliamentary committee required action) [0.299 ($p=.068$), $N=38$].³

What Kind of Factors Contribute to the Perceived Usefulness of Performance Audit

In order to test hypotheses 2.1–2.7 outlined in [Analytical Framework](#) section, we conducted a correlation analysis – in order to uncover whether there were correlations between the respondents' evaluations of the various aspects of the performance audit and its perceived usefulness. The results of the analysis are reported in [Table 4](#).

The results reported in [Table 4](#) indicate that the perceived usefulness of the audit is indeed correlated with the perceived expertise of the auditors (as hypothesized in H 2.1), high quality of the audit report (as hypothesized in H 2.2) and the perceived openness of the audit team to dialogue with the auditees (as hypothesized in H 2.3). The correlation analysis also shows that the higher the agreement with the audit criteria and the conclusions in the audit report by the auditees, the higher the perceived usefulness of the audit (as hypothesized in H 2.5). It can also be observed that the more concrete the recommendations made in the audit report, the higher the perceived usefulness of the audit by the auditees (as hypothesized in H 2.6). There was, however, no correlation between the explanation of the shortcomings in the audit report and the perceived usefulness of the performance audit (i.e., H 2.4 was not corroborated). With regard to the association between the

³ The same result was found in Reichborn-Kjennerud (2013a) .

Table 4 Bivariate correlations between various factors and *perceived usefulness of the audit*

	Pearson correlation	<i>p</i> -value	Number
The performance audit report proved that the auditors had good sector expertise	0.426***	0.000	64
The performance audit report held good quality	0.660***	0.000	66
The audit team open to dialogue	0.416***	0.000	68
Agreement with the audit criteria and the way they were used	0.418**	0.002	53
Agreement with the NAO's conclusions in the performance audit report	0.453***	0.000	69
Recommendations in the performance audit report were sufficiently concrete	0.353**	0.004	65
The performance audit report provided explanations about the shortcomings found	0.124	0.314	68
The audited entity make changes as a consequence of the assessments in the performance audit report	0.408***	0.001	67

*0.05 level; **0.01 level; ***0.001 level

Answers to all questions were given on a 5-point Likert scale

PA-induced changes, the analysis shows that the more the PA was seen as having led to changes, the more useful it was perceived to be (corroborating H 2.7).

Which Factors Contribute to PA-Induced Change?

In order to test hypotheses 3.1–3.8, we examined whether there were correlations between the extent to which PA was seen as having led to change in the audited organizations and the various factors outlined in *Analytical Framework* section. The results of the analysis are reported in Table 5.

Of all the variables that were expected to be conducive to PA-induced change, the only “internal” factors with a significant correlation were the perceived competence of the audit team and the quality of the audit report. However, both of these were significant only at the 0.1 level. Thus, we can argue that while hypotheses H 3.1. and H 3.2 found some corroboration, hypotheses H 3.3–3.6 were not confirmed. With regard to the “external” variables, the degree of media attention *alone* was not correlated with the adopted changes but when media attention had led to political debate and increased pressure on the minister and administration from the political opponents, the media’s role appears significant (hence offering at least partial confirmation of hypothesis 3.7). Also, it can be observed that there is a correlation between the extent of the demands of the parliamentary committee and the extent of reported changes (hence corroborating H 3.8).

In order to explore what may have prevented audited entities from implementing the changes recommended by the performance audit, respondents whose organizations did not make changes were asked to list the reasons for their inaction (reported in Table 6).

As can be seen from Table 6, the predominant reason why the respondents did *not* think that the performance audit had *led* to changes in the audited organization was that

Table 5 Bivariate correlations between various factors and *the extent to which the audited agency made changes as a consequence of the assessments in the performance audit report*

	Pearson correlation	<i>p</i> -value	Number
The performance audit report proved that the auditors had good sector expertise	0.251 ⁺	0.070	53
The performance audit report held good quality	0.212 ⁺	0.082	68
The audit team was open to dialogue	0.041	0.734	70
Agreement with the audit criteria and the way they were used	0.177	0.201	54
Agreement with the NAO's conclusions in the performance audit report	0.071	0.554	71
Recommendations in the performance audit report were sufficiently concrete	0.063	0.613	66
The performance audit report provided explanations about the shortcomings found	-0.082	0.501	69
The committee of the parliament demanded that the government improve the situation as a response to the performance audit	0.307*	0.038	46
The media showed an interest in the performance audit	0.009	0.945	57
As a result of the media attention, a political debate was launched	0.307*	0.028	51
As a result of the media attention, the minister and/or the administration were put under pressure from political opponents	0.284*	0.050	48

⁺ 0.1 level; *0.05 level; **0.01 level; ***0.001 level

Answers to all questions were given on a 5-point Likert scale

the organization was *already aware* of the challenges and had put measures in place to correct them. Given that, on average, an audit lasts about 13 months it is not surprising

Table 6 Reasons for not implementing changes

If you didn't make changes, what was the reason that changes weren't made? (you can choose more than one option)

Answer options	Response percent	Response count
We didn't have the possibility to make all these changes as they also involved other levels of public administration	39.0 %	23
We didn't think that the facts in the report were presented accurately	23.7 %	14
We didn't agree with the assessments of the NAO	32.2 %	19
We weren't responsible for making the changes	32.2 %	19
It is too early make changes	8.5 %	5
We were already aware of the challenges and had put measures in place to correct them	50.8 %	30
I don't know.	8.5 %	5
Answered question		59
Skipped question		59

that the audited organizations started undertaking changes already during the audit itself (and hence did not view the changes as something that had been undertaken strictly *in response* to the audit report itself). In addition, frequently mentioned reasons were also: “we didn’t agree with the assessments of the NAO”, “we didn’t have the possibility to make all these changes as they also involved other levels of public administration” and “we weren’t responsible for making the changes”. As pointed out by previous studies about public administration in Estonia, the coordination of activities between the various public sector organizations in Estonia is one of the most pressing challenges in the Estonian public sector (OECD 2011). These challenges are clearly reflected in the results of our study as well.

Conclusions

Given the spreading use of performance audit and increasing expectations about what it is supposed to achieve, it is only fair to ask what the impacts of PA have been, whether it been “delivering” its promises, and how it has been performing. While we agree with Pollitt and Summa (1997) and Lonsdale (2000) that measuring the impact of performance audit on a precise scale is a very difficult (if not impossible) exercise, we believe that assessing the impacts is still an endeavor worth undertaking. This article explores the impact of PA in Estonia by looking at its perceived usefulness of PA in the eyes of the auditees and by examining to what extent PA (as perceived by the auditees) led to changes in the audited organizations. In addition, we attempted to uncover which factors increased the perceived usefulness and the extent of changes undertaken in response to PA. Table 7 summarizes the hypotheses we tested and whether they received support or not.

This article contributes to existing knowledge about the impacts of performance audit in the following ways:

First, the existing empirical studies have not explicitly examined the question of whether there is a trade-off between the accountability and improvement functions of performance audit. Our study sought to test that and the results suggest that no such trade-off could be observed in the Estonian context. A surprising finding of our survey was that less than 10 % of the respondents considered performance audit to have been used for holding the audited organization *accountable* for their actions. Given that the literature on performance audit considers accountability to be one of the main functions of performance audit, it would be highly interesting to explore in future studies whether similar patterns could be found in other countries as well.

Second, our study shows that when investigating the impacts of performance audit on the audited organizations, it is fruitful to look at *both* types of impacts – perceived usefulness and resulting changes – because their scope may vary. In the Estonian case, while 40 % of the auditees found PA to have been useful, only 21 % agreed that it led to the adoption of changes in the audited organizations. While it could be argued that scrutinizing the *changes* PA has brought about captures the more direct impacts of PA, our research indicates that examining the *perceived usefulness* of PA helps to gauge the impacts of the audit more broadly and it could be used to provide a more holistic picture of the role of PA in the public sector. Indeed, our analysis suggests that although there is a correlation between the perceived usefulness of the audit and the extent of

Table 7 The hypotheses and the empirical support

Hypotheses	Support
H 1: There is a trade-off between the two functions of performance audit (accountability and improvement)	No
H 2.1: The higher the perceived expertise of the members of the audit team the higher the perceived usefulness of PA	Yes
H 2.2: The higher the perceived quality of the audit report, the higher the perceived usefulness of PA	Yes
H 2.3: The higher the perceived willingness of the auditors to engage in a dialogue during the audit process, the higher the perceived usefulness of PA	Yes
H 2.4: The more extensive the explanation of the shortcomings in the audit report, the higher the perceived usefulness of PA	No
H 2.5: The higher their agreement with audit criteria and the audit conclusions by the auditees, the higher the perceived usefulness of PA	Yes
H 2.6: The more concrete the recommendations made in the audit report, the higher the perceived usefulness of PA	Yes
H 2.7: The more changes made as a result of the performance audit, the higher the perceived usefulness of PA	Yes
H 3.1: The higher the perceived expertise of the members of the audit team, the more extensive the changes adopted in response to PA	Yes
H 3.2: The higher the perceived quality of the audit report, the more extensive the changes adopted in response to PA	Yes
H 3.3: The higher their perceived willingness of the auditors to engage in a dialogue during the audit process, the more extensive the changes adopted in response to PA	No
H 3.4: The more extensive the explanation of the shortcomings in the audit report, the more extensive the changes adopted in response to PA	No
H 3.5: The higher the auditees' agreement with audit criteria and audit conclusions, the more extensive the changes adopted in response to PA	No
H 3.6: The more concrete the recommendations made in the audit report, the more extensive the changes adopted in response to PA	No
H 3.7: The more attention performance audit received from the media, the more extensive the changes adopted in response to PA	Partial
H 3.8: The more attention the parliamentarians paid to the performance audit, the more extensive the changes adopted in response to PA	Yes

changes it brought about, the auditees can perceive the performance audit to be useful even if it does not lead to specific changes in policies or organizational practices.

Third, our analysis indicates that when examining the factors that facilitate the implementation of changes recommended in the audit report, it is useful to distinguish explicitly between *internal* and *external* factors because these factors may play a different role and have different weights in terms of their influence. Such an analytical approach can help to better disentangle the factors that can enhance the impacts of PA and to develop better normative suggestions for improving PA. As our empirical analysis suggests, while the internal factors – the expertise of the audit team, the quality of the audit, the willingness of the auditors to engage in a dialogue during the audit process, the auditees' agreement with the audit criteria and audit conclusions, and the concreteness of the audit recommendations – played a significant role influencing the

perceived utility of the performance audit, they appeared to have less bearing on inducing changes in response to PA.

Fourth, our survey results corroborate the findings of the existing empirical studies (undertaken in Canada and the US) with regard to the role of parliamentarians in facilitating the adoption of changes in response to PA (Brown and Craft 1980; Morin 2004, 2008, 2014). Finding a similar pattern in a very different context can be viewed as offering additional support to the conclusion that explicit actions by the members of the legislature can play an important role in increasing the effectiveness of PA in bringing about changes in public sector organizations.

Fifth, our study demonstrates that when scrutinizing the influence of media coverage on the impacts of PA, it is worth paying attention to whether media attention led to more extensive political debates and increased pressure on the minister and administration from political opponents. Our survey results suggest that degree of media attention *alone* was not correlated with the adopted changes, but when the media attention led to political debate and increased pressure from the opposition, the media's role became significant.

Sixth, our survey reveals that among the explanations offered by the auditees themselves the predominant explanation for *not* undertaking *changes* in response to the recommendations in the audit report was that the audited organization was *already aware* of the challenges and had put measures in place to correct them. This is a novel finding that has not been observed by the existing empirical studies. It suggests that in further research about the impacts of PA it would be useful to take an in-depth look at the changes that are implemented during the time frame in which the audit is taking place and what kind of factors trigger these changes. In particular, one could pay attention to the mechanisms that are at play here. First, it is possible that the attention of the auditors to a specific sector or policy area can act as an *accelerator* of change that had been contemplated by the auditees already. Second, it is also plausible (especially when the audits are lengthy) that during the early meetings between the auditors and auditees, the latter internalize the ideas of the auditors and come to view them “as their own” (with the result that they view themselves as the main change actors rather than the audit(ors) when asked about the reasons for change).

Finally, our empirical findings suggests that in many cases PA does not lead to the implementation of changes within the audited organizations, thus providing contradictory evidence to several existing empirical studies which have found PA to produce changes (see Van Looke and Put 2011). Thus, we would argue that still more studies are needed to investigate whether and how PA gives rise to changes in public sector organizations. In future studies it would also be useful to distinguish, for example, between short-term and long-term impacts of PA, small-scale and large-scale changes, and symbolic and actual steps taken in the audited organizations in response to the audit.

In interpreting and drawing conclusions from our empirical study – especially with regard to the impacts of performance audits on organizational changes – it has to be kept in mind that the survey measured the *perceptions* of the audited officials, which, we are aware, is a clear limitation of our study. In future research about the impacts of performance audit, it would be useful to triangulate survey findings with more in-depth qualitative case studies about specific audits and to use process-tracing to examine closer the mechanisms through which the factors we have discussed influence the impacts of performance audit.

References

- Alwardat, Y.A. (2010). External auditors and clients: an investigation of perceptions of value for money (VfM) Audit Practices in the UK Public Sector. PhD thesis, University of Westminster, Harrow Business School, London.
- Behn, R. (2001). *Rethinking democratic accountability*. Washington: Brookings Institution Press.
- Bemelmans-Videc, M.-L., Lonsdale, J., & Perrin, B. (2007). *Making accountability work*. New Brunswick: Transaction Publishers.
- Bendor, J., Taylor, S., & Van Gaalen, R. (1987). Politicians, bureaucrats and asymmetric information. *American Journal of Political Science*, 31, 796–828.
- Brown, R., & Craft, J. (1980). Auditing and public administration: the unrealized partnership. *Public Administration Review*, 40(3), 259–265.
- Dubnick, M. (2005). Accountability and the promise of performance: in search of mechanisms. *Public Performance and Management Review*, 28(3), 376–417.
- Funkhouser, M. (2011). Accountability, performance and performance auditing: reconciling the view of scholars and auditors. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contribution to accountability in democratic government*. Cheltenham: Edward Elgar.
- Furubo, J. E. (2011). Performance auditing: Audit or misnomer? In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contributing to accountability in democratic government*. Cheltenham: Edward Elgar.
- Hatherly, D. J., & Parker, L. D. (1988). Performance auditing outcomes: A comparative study. *Financial Accountability & Management*, 4(1), 21–41.
- Justesen, L., & Skaerbek, P. (2010). Performance auditing and the narrating of a new auditee identity. *Financial Accountability & Management*, 3(26), 325–343.
- Kells, S. (2011). The Seven deadly sins of performance auditing: implications for monitoring public audit institutions. *Australian Accounting Review*, 21(59), 383–396.
- Lapsley, I., & Pong, C. K. M. (2000). Modernization versus problematization: value-for- money audit in public services. *The European Accounting Review*, 9(4), 541–567.
- Leeuw, F. L. (2009). Evaluation: a booming business but is it adding value? *Evaluation Journal of Australasia*, 9(1), 3–9.
- Leeuw, F. L. (2011). On the effects, lack of effects and perverse effects of performance audit. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contribution to accountability in democratic government*. Cheltenham: Edward Elgar.
- Lonsdale, J. (1999). Impacts. In C. Pollitt, X. Girre, J. Lonsdale, R. Mul, H. Summa & M. Waerness (Eds.), *Performance or compliance? Performance audit and public management in five countries*. Oxford: Oxford University Press.
- Lonsdale, J. (2000). Developments in value-for-money audits: impacts and implications. *International Review of Administrative Sciences*, 66, 73–89.
- Lonsdale, J. (2008). Balancing independence and responsiveness: a practitioner perspective on the relationships shaping performance audit. *Evaluation*, 14, 227–248.
- Lonsdale, J. (2011). Introduction. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contribution to accountability in democratic government*. Cheltenham: Edward Elgar.
- Lonsdale, J., & Bechberger, E. (2011). Learning in an accountability setting. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contribution to accountability in democratic government*. Cheltenham: Edward Elgar.
- March, J. G., & Olsen, J. P. (1989). *Rediscovering institutions: The organizational basis of politics*. London: Collier Macmillan Publishers.
- Meyer, J. W., & Rowan, B. (1977). Institutionalised organizations: formal structure as myth and ceremony. *American Journal of Sociology*, 83, 340–363.
- Moe, T. M. (1984). The new economics of organization. *American Journal of Political Science*, 28, 739–777.
- Morin, D. (2001). Influence of value for money audit on public administrations: looking beyond appearances. *Financial Accountability and Management*, 17(2), 99–118.
- Morin, D. (2003). Controllers or catalysts for change and improvement: would the real value for money auditors please stand up? *Managerial Auditing Journal*, 18(1), 19–30.
- Morin, D. (2004). Measuring the impact of value- for- money audits: a model for surveying audited managers. *Canadian Public Administration*, 47(2), 141–164.
- Morin, D. (2008). Auditors general’s universe revisited. An exploratory study of the influence they exert on public administration through their value for money audits. *Managerial Auditing Journal*, 23(7), 697–720.

- Morin, D. (2014). Auditors general's impact on administrations: a Pan-Canadian study. *Managerial Auditing Journal*, 29(5), 395–426.
- Moynihan, D. P. (2006). What do we talk about when we talk about performance? dialogue theory and performance budgeting. *Journal of Public Administration Research and Theory*, 16(2), 151–168.
- Moynihan, D. P. (2008a). *The dynamics of performance management: Constructing information and reform*. Washington: Georgetown University Press.
- Moynihan, D. P. (2008b). Advocacy and learning: An interactive-dialogue approach to performance information use. In W. Van Dooren & S. Van de Walle (Eds.), *Performance information in the public sector: How it is used*. Basingstoke: Palgrave Mc Millan.
- Nõmm, K., & Randma-Liiv, T. (2012). Performance measurement and performance information in new democracies. *Public Management Review*, 14(7), 859–879.
- OECD. (2011). *Estonia: towards a single government approach. Assessment and recommendations*. Accessible at: http://www.valitsus.ee/UserFiles/valitsus/et/riigikantslei/uldinfo/dokumendiregister/Uuringud/PGR_Estonia_A&R.pdf.
- Pollitt, C., & Summa, H. (1996). Performance, audit and evaluation: similar tools, different relationships? *New Directions for Evaluation*, 71, 29–50.
- Pollitt, C., & Summa, H. (1997). Reflexive watchdogs? How supreme audit institutions account for themselves. *Public Administration*, 75(2), 313–336.
- Pollitt, C., Girre, X., Lonsdale, J., Mul, R., Summa, H., & Waerness, M. (1999). *Performance or compliance? Performance audit and public management in five countries*. New York: Oxford University Press.
- Pollitt, C., & Bouckaert, G. (2004). *Public management reform: a comparative analysis* (2nd ed.). New York: Oxford University Press.
- Popper, M., & Lipshitz, R. (1998). Organizational learning mechanisms: a structural and cultural approach to organizational learning. *The Journal of Applied Behavioral Science*, 34(2), 161–179.
- Preskill, H., & Torres, R. (2000). The learning dimension of evaluation use. In V. Caracelli & H. Preskill (Eds.), *The expanding scope of evaluation use*. San Francisco: Jossey-Bass Publishers.
- Raven, B. H. (1992). A power/interaction model of interpersonal influence: French and Raven 30 years later. *Journal of Social Behavior & Personality*, 7(2), 217–244.
- Reichborn-Kjennerud, K., & Johnsen, A. (2011). Auditors' understanding of evidence: a performance audit of an urban development programme. *Evaluation*, 17(3), 217–231.
- Reichborn-Kjennerud, K. (2013a). Political accountability and performance audit: the case of the auditor general in Norway. *Public Administration*, 91(3), 680–695.
- Reichborn-Kjennerud, K. (2013b). Resistance to control: Norwegian ministries' and agencies' reactions to performance audit. *Forthcoming in Public Organization Review*. doi:10.1007/s11115-013-0247-6.
- Talbot, C. (2010). *Theories of performance: Organizational and service improvement in the public domain*. Oxford: Oxford University Press.
- Tedeschi, J. T. (Ed.). (1972). *The social influence processes*. Chicago: Aldine Atherton.
- Van Der Meer, F. B. (1999). Evaluation and the social construction of impacts. *Evaluation*, 5(4), 387–406.
- Van der Meer, F.-B., & Edelenbos, J. (2006). Evaluation in multi-actor policy processes. *Evaluation*, 12(2), 201–218.
- Van Loocke, E., & Put, V. (2011). The impact of performance audits: A review of existing evidence. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contribution to accountability in democratic government*. Cheltenham: Edward Elgar.
- Weets, K. (2011). Impact at local government level: A multiple case study. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing: Contribution to accountability in democratic government*. Cheltenham: Edward Elgar.
- Weiss, C. (1998). *Evaluation* (2nd ed.). New Jersey: Prentice-Hall.
- Wilkins, P., & Lonsdale, J. (2007). Public sector auditing for accountability: New directions, new tricks? In M.-L. Bemelmans-Videc, J. Lonsdale, & B. Perrin (Eds.), *Making accountability work: Dilemmas for evaluation and for audit*. New Jersey: Transaction.

Ringa Raudla is a professor of public finance and governance at Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Estonia. Her main research interests are fiscal governance, budgeting, institutional economics, performance management, public sector accounting and public management reforms. Her recent publications include articles in *Public Administration*, *Governance*, *Journal of Public Policy*, *Policy Studies*, *European Journal of Law and Economics*, and *Accounting, Auditing & Accountability Journal*.

Külli Taro is a PhD candidate at Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Estonia. Her main research interests include performance management, measurement and auditing in the public sector. She worked at the National Audit Office of Estonia for many years. Currently, she is the head of Public Governance Program at a think tank founded by the President of Estonia. She has also published in *Public Management Review*.

Cherlin Agu is graduate student at Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology. Her research interests are budgeting, fiscal policy and public management. She is currently working as a chief specialist at the central training program in the Estonian Ministry of Finance.

James W. Douglas is a professor in the Department of Political Science & Public Administration at University of North Carolina at Charlotte. His research specialization is in the area of public budgeting and finance, and his articles have appeared in journals such as *Public Administration Review*, *Journal of Public Administration Research and Theory*, *Public Budgeting and Finance*, *American Review of Public Administration*, *American Journal of Political Science*, and *Journal of Policy Analysis and Management*.

III Taro, Külli and Eerik Hanni. 2015. “Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite-Sports Policy: The Case of Estonia.” *Administrative Culture* 16 (1), 24-46. (1.1)

Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite-Sports Policy: The Case of Estonia

Küllli Taro and Eerik Hanni

ABSTRACT

In 2011, the Estonian National Audit Office conducted a performance audit of the Estonian national elite-sports policy. The elite-sports system was assessed based on the international research methodology called SPLISS (Sports Policy Factors Leading to International Sporting Success) developed by a group of international scholars. The article demonstrates how scientific research can contribute to performance auditing. It helps auditors to fulfill the task of giving more in-depth exploratory, not only descriptive, analysis. However, the Estonian case demonstrates that success in the highly competitive international sporting arena may occur despite the shortcomings in public policy. Other contextual factors outside public policy and the availability of resources may lead to success. Although SPLISS aims for comprehensive and systematic diagnosis of a sports system, it does not fully explain why some nations are more successful in international competitions than others. SPLISS is useful in developing an audit methodology, but the unclear cause-and-effect relationship between sports policy and success of athletes creates difficulties in giving policy advice. But advising governments on better policy performance is considered to be an essential part of a modern performance audit.

Keywords: elite-sports policies, SPLISS, Estonia, performance audit, performance measurement

1. Introduction

Sports development is increasingly becoming a public-policy priority. Governments, with the help of different state agencies, are getting more and more involved in the delivery and management of sports. The trend of governmentalization is accompanied with globalization and commercialization (Houlihan and Green 2008). Competition in international sports is increasing, and more nations are adopting strategic approaches to develop world-class athletes. This phenomenon has resulted in increasing amounts of money being invested in elite-sports development by many

nations. There is also a growing interest in tracing the factors leading to international sports success (De Bosscher et al. 2008, 2011; Green and Oakley 2001; Grix and Carmichael 2012). Increasing public funds invested in elite sports call for more in-depth analysis of the effective use of those public resources and sports-policy performance.

In 2011, the Estonian National Audit Office conducted a performance audit of the Estonian national elite-sports system. The effectiveness of the sports system was assessed based on a comparative international research methodology called SPLISS (Sports Policy Factors Leading to International Sporting Success) (Riigikontroll 2012). In 2015, a comparative SPLISS study involving 15 nations (including Estonia) will be published (De Bosscher et al. 2015). Estonian data in this forthcoming study was collected during the performance audit in 2011. The authors of the article at hand also conducted the audit.

The results of the SPLISS study indicated several underdeveloped elite-sports policy areas in Estonia (e.g. talent identification, athletic and post-career support) that are considered to be critical for sporting success. The overall assessment of the Estonian elite-sports policy case does not promise sporting success in the international arena. But contrary to these expectations, Estonian athletes have been fairly successful in the international arena, especially in per-capita results. In this study, success in elite sports is defined as rankings in international sporting indexes and change in the number of medals won at the Olympic Games and international championships.

According to the World Sporting Index developed by the UK Sport, in per-capita results Estonia ranked 7th globally in 2006 and 15th in 2010 (Riigikontroll 2012). Furthermore, a very recent alternative analysis by Sporting Intelligence states that Estonia is even the world's second best nation at winning gold medals when the size of the population is taken into account (the first is Jamaica) (Harris 2012).

In 2000, Estonian athletes won a total of 62 medals at the Olympic Games, world championships and European championships. Taking into account some setback years this number has increased to 138 in 2014 (Estonian Sports Register). As the competition has increased as well as the number of events, the market share of medals has not risen accordingly.

Since regaining its independence in 1991, the most successful Olympic Games for Estonia were the Torino Winter Games in 2006. In Torino, Estonian athletes won three gold medals (2.04% of all medal points, 42nd place in country ranking) (Raudsepp et al. 2013). Since 2000 Estonian athletes have won at least one medal in every Olympic Games, except in Sochi where 24th place was the best result. The Estonian Olympic Committee declared that the performance of the Estonian athletes at the Sochi Olympic Games was "below expectations" (Estonian Olympic Committee 2014b).

The performance audit of the Estonian elite-sports system concentrated on the use of public funds. The SPLISS framework was utilized to identify and describe the essential parts of the sports system and develop audit criteria. Audit criteria are needed for comparing facts ("what is") against norms ("what should be"). SPLISS has defined the desirable "what should be" state of elite-sports policy (Riigikontroll 2012, Put 2011). This audit was a first attempt to put Estonian elite-sports policies

into such an all-encompassing and multi-dimensional framework. But it did not go into the conceptual issues of performance measurement and did not aim to explain the gap between the high status of Estonia in international rankings and its moderate to low scores on sports policy in the SPLISS framework.

The article at hand focuses on the usefulness of the SPLISS framework in analyzing elite-sports-policy effectiveness. The research question is whether the comparative methodology of SPLISS is suitable for use in the performance audit. Does it help to define good performance of the Estonian elite-sports policy and fulfil the performance-audit task of giving advice on how to improve policy performance?

Based solely on the Estonian case, we aim to conceptualize the performance measurement in the field of elite-sports policies in a context of a small state where resources are scarce. The case of elite-sports policy is especially interesting because the policy field could be considered easily measurable by the results in international competitions. This academic insight into the Estonian case could contribute to the performance-audit methodology as well as SPLISS-framework development. It also helps to further the significantly meager discussion on elite-sports public policy in Estonia.

In general, the effectiveness of a public policy can be defined as the extent to which the policies are achieving the objectives that they are supposed to achieve (Pollitt and Bouckaert 2004; European Commission 2009). Objectives can be described as outputs (direct results of an activity) or outcomes (wider societal aims of public policy). In the context of the current article and the SPLISS study, the objective of elite-sports policy is defined as success in international competitions.

At first, some conceptual issues of elite-sports-policy performance measurement in the context of elite-sports policy are discussed. Secondly, the SPLISS model, Estonian data collection, data analysis and results are described in more detail. This is followed by the discussion on the usefulness of the SPLISS model for performance measurement in a performance audit and its limitations in providing governments with policy advice.

2. Conceptualizing performance audit in the public sector

Public-sector performance has been at the very center of public-management debate in the last 15-20 years (Pollitt and Bouckaert 2004; Talbot 2010). Alongside the rise of New Public Management reforms and the more recent economic crisis, the public sector worldwide has been under constant pressure to improve its performance in pursuit of more efficiency and effectiveness, and in order to revive the citizens' trust in public institutions. Despite the conceptual controversy and severe academic criticism (see e.g. Bevan and Hood 2006; Pollitt and Bouckaert 2004; Radnor 2008; Talbot 2010; Van de Walle 2009), performance management and measurement have become more extensive, intensive and external during the last few decades (Pollitt and Bouckaert 2004).

Following the rise of performance movements, Supreme Audit Institutions (SAI) of many countries have introduced the practice of performance auditing. Nowadays, performance audit is considered an essential element of "an audit portfolio" for an SAI (Lonsdale 2011). By holding a substantial, if not predominant, position and

gaining much public attention it has even shifted the overall balance of the audit work in SAIs (Lonsdale 2011; Pollitt and Summa 1999). For example, in the Estonian SAI, performance audit takes up nearly 60 percent of audit staff. In recent years, over one-third of the published audit reports have been performance audits.

There are many definitions of performance audit (see Furubo 2011, INTOSAI 2003, Pollitt and Summa 1999, for example). The definition by the International Organization of Supreme Audit Institutions (INTOSAI), probably the best-known one, is based on the concept of the three E's (economy, efficiency and effectiveness)

Performance auditing is an independent and objective examination of government undertakings, systems, programs or organizations, with regard to one or more of the three aspects of economy, efficiency and effectiveness, aiming to lead to improvements.

The concept of the three E's is derived from the performance management "input-output-outcome" model (see Pollitt and Bouckaert 2004; Pollitt and Summa 1999, 9-10). The model envisages public policy as a process in which certain inputs (e.g. financial support, skills, organizational authority in the case of sports policy) are combined using a defined process (talent identification, training facilities, coaching etc.) with the aim to produce a specific "product" or "output" (performance of elite athletes) that will hopefully lead to desired outcomes (e.g. increased participation in leisure sports, country's prestige in the international arena). Outputs are the final products of government activity. Outcomes are the consequences of outputs and describe wider societal aims, often influenced by other externalities outside government policy. There is also a division between intermediate (in the short term) and final (in the long term) outcomes (Van Dooren et al. 2010, 16-36).

Government activities in elite-sports policy are clearly targeted mainly at one objective (output), that is, to perform successfully against the best athletes, mostly during international competitions (De Bosscher et al. 2011). There are various methods by which the success of elite athletes and national sports systems can be measured: the number of medals won during the Olympic Games or other events; top six or eight places; the relative success (e.g. controlling for population, wealth) or even the number of participants qualifying to take part (De Bosscher et al. 2010). All of these methods appear to correlate significantly (ibid.).

Though the output of elite-sports policies is clearly measured by medals won, the wider societal objectives (outcomes) of elite sports are hazy (Shibli et al. 2013). It is evidential that "sports for all" policy is part of a healthy lifestyle extending peoples' lives. But what is the societal aim of elite sports? According to common assumption, elite sports support participation in leisure sports. Olympic and world champions are expected to be role models encouraging thousands of people to take up some form of sports activity. Elite-sports success is thought to lead to a better image abroad, to bolster national identity and to stimulate domestic mass participation. This leads to a healthy nation and a wider pool from which to choose future elite stars, which, again, leads to increased elite-sports success (Grix and Carmichael 2012). However, there are some scholars who state that a causal relation between the "sports for all" and elite sports is disputable (van Bottenburg 2002; Hanstad and Skille 2010).

Traditional definitions of the three E's operationalize the production process and input-output-outcome ratio. Economy is associated purely with the measurement of inputs. The purpose of economic activity is to minimize the resource consumption. The combination of inputs and outputs defines the efficiency of the public policy. Efficiency is explained as "keeping the costs down" or "getting the most out of the given input". Effectiveness means achieving the intended outcomes from an activity. It is also defined as output and outcome ratio. (Pollitt and Summa 1999; Van Dooren et al. 2010)

Although the INTOSAI standard quotes the three main performance-audit criteria (sometimes also called norms) of economy, efficiency and effectiveness, modern performance audits have gone much further in interpreting what is "good management" or "good performance". In practice, most performance audits concern an assessment of management and policy practices, rather than a direct assessment of effectiveness and efficiency (Pollitt and Summa 1999; Put 2011). As performance audit is first and foremost targeted at evaluating government activities and the use of public funds, it is expected to be concerned with output evaluation, in the sense that it can assess only what is directly influenced by the government. Output evaluation refers to the use of efficiency criteria, but evaluating efficiency has turned out to be especially complicated in the public sector because of the absence of adequate comparative data to evaluate whether there is enough output for a given input (Stone 2001, 61-85). In reality, efficiency evaluation is replaced with the assessment of effectiveness, and the definition of outcome has been shifted to the final goal of public policy, which is under the control of the government activity. Effectiveness evaluation, therefore, has obtained the meaning of assessing whether public policy has achieved the defined goals (output or outcome) (Riigikontroll 2009; Van Dooren et al. 2010).

Regardless of the specific definition of outputs, outcomes, efficiency and effectiveness, any assessment in performance audit is based on a comparison of facts ("what is") against norms ("what should be") (INTOSAI 2003, Pollitt and Summa 1999, Put 2011), thus evaluating programs, processes or effects of policy using specified criteria. Without audit criteria auditors would not know what constitutes good public management and could not come up with clear conclusions in an audit. This search for adequate and applicable audit criteria has led SAIs to turn to scientific research.

Also the more recent academic discourse on performance audit has retreated from the concept of the three E's and application of the production-process logic. Nowadays, the discourse on performance audit is more focused on its role and tasks. One of the most recent definitions by Furubo (2011) states that

Performance auditing is an evaluative activity which produces assessments regarding performance or information about performance, of such a reliable degree and with such a freedom from investigatory and reporting constraints, that they can be used in the realm of accountability.

This emphasizes that performance audit is an evaluative activity. It has gone beyond concerns with regularity and compliance. Performance audit has developed from

being a mere accountability mechanism towards contributing to the wider policy debate. In addition to reporting on performance, it is also expected to give policy advice to governments and to give guidance on how public administration can improve in keeping up with the wider performance agenda. To achieve this end, audit reports should include explanatory analysis of a policy process. However, most performance audits contain information that explains what went wrong, as opposed to why things went wrong (Put 2011). Therefore, in order to formulate relevant recommendations and contribute to improving the performance of the government agencies, SAIs need information that explains shortfalls in the expected results. These more “diagnostic” performance audits (Put 2011) should overcome the explanatory deficiency of auditing.

As one of many forms of policy analysis and evaluation, performance audit is “a hybrid activity, making use of whatever means it can find to generate sufficiently robust evidence for its purposes” (Lonsdale 2011, 15). Although the role of a performance auditor may be crucially different from that of an evaluator by being part of the authoritative system of control (see Pollitt and Summa 1999), performance auditors turn more and more to the wide array of evaluation and scientific methods, making use of the best available qualitative as well as quantitative methods.

3. The SPLISS research model

While the literature on elite-sports development has been rapidly growing during the last decade, scholarly research lacks an all-encompassing framework for the evaluation of the effectiveness of elite-sports policies in relation to its goals (De Bosscher et al. 2011). The deficiency in the literature is caused by the difficulty of directly comparing nations and sports systems on a comparable basis, due to cultural differences and the uniqueness of each sports system. Furthermore, there is also a lack of publicly available and quantifiable data on sports policies (De Bosscher et al. 2010).

In 2002 a consortium of research groups from three nations (Belgium, the Netherlands and the United Kingdom) launched an international comparative study on elite-sports policies to fill the gap in scientific research on the relationship between elite-sports policies and international sporting success and to benchmark their nation against other competitors. This initial project was named SPLISS 1.0 (“SPLISS” stands for Sports Policy factors Leading to International Sporting Success) and lasted until 2008. It involved an overall elite-sports comparison of elite-sports policies in six nations (Belgium, Canada, Italy, the Netherlands, Norway and the United Kingdom). The pilot project was aimed at developing a research model to analyze why some nations succeed and others fail in high-performance sports (De Bosscher et al. 2010).

The second stage of the benchmarking study, called SPLISS 2.0, was initiated to better understand which (and how) sports policies lead to international sporting success and to obtain a better insight into the effectiveness and efficiency of elite-sports policies of nations at an overall sports level. Data collection was organized between 2010-2014 from 15 countries (Australia, Belgium, Brazil, Canada, Denmark, Estonia, Finland, France, Japan, the Netherlands, Northern Ireland, Portugal, Spain, South Korea and Switzerland). The SPLISS 2.0 study builds on the model and meth-

odologies developed in SPLISS 1.0, and its comparative results will be published by the end of the year 2015.

There are several factors determining a country's success in elite sports. These factors could be classified at three levels: micro-level factors, such as athletes' genetic predisposition or family support, macro-level factors such as macro-economic well-being, climate and population, and finally meso-level factors such as sports policies and politics. Macro-level factors have a major impact on elite-sports success, but they are largely out of the control of policy-makers (De Bosscher et al. 2010, 2015). In recent decades macro-level factors, such as GDP or population, have become less good predictors of nations' performance than they had been earlier. The main reason is that nations put more resources into elite-sports development (Green and Oakley 2001). In successful sporting countries elite-sports development models become increasingly similar. Sports have grown increasingly central to policy or the strategic-level role. Sports success enhances the sports organizations and makes governments increase their spending on elite-sports support. Along with intensified competition sports organizations around the world are looking for more effective operating models. These factors are determined by sports policies. Elite athletes have a greater chance of success depending on the effectiveness of policy and the investment made in elite sports. It is suggested that besides the factors determining the personal success of an individual athlete it is also possible to distinguish certain sports-policy factors that have an impact on the overall sporting success of a nation (De Bosscher et al. 2010).

The SPLISS methodology is based on the concept of nine sports policy areas or "pillars" which have an impact on the success of the elite-sports system. The influencing policy areas are as follows:

- Pillar 1: financial support – the public expenditure at the national level of elite sports;
- Pillar 2: organization and structure of sports policies (integrated approach to policy development) – effective management, working communication system and clear task descriptions, simple administration through common sporting and political boundaries;
- Pillar 3: participation in sports – a broad base of sports participation provides a supply of young talents and the opportunity for training;
- Pillar 4: talent identification and development system – a well-organized scouting and development system provides new talents;
- Pillar 5: athletic and post-career system – in only a few sports can athletes make a living from their sports earnings, some countries have programs for supporting their athletes financially and also support their university studies or retraining after finishing their elite-sports career – it helps athlete to concentrate on his/her elite-sports career while needed;
- Pillar 6: training facilities – training in a high-quality environment;
- Pillar 7: coaching provision and coach development – the qualification and motivation system for coaches;

The Case of Estonia

- Pillar 8: (inter)national competition – sufficient competition experience for athletes;
- Pillar 9: scientific research – the systematic gathering and dissemination of scientific information, innovation.

The nine-pillar model is based on a review of literature, pilot studies and interviews with athletes and high-performance experts. Each pillar is made measureable by 6-21 critical success factors. Each critical success factor is measured by one to four different research instruments: 1) the overall sports-policy inventory – a specific questionnaire on every pillar assembled by experts to be answered by researchers collecting data via interviews or using secondary data sources, document analysis etc; 2) an elite athletes' survey; 3) an elite coaches' survey and 4) a sports federations' survey.

The researchers behind SPLISS model suggest that the application of theories derived from the organizational-effectiveness literature in relation to national sports organizations (micro-level) and applying them at a macro or country/region level could present a viable option through which one can develop a method to evaluate the effectiveness of elite-sports policies (De Bosscher et al. 2010, 2011). The operational research model of SPLISS is similar to the process model of “input-output-outcome” that performance audit is based on. In SPLISS it is described as an “input-throughput-output-(outcome) model” (De Bosscher et al. 2010, 2011, 2015; Shibli 2013; see Figure 1).



Sources: De Bosscher et al. 2010, 2011, 2015; Shibli 2013

Figure 1: **SPLISS conceptual model**

* The SPLISS framework does not involve outcome analysis. It is admitted that this needs further research.

However, the confusion with the definition of efficiency and effectiveness described in the previous section of this article is also reflected in the SPLISS study on elite-sports policy. It is claimed that SPLISS evaluates the “effectiveness of elite-sports policies from a multidimensional perspective” (De Bosscher et al. 2010, 2011, 2015), but it is actually targeted to output, not outcome, evaluation (success of elite athletes in international competitions) (see Figure 1). Outcomes or wider objectives of elite-sports policy on society are not the subject of the SPLISS study (ibid.). Therefore, similarly to the developments in the performance-audit concept, effectiveness is defined through a policy objective to perform successfully against the best athletes in competitions, not as the output and outcome ratio. Also, the term efficiency is adapted to the SPLISS framework. Throughputs (see Figure 1) refer to the efficiency of sports policies, that is, the optimal way the inputs can be managed to produce the

required outputs in elite sports (De Bosscher et al. 2008). Efficiency in this context is not merely a “keeping the costs down” activity or “cost of a medal” indicator. The SPLISS study does not go into the discussion of the most efficient sports system, but concentrates on achieving better output (success of athletes) of elite-sports policies by defining it as an effectiveness evaluation.

4. Data collection and analysis in Estonia

In Estonia, the data was collected and analyzed by the National Audit Office of Estonia (ENAO) in 2011-2012 as part of a performance audit of elite-sports policy and spending (see Riigikontroll 2012). The data collection corresponded to the methodology presented in the original SPLISS framework used in the SPLISS 1.0 and 2.0 studies (De Bosscher et al. 2008, 2010, 2015). Only the scoring system differs from the forthcoming SPLISS 2.0 study in two aspects. Firstly, those 117 critical success factors (6-21 factors in 9 pillars) containing quantitative and qualitative data are aggregated into a final percentage score for each pillar, but the methodology foresees that not all factors have equal weight in the final score. Prior to the data collection in Estonia, the top-level managers of the Estonian Ministry of Culture and the Estonian Olympic Committee were interviewed. Interviewees were asked to give their evaluation to critical success factors and the appropriateness of critical success factors describing and assessing the development of the Estonian sports system. The evaluation was given on a scale from 1 to 10 and it was used later for weighting data collected by surveys and inventory. But in SPLISS studies these weights are given by international experts, and they may considerably differ from the Estonian scoring system, and those weights given have also changed in the SPLISS 2.0 study, compared to the pilot SPLISS 1.0 study. Secondly, some success factors presume that there is available comparative data from other countries because evaluation is given on a relative not an absolute scale. Scores are given compared to the situation in other countries, not only as expected by the experts and participants in the system. But by the time of the Estonian analysis this comparative data was scarce. The data collection for the SPLISS 2.0 study (which also includes the Estonian data) took place in 15 different countries at the same time, and it was not possible to use collected data for comparison in the Estonian analysis due to different data-collecting and -processing time schedules. Therefore, the Estonian scores were calculated using the so-called absolute scale; scales minimum and maximum values were not based on the reference data (with the exception of the first pillar). The only exception is the first pillar (financial support), where data from the SPLISS pilot study (2003) was available for reference.

This empirical work involved a detailed questionnaire answered by athletes, coaches and national governing bodies' performance directors. This large-scale survey was complemented by additional interviews and secondary data analysis. Prior to the data collection a small number of key persons in the Estonian sports system was also asked to give an assessment of the relevancy of those success factors proposed by SPLISS in the case of the Estonian system.

The data on critical success factors was collected by 1) overall sports-policy

inventory (214 questions); 2) the elite athletes’ survey (61 questions); 3) the elite coaches’ survey (62 questions) and 4) the sports federations’ survey (56 questions).

The data for the overall sports-policy inventory was collected by semi-structured interviews, secondary data (statistics, surveys) and document analysis (legislation, state budgets). The elite athletes’, coaches’ and sports federations’ surveys aimed to collect the information on factors not easy to assess on the binary scale (yes/no). Respondents were asked to give an assessment, usually on a 5-point ordinal Likert scale ranging from very good to very bad. In the case of dichotomous questions scores were calculated according to the proportion of positive responses, in the case of rating-scale questions, according to the difference between the proportions of positive and negative ratings (see Table 1).

Table 1: Illustration of the Points Attributed to Two Types of Questions in the Elite-Sports Climate Survey: Dichotomous Questions and Ratings

Dichotomous questions (yes/no)		Ratings	
Yes	Score	% (positive–negative)	Evaluation
0-20%	1	<-19.9%	1
20.1-40%	2	-19.9-0%	2
40.1-60%	3	0.1-20.0%	3
60.1-80%	4	20.1-50.0%	4
80.1-100%	5	> 50.0%	5

Source: De Bosscher et al. 2010

All athletes participating in the Estonian Olympic Committee’s Olympic preparation program for the London and Sochi Olympics, also the promising talents suggested by the sports federations and athletes that finished their career in the last two years, were asked to answer the elite athletes’ survey. A questionnaire was sent to a total of 122 athletes, 82 or 67% of whom answered.

The elite coaches’ survey was sent to 420 highly qualified coaches (in the two highest qualification categories). A total of 190 coaches responded to the survey (45% of those who were sent a questionnaire). According to SPLISS the elite coaches’ samples should have focused more narrowly on the elite athletes’ coaches and coaches training young talents only. But in the Estonian case the sample was expanded to all coaches holding a higher-qualification certificate to increase the number of respondents and get a somewhat broader picture of the situation. If only those coaches who trained elite athletes during the survey period would have been engaged, the number would have been too small for adequate conclusions to be drawn. The results of the survey showed, however, that most of the coaches responding to the questionnaire were at some point of their career engaged in elite athletes’ or young-talent training.

All heads of sports federations represented at the Olympics were asked to respond to the sports federations’ survey. A questionnaire was sent to a total of 34 sports federations, 19 or 56% of whom responded.

5. Results of the SPLISS study in Estonia

The evaluation according to the SPLISS framework revealed that most of the pillars were moderately developed. The best score was achieved by the financial-support pillar meaning that compared to other countries the amount of money allocated to sports was greater *in the comparative scale*. However, no single pillar was very underdeveloped; the lowest scores were given to the talent-identification and -development system, the athletic and post-career system and to scientific research.

Preliminary results of the SPLISS 2.0 study (which will be published at the end of 2015) indicate that Estonian scores will be even lower than calculated in the analysis at hand (except for the talent-identification and -development system). Scores for most pillars will be below the average results of the 15 countries participating in the SPLISS 2.0 project. It shows that when results are compared with more recent data from other countries and factors are given different weights by foreign experts, then the Estonian elite-sports policy scores even lower than when evaluated by domestic athletes, coaches, sports federations and other experts.

Table 2: Scores given to elite-sports policy pillars in Estonia

Pillar	Score in scale 1(min)-5(max)	Description
Financial support*	4	Good level of development
Integrated approach to policy development	3	Moderate level of development
Participation in sport	3	Moderate level of development
Talent-identification and -development system	2	Limited development
Athletic and post-career support	2	Limited development
Training facilities	3	Moderate level of development
Coaching provision and coach development	3	Moderate level of development
(Inter)national competition	3	Moderate level of development
Scientific research	2	Limited development

* Evaluation given on a comparative scale according to the SPLISS pilot study data from 2003. Other pillars' scores were calculated using the absolute scale.
Source: Riigikontroll 2012; authors

The SPLISS methodology combines information gathered from many sources into a single rating. Ratings of one pillar are implicated by a number of critical success factors, some of which may receive the highest scores and others the lowest scores, giving the whole pillar the total moderate score. For example the coaching-provision and coach-development pillar's critical success factors got extremely bipolar scores. By the qualification the coaches were assessed as highly professional, but coaches' working conditions were assessed as poor.

Thorough analysis of survey data on pillars which got lower scores exposed a number of problematic areas. Therefore, we took a closer look at the pillar concerning coaching provision and coach development due to the extremely diverse estimates by athletes, sports-federation leaders and coaches themselves. These success factors were also estimated as important by the managers from the Ministry of Culture and the Estonian Olympic Committee during the critical success-factor estimation process prior to conducting the surveys. In addition, the financial-support pillar as the most controversial factor in comparing Estonian and international data needs more explanation.

5.1. Financial support

The central government of Estonia spends 15-17 million Euros a year on sports development and management. Approximately half of the budget is spent on elite sports. Additionally, local-government expenditure on sports activities fluctuates between 40 and 60 million Euros per year. This is mostly allocated to grassroots sports, but local governments may occasionally support elite athletes as well. In 2008, the Estonian central-government spending on sports activities was double the amount of money spent for the same purpose in the year 2000. However, as a result of the rapid growth in total budget during the same period, the proportion of the overall state budget spent on sports activities decreased from 1% to 0.4%. Following the financial crises, the total government spending decreased after 2008 by almost 30% in absolute terms. The percentage of the state budget spent on sports activities was rather stable between 2008 and 2013 but decreased to 0.2% in 2014 as a result of the increase in total budget (Estonian Olympic Committee 2014a; Riigikontroll 2012). Sports activities in Estonia are primarily funded by the public sector. This is similar to the tradition in Eastern European countries, where public funding of sports has a much larger role than in many Western countries (Eurostrategies et al. 2011). According to the official statistics private sponsorship constitutes only 6-8% of the total funding of sports activities.

The amount of financial support to elite-sports policies is probably the most controversial factor in international comparison, as this is directly dependent on the size and wealth of the country (pool of talents is the other factor dependent on the country's size/population). The government can decide on its public-policy priorities and favor some policy areas for others, but the total amount of resources available stays limited to smaller and less wealthy countries regardless of the priorities set.

As the study at hand is focused on government activities, more weight in evaluation is put on the relative amount of finances spent on elite-sports policies (as per head of population) and government expenditure. The SPLISS framework also takes into account the total national expenditure and the absolute cash terms. In Estonia, half of the government expenditure on sports is targeted towards elite sports; this proportion is lower on average in the 15 countries covered by the SPLISS study. It indicates that the government considers elite sports to be a policy priority. But in absolute terms, Estonian expenditure on elite sports is lacking when compared to other countries (De Bosscher et al. 2015). This explains why financial support is

considered to be at a good level of development in our study but receives much lower scores in SPLISS when put into an international comparative perspective. One of the main findings in the SPLISS 2.0 analysis was that the best predictor of output appears to be the absolute amount of funding allocated to elite sports. More money generally means more medals, though not quite automatically, as elite-sports success is the result of the way the resources are invested in a blend of factors (De Bosscher et al. 2013, 2015).

5.2. Coaches' professionalism

The SPLISS pilot study confirmed that critical success factors linked to coaches' work had a slightly more significant correlation towards elite-sports success – countries getting higher scores in the coaching-provision and coach-development pillar also proved to win more medals (De Bosscher et al. 2008).

In many countries most of the coaches work as volunteers, but in Estonia due to several reasons only few coaches do voluntary work, at least formally (Eurostrategies et al. 2011). Most of the coaches working as volunteers have lower qualification than respondents in the present survey. Coaches answering to the questions in the present survey should be highly qualified in order to be able to train elite athletes or prepare young talents to become elite athletes.

Issues related to the appreciation of coaches' work and work conditions have been a subject of the public discussion in Estonia for the last decade. The sports-system reform at the beginning of the present century abolished a number of government sports schools and created the environment for the establishment of private sports clubs. The reform was meant also to attract more private funding to the sports system but actually led to a reduction of social guarantees and was associated with social insecurity for coaches.

In the surveys all three target groups (elite athletes, coaches and sports-federation leaders) were asked to name the areas which needed most effort by decision makers and resources put into development. All three named coaching provision and coach development most frequently. The coaches' survey helped bring out some of the most problematic issues which need to be dealt with primarily. Firstly, 31% of the coaches work without an employment contract, probably getting paid via different scholarships schemes, which does not guarantee any social benefits, for example unemployment and health insurance, in case they will be needed. Shortly after the SPLISS analysis in Estonia was conducted the Estonian Tax and Customs Board took measures to stop coaches' work-financing through scholarship schemes as this is not in accordance with the tax law. In 2014 and 2015 the government allocated some extraordinary public funds to compensate the additional tax costs to the sports clubs and federations. The issue of financing coaching, especially for young athletes, is still open for discussion and requires a public-policy decision from the government.

The survey showed that 30% of the coaches spent personal money for their training job – a common situation in case of voluntary work or hobby but slightly surprising in the case of high-level professionals. 60% of the coaches could not find enough time to adequately supervise their most talented athletes. Finally 66% of the coaches who responded said that their living and employment conditions were poor. Also all

the sports-federation leaders answered that their federation could not offer world-class living conditions to the coaches. At the same time 74% of the athletes responded that their coaches' competence was at the world's highest level.

It is possible to conclude that nearly one-third of the highly qualified coaches are working factually as volunteers or part-time workers because they do not get paid at all or for a full job. Most coaches need to find other jobs to get enough income for a living. The success or failure of Estonian athletes is based largely on coaches' enthusiasm. However, there is resentment towards the insufficient state activities to adequately support professional coaching.

5.3. Talents

Talent identification and development with the help of special training and support services is considered a key factor of a successful sports system. However, for several reasons, this is an underdeveloped area in most countries. Attention was drawn to the need for a systematic talent-identification system, financial and scientific support to sports federations in creating such a system, special support services or programs for the young talents and the existence of a legal framework that would allow young talents to be treated, taking into account the elite athletes' needs (for example a flexible education system).

Out of 19 sports federations that responded to the questionnaire only 7 stated that they had some sort of monitoring system for talent search and only 3 estimated it to generate results. Only 1 monitoring system relied on scientific evidence, and only 1 sports federation systematized collected information on the development of young athletes in a database. However despite the lack of systematic talent identification the talented young athletes are given more attention by the sports federations, clubs and coaches. Most often more frequent and more intensive training is available to them (73% of the athletes responded) as well as medical support services (70%) and combining education and training (26%). 60% of the young athletes had been offered special treatment in recognition of their status as an elite athlete by their secondary school or university. In most cases special treatment meant flexible class schedules, minimized attendance obligations, flexible examinations or individual study; only 25% of the athletes responded that they considered the support offered to them as young athletes sufficient to support their career.

It appeared that the most difficult, but crucial period for becoming a successful athlete were the years right after graduating secondary school (starting from the age of 18-19 in Estonia). Amateur athletes at that age are not usually able to achieve internationally outstanding results that could guarantee them special funding by the Olympic Committee. According to elite athletes it takes 3-5 years to be clear if the young athlete could be successful at the international level. During this time, it is extremely difficult for young athletes to combine their training schedule, follow the right nutrition plan, attend university/vocational school and also have essential time for resting. These years are the most important during an athlete's career. However, during these years it is also most difficult to obtain supporting services and funding.

Talent identification and development is the only pillar that will probably score better in the SPLISS 2.0 study than in the evaluation at hand. It may indicate that

though there is no generic public system to identify young talents, the relative sporting success at international competitions proves that talents are found from a small pool of people (the Estonian population is 1.3 million). And once young talents are identified, then they get extra attention.

It needs to be pointed out that many successful athletes in Estonia come from families where parents are coaches themselves and often train their children (e.g. Olympic medalists Andrus Veerpalu and Kristiina Šmigun-Vähi or more recent successful athletes like Rasmus Mägi or Kelly Sildaru). Talent identification and development thus comes from within the family.

5.4. Athletic and post-career support

The support services ensure athletes' good physical and mental preparation for competition. Post-career support is aimed at preparing athletes for the activities after the end of their career – legal advice, direct financial support or support for their studies, for example. Support services were a fairly well-developed area in most countries participating in the SPLISS study.

The athletes' survey confirmed that participation in international competitions was most commonly supported. Financial support for this purpose was received by 61% of the surveyed athletes. Travel costs were covered for 57%. Somewhat less common was the reimbursement of purchases of sports equipment (43%) and training-related costs (43%). Rarely did benefits occur that were indirectly related to sporting activities, such as free or favorably priced cars (13%), as well as free meals at training facilities (17%).

The most accessible support services oriented towards athletes' physical conditions were massages, which have been used by 95% of the surveyed athletes, sports-medical service (86%), strength and endurance tests (73%) and physiotherapy (71%). All other services that were included into the questionnaire were used with lower frequency and were not available to most of the athletes (biomechanics analysis, nutrition coaching, physiological and psychological counseling, legal, financial and career counseling). For young athletes it is more difficult to access support services. More than half of the athletes (59%) had massage services available at their main training base, to get any other service athletes would usually have to go to find it elsewhere.

Career-planning services are offered for athletes only in exceptional cases. More than half (53%) of the athletes surveyed acknowledged, however, that they were concerned about their post-career activities, 36% admitted that these concerns were interfering with, and had negatively affected, their career as elite athletes. While career planning services are usually not available for everybody, the Estonian Olympic Committee has initiated several actions over the last decade to improve the athletes' post-career opportunities. For example, universities and defense forces have increased the number of athletes on their payroll or using scholarships provided by them; also special grants are offered in collaboration with the International Olympic Committee for the athletes, including free training and support in their search of employment. However, there was no information about athletes' use of the latter.

In Estonia, many athletes have used the opportunity to choose and hire their support staff themselves, independently from the national team. These personalized

teams are not directly connected to the government activities or official sports policies, but athletes may also get support services and funding from the government and the Olympic Committee. This is a common feature in elite sports worldwide. In addition to a personalized approach, in the Estonian case a reason for this is also the small total number of athletes. There is often only a single or few athletes in a discipline, especially competing internationally.

6. SPLISS as a tool for performance evaluation

Models like SPLISS help to define “good performance” and “good management” in order to assess the performance of government institutions. As noted earlier, the search for adequate and applicable audit criteria in order to evaluate and explain, as well as improve, government policy has lead SAIs to turn to scientific research (Pollitt and Summa 1999; Lonsdale 2011). Although SPLISS recognizes the importance of other factors situated at the macro level (e.g. population, wealth, natural resources, cultural factors, religion) and the micro level (e.g. the individual athlete and their close environment), it focuses on the meso-level factors that could be influenced by the sports policies (De Bosscher et al. 2010). This makes it a beneficial systematic tool for policy analysis and particularly suitable for evaluating the activities of state institutions in public-sector performance audits.

In the case of the audit “Activities of the State in Supporting Elite Sports” SPLISS was used as a starting point to develop audit methodology, as well as a source of comparative information about elite-sports policies in other countries. The most important benefit for the audit was the exploitation of a mapping system (definition of sports-policy components) created by the SPLISS group. This contributed significantly to the preliminary analysis and audit planning phase. As a result, it was relatively simple to get a compendious picture of the national sports system.

From the methodological point of view, it has to be noted that all of the SPLISS pillars definitely do not have an equal role in leading to sporting success. One can argue whether one or the other critical success factor is important enough and should affect the total score of some pillars (whether the weighting system of different sub-factors is adequate), but taking into account the systematic approach to the sports system, the individual sub-factors do not have too much significance, because sports leaders were directed to assess separately the importance of any single factor. It is still possible to get a good overview of the sports system. The model helps to ascertain its essential parts and problems, though what the importance of those single factors is, is still not clear. Apparently, this downside will benefit from the international comparisons.

The SPLISS framework is also useful for comparing nations. Nevertheless, it should not be isolated from qualitative descriptions and from a broader understanding of an elite-sports system. Therefore, for the purpose of an in-depth analysis of the Estonian elite-sports support system, we added to the methodology some more interviews with politicians, (former) athletes, sports managers, coaches, researchers and public servants responsible for sports administration. In addition to that, a focus group of experts, public servants and politicians was organized to discuss the results of SPLISS.

The main critique on the SPLISS analytical model stems from the conceptual problem of policy analysis: how to differentiate between cause-and-effects relationship and pure correlation. This is also called the “attribution problem” (Talbot 2010, 48). In public-policy analysis, correlation is often mistaken for causation. Although, the case of elite-sports policies is unique because outcomes are clearly measurable (international success), it is still difficult to attribute changes to public-sector activities in a reliable and valid way. The developers of the SPLISS framework also remain somewhat vague in the discussion whether the model explains causality in sports policy-making or not. On the one hand, they claim that SPLISS is a helpful tool to find the causal relationship between sports policy and success of the athletes at an international level (De Bosscher et al. 2010). On the other hand, in the Flanders case, it was found that in spite of the increasing elite-sports expenditures in Flanders (inputs) and the development of the throughputs (processes), this has not as yet led to better results (outputs) at an international level (De Bosscher et al. 2011). And the Estonian case demonstrated that a country may be highly ranked in international sporting indexes despite the deficiencies in its sports-policy factors. It seems that the SPLISS model presents an opportunity to compare what different countries are doing, but actually does not explain what determines success or failure in elite sports. Even the developers of the SPLISS model admit the limitation. Inputs and throughputs (policy factors) in this model are the sports-policy factors that may increase chances of international success (outputs), but do not guarantee success (De Bosscher et al. 2010).

In most public policy areas, outputs and outcomes (for the difference, see Pollitt and Summa 1999) are difficult to measure. Researchers also disagree on elite-sports policies’ outcomes (e.g. boosting grassroots sports participation, enhancing the international reputation of a country, see van Bottenburg 2002; Grix and Carmichel 2012; Hanstad and Skille 2010). But the output of elite-sports policies is clearly defined, that is, medals won at international competitions. A widely accepted output definition of elite sports, however, does not make the policy analysis easy. It also needs to be taken into account that in sports, the results are indeed usually clearly measurable, but the result is mainly only a position compared to other athletes. The accomplishments of an athlete are therefore relative to the success of other athletes and the level of competition. For example, the number of events at the Olympic Games has risen from 145 in 1956 and to 302 in 2012. During this period there has been an increase by more than three times in the number of participating countries and participating athletes. Also the number of medal-winning countries has risen from 37 in 1956 and 64 in 1992 to 85 in 2012 (Sports Reference). These numbers suggest that the competition for medals has increased significantly. There is direct correlation between the number of participating countries and the number of countries winning medals and as a result, many nations have seen their share of international success reduced (De Bosscher et al. 2008).

7. SPLISS as a tool for policy advice

Performance audits are expected not only to evaluate the use of public resources but also to give advice to the government in approving its activities and increasing the efficiency and effectiveness of public policy. Governments aim for long-term strategic planning and more effective, evidence-based policy-making. There is a growing belief that nations that increasingly plan for success will increase their chances of success (De Bosscher et al. 2008). This puts pressure on performance audits to contribute to policy-making.

SPLISS is a useful tool to identify which factors (pillars) of the sports system are well developed, and which are deficient, but it still does not really explain why some nations are more successful in international sports competitions than others. This explanatory weakness refers to the fact that recommendations given based on SPLISS need to be carefully considered.

The role of macro- and micro-level factors also raises the discussion about the possibilities of cross-national sports studies. Cross-national sports studies are often considered to be problematic because sports are embedded in a broader cultural context, where beliefs, norms, and values have been shown to have had a marked impact on the character of sports policy (Houlihan and Green 2008). Successful policy instruments are often dependent on the local situation (e.g. politics, geographical situation, ethical values), which implies that similar policy actions may have different outcomes in different nations (De Bosscher et al. 2010). Consequently, there is no single recipe for competitiveness in sports. As micro- and macro-level factors differ and are independent of the government's policy in every country, there could not be a unified single approach to achieve international sporting success. "One size fits all" recommendations may ignore the danger of policy failure caused by uninformed, incomplete or inappropriate policy transfer (Dolowitz and Marsh 2000; Rose 1993).

The SPLISS study has shown that countries have to invest more in order to be successful in the escalating global sporting arms race. This means a need for more money in cash terms, as the best predictor of output appears to be the absolute amount of funding allocated to elite sports (De Bosscher et al. 2013, 2015). It makes it increasingly difficult for small countries with fewer resources to compete in the global arena. Estonia will never be able to invest as much money as Canada or Finland. And just making a recommendation to increase resources spent on elite sports is inappropriate in the Estonian context. Also, with a population of only 1.3 million, the pool of talent stays limited. But in elite sports, the rules of the game are dictated by what rival nations are doing, not on the basis of what an individual country is doing (De Bosscher et al. 2013, 2015). The Estonian elite-sports policies may be efficient, in terms of their use of available resources, but they will not compete with big countries.

Taking into account the importance of money for international sporting success, and the relatively low scores given to its elite-sports policies by the SPLISS methodology, Estonia should not have a considerable chance in this sporting arms race. But medals won at the Olympic Games and other championships indicate that this could be an irregular case. This finding suggests that the other factors outside the meso level (sports policies) may have a significant influence on the success of nations.

Meso-level factors form only a small part of the overall factors leading to international sporting success. Arguably an additional three highly influential dimensions were identified in the preliminary explorative surveys of SPLISS (media attention, specific sports culture and athletes' personal environments, e.g. parental support). Though important, but outside the direct impact of the government's action, they were therefore excluded from the SPLISS study (De Bosscher et al. 2010). In these macro- and micro-level factors the chance may hide for success for small countries like Estonia. Does this mean that the role of the state becomes minor? Not necessarily. SPLISS findings still indicate that, although a minimum amount of funding seems necessary for success, elite-sports success is the result of the way the resources are invested in a blend of factors (De Bosscher et al. 2013). Instead of focusing on what it does not have (i.e. resources), public policies should enhance the opportunities (i.e. personalized approach to talents, small and efficient teams, working ethics, parental support etc.) of a small state.

8. Conclusion

In Estonia, a considerable share of the elite-sports system is financed with public funds. That puts great expectations on the performance of elite-sports policies as well as pressure for the efficient and effective use of public funds. It is the task of the performance audit to evaluate the performance of the government in achieving goals of public polices and give advice in improving policy performance.

The Estonian National Audit Office conducted a performance audit of the Estonian elite-sports system and used the international comparative study framework called SPLISS to develop the audit methodology. SPLISS turned out to be useful in identifying the essential parts of an elite-sports policy and the weaknesses of the sports system compared to what is expected, based on the practice in other countries and "best practice" defined by policy experts, athletes and coaches. In audit terminology it means that SPLISS was used for formulating audit criteria to identify what is considered "good management" or "good performance". However, the use of SPLISS is limited in fulfilling the advisory function of performance audit. SPLISS aims for comprehensive and systematic diagnosis of a sports system, as well as a comparison of countries, but it does not fully explain why some nations are more successful in international competitions than others.

The key findings of the Estonian SPLISS study demonstrated that despite the relative success of elite athletes in international sports competitions most of the factors crucial to the sports system are at the moderate or even limited level of development. There is no talent-identification and -development system in the case of most sports. Post-career support for athletes is not sufficient. Mainly financial support for highly qualified coaches is lacking. There are problems with implementing results of scientific research in order to innovate in the training process and there is not enough local applied research. Therefore, the success in elite sports is not coherently the result of the Estonian sports system, i.e. government policy and activities. The Estonian case revealed a controversial situation. The country has been relatively successful in international competitions, which is the main aim of elite-sports policy, but this does not represent the effectiveness of the public policy.

Though Estonian governmental financial support for sports per capita is greater than in several old EU countries, the total amount of funds invested in elite sports is lacking. And the SPLISS study has revealed that the best predictor of output is the absolute amount of funding allocated to elite sports. This finding is especially crucial for small states like Estonia. Elite-sports policies and government actions may be efficient and effective, taking into account the resources available, but this will not be sufficient for participating in the escalating global sporting arms race. Elite-sporting success appears to be the outcome of a multivariate process involving many pillars, not only the amount of financing. A good organization of other critical factors also contributes to the success. And finally, success may occur at random or unsystematically.

Elite-sports policy is probably one of few public-policy fields where measuring policy outputs seems to be fairly straightforward. Methods vary, but in principle, success can be clearly measured in medals won at international competitions. However, this does not explain whether the success of athletes is the result of public policies and government involvement. The analysis of the causal relationship between public policies and elite-sports outputs needs more understanding of sports-policy factors leading to international sporting success. Also, views about elite-sports public outcome or broader societal aims, i.e. reasoning for government involvement, vary significantly.

The SPLISS model stands out among other research in the field because of the use of quantitative data and the introduction of “measurement” in an area (comparing elite-sports policies and systems) that is essentially qualitative in nature (De Bosscher et al. 2010). For a performance audit, the value of the SPLISS model is not so much the exact score of different pillars but the identification of the success factors. By focusing on the meso-level (i.e. public-policy) analysis, it is a useful analytical tool to evaluate the government performance.

While appreciating the work of scholars in developing this comprehensive SPLISS framework, the Estonian case pointed out the limits of scientific methods in explaining the cause-and-effect relationship of sports policies and the success of elite athletes. Elite athletes may still be successful despite the existing sports system in their country. And on the other hand, even a good system cannot overrule the importance of micro- and macro-level factors, such as personal abilities, cultural background, natural resources or climate.

REFERENCES

- Bevan, G. and C. Hood. 2006. "What's Measured is What Matters: Targets and Gaming in the English Public Health Care System." *Public Administration* 84(3), 517-538.
- van Bottenburg, M. 2002. "Sport for All and Elite Sport: Do They Benefit One Another?" Paper presented at the IX World Sport for All Congress, Papendal, the Netherlands, 27-30 October 2002.
- De Bosscher, V., J. Bingham, S. Shibli, M. van Bottenburg and P. De Knop. 2008. *The Global Sporting Arms Race: An International Comparative Study on Sports Policy Factors Leading to International Sporting Success*. Aachen: Meyer & Meyer.
- De Bosscher, V., S. Shibli, M. van Bottenburg, P. De Knop and J. Truyens. 2010. "Developing a Method for Comparing the Elite Sport Systems and Policies of Nations: A Mixed Research Methods Approach." *Journal of Sports Management* 24, 567-600.
- De Bosscher, V., S. Shibli, H. Westerbeek and M. van Bottenburg. 2015 (forthcoming). *Successful Elite Sport Policies: An International Comparison of the Sports Policy Factors Leading to International Sporting Success (SPLISS 2.0) in 15 Nations*. Aachen: Meyer & Meyer.
- De Bosscher, V., D. Shilbury, M. Theeboom, J. Van Hoecke and P. De Knop. 2011. "Effectiveness of National Elite Sport Policies: A Multidimensional Approach Applied to the Case of Flanders." *European Sport Management Quarterly* 11(2), 115-141.
- De Bosscher, V., M. van Bottenburg, S. Shibli and H. Westerbeek. 2013. "Does More Money in Equal More Medals out? An International Comparison in 15 Countries, Results of the SPLISS 2.0 Study." Abstract available at www.easm.net (last accessed 25.02.2015).
- Dolowitz, D. and D. Marsh. 2000. "Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making." *Governance: An International Journal of Policy and Administration* 13(1), 5-24.
- Estonian Olympic Committee. 2014a VIII Sport Congress materials. Available in Estonian at <http://www.eok.ee/viii-eesti-spordi-kongress> (last accessed 25.02.2015).
- Estonian Olympic Committee. 2014b. Press release. Available in Estonian at <http://www.eok.ee/eok-esindajate-kogu-kinnitas-rio-ol%C3%BCmpiam-%C3%A4ngude-kvalifikatsiooninormid> (last accessed 25.02.2015).
- Estonian Sports Register. Sports statistics available in Estonian at www.spodiregister.ee (last accessed 25.02.2015).
- European Commission. 2009. Impact Assessment Guidelines, SEC(2009)92.
- Eurostrategies, Amnyos, CDES and German Sport Institute of Cologne. 2011. *Study on the Funding of Grassroots Sports in the EU: With a Focus on the Internal Market Aspects Concerning Legislative Frameworks and Systems of Financing*. Report prepared for the European Commission DG Internal Market and Services. Final Report Volume II – Country Reports. Available

- at http://ec.europa.eu/internal_market/top_layer/docs/FinalReportVoll_en.pdf (last accessed 25.02.2015).
- Furubo, J.-E. 2011. "Performance Auditing: Audit or Misnomer?" In J. Lonsdale, P. Wilkins and T. Ling (eds). *Performance Auditing: Contributing to Accountability in Democratic Government*. Cheltenham: Edward Elgar Publishing, 22-50.
- Green, M. and B. Oakley. 2001. "Elite Sport Development Systems and Playing to Win: Uniformity and Diversity in International Approaches." *Leisure Studies* 20, 247-267.
- Grix, J. and F. Carmichael. 2012. "Why do Governments Invest in Elite Sport? A Polemic." *International Journal of Sport Policy and Politics* 4(1), 73-90.
- Hanstad, D.V. and E.Å. Skille. 2010. "Does Elite Sport Develop Mass Sport? A Norwegian Case Study." *Scandinavian Sport Studies Forum* 1, 51-68.
- Harris, N. 2012. "The Alternative Best Olympic Nations: Jamaica, Estonia, Mongolia, NZ, Georgia ... and Australia." *Sporting Intelligence*. Available at <http://www.sportingintelligence.com/2012/03/01/the-alternative-best-olympic-nations-jamaica-estonia-mongolia-nz-georgia-and-australia-010302/> (last accessed 25.02.2015).
- Houlihan, B. and M. Green. 2008. *Comparative Elite Sport Development: Systems, Structures and Public Policy*. London: Elsevier.
- INTOSAI (International Organization of Supreme Audit Institutions). 2003. Performance Audit Guidelines – Key Principles. The International Standards of Supreme Audit Institutions (ISSAI 3100).
- Lonsdale, J. 2011. "Introduction." In J. Lonsdale, P. Wilkins and T. Ling (eds). *Performance Auditing: Contributing to Accountability in Democratic Government*. Cheltenham: Edward Elgar Publishing, 1-21.
- Pollitt, C. and G. Bouckaert. 2004. *Public Management Reform: A Comparative Analysis*. 2nd edn. New York: Oxford University Press.
- Pollitt, C. and H. Summa. 1999. "Performance Audit and Public Management Reform." In C. Pollitt, X. Girre, J. Lonsdale, R. Mul, H. Summa and M. Waerness (eds). *Performance or Compliance? Performance Audit and Public Management in Five Countries*. New York: Oxford University Press, 1-10.
- Put, V. 2011. "Norms Used: Some Strategic Considerations from the Netherlands and the UK." In J. Lonsdale, P. Wilkins and T. Ling (eds). *Performance Auditing: Contributing to Accountability in Democratic Government*. Cheltenham: Edward Elgar Publishing, 75-94.
- Radnor, Z. 2008. "Hitting the Target and Missing the Point? Developing an Understanding of Organizational Gaming." In W. Van Dooren and S. Van de Walle (eds). *Performance Information in the Public Sector: How It Is Used*. Basingstoke: Palgrave Macmillan.
- Raudsepp, L., V. Ööpik and P. Lusmägi. 2013. "Estonia." In K. Hallmann and K. Petry (eds). *Comparative Sport Development: Sports Economics, Management and Policy*. New York: Springer, 33-45.
- Riigikontroll [National Audit Office of Estonia]. 2009. Auditikäsiraamat [Audit Guidelines]. Available in Estonian at <http://www.riigikontroll.ee> (last accessed 25.02.2015).

- Riigikontroll [National Audit Office of Estonia]. 2012. "Riigi tegevus tippspordi toetamisel" [Activities of the state in supporting elite sports: Audit report]. Summary in English available at <http://www.riigikontroll.ee> (last accessed 25.02.2015).
- Rose, R. 1993. *Lesson-drawing in Public Policy: A Guide to Learning across Time and Space*. New Jersey: Chatham House Publishers, Inc.
- Shibli, S., V. De Bosscher, M. van Bottenburg and H. Westerbeek. 2013. "Measuring Performance and Success in Elite Sports." In P. Sotiriadou and V. De Bosscher (eds). *Managing High Performance Sport*. London: Routledge, 30-45. Sports Reference. Sports statistics available at <http://www.sports-reference.com/> (last accessed 25.02.2015).
- Stone, D. 2001. *Policy Paradox: The Art of Political Decision Making*. New York: W.W. Norton & Company.
- Talbot, C. 2010. *Theories of Performance: Organizational and Service Improvement in the Public Domain*. Oxford, New York: Oxford University Press.
- Van de Walle, S. 2009. "International Comparisons of Public Sector Performance: How to Move ahead." *Public Management Review* 11(1), 39-56.
- Van Dooren, W., G. Bouckaert and J. Halligan. 2010. *Performance Management in the Public Sector*. London/New York: Routledge.

KÜLLI TARO is a PhD candidate at the Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Estonia. Her main research interests include performance management, measurement and auditing in the public sector. She worked at the National Audit Office of Estonia for many years. Currently, she is the head of the Public Governance Program at a think tank founded by the President of Estonia. Correspondence: Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Akadeemia tee 3, Tallinn 12618, Estonia; E-mail: kylli.taro@gmail.com. **EERIK HANNI** is a performance auditor at the National Audit Office of Estonia and a contact person for the SLPiSS project in Estonia. Correspondence: National Audit Office of Estonia, Narva mnt 11a, 15013 Tallinn, Estonia; E-mail: eerik.hanni@riigikontroll.ee.

CURRICULUM VITAE

1. Personal data

Name: Külli Taro (Nõmm)
Date and place of birth: 18 January 1980, Pärnumaa
Citizenship: Estonian
E-mail address: kylli.taro@gmail.com

2. Education

2003 - 2007 University of Tartu, Master of Arts in public administration and social policy
2001 - 2002 University of Konstanz, Department of Public Administration and Political Science (exchange student)
1999 - 2003 University of Tartu, Bachelor's Degree (BA) in public administration *cum laude* (BA degree equivalent to Bologna Master's Degree)

3. Language skills

Estonian native language
English proficient level
German mediate level
Russian intermediate level
Finnish basic level

4. Noteworthy international special courses

2009 "Restructuring Government to Overcome Crisis", UNDP Regional Centre for Public Administration Reform and State Chancellery of Republic of Latvia
2007 International Training Course on Financial and Value for Money Auditing, UK National Audit Office
2004 Policy Workshop on Burden of Disease and Cost Effectiveness Analysis, WHO and Ministry of Social Affairs

2002 “Imperfections of Governance: Challenges Posed by Environment, Economics and Globalisation”, University of Helsinki Summer School

5. Professional employment

2013 - 2015 Estonian Cooperation Assembly, Head of Public Governance Program

Member of the Local Government Reform Expert Committee, established by the Minister of Public Administration

Member of the Administrative Capacity Committee, established by the Secretary of State

2005 - 2012 National Audit Office of Estonia, Performance Audit Department, Audit Manager

Member of INTOSAI working Group for the Audit of Privatisation, Economic Regulation and Public Private Partnerships

2004 - 2005 National Audit Office of Estonia, Performance Audit Department, Auditor

2002 - 2004 Parliament of Estonia, Constitutional Committee, consultant

2001 PRAXIS Center for Policy Studies, Researcher

2000 - 2001 Ministry of Justice of Estonia, Member of the Working Group for drafting the Administrative Procedure Act

6. Defended theses

2007 “Teiste riikide näidete kasutamine poliitikakujundamisel ja muudatuste põhjendamisel” [Use of lesson-drawing in policy-making and political argumentation] (*Magister Artium*), University of Tartu

2003 “Poliitika ülevõtmine: pearahasüsteem Eesti perearstide rahastamisel” [Policy Transfer: The introduction of capitation-based payment financing scheme for family doctors] (*Baccalaureus Artium*), University of Tartu

7. Research activity

Peer-reviewed publications

Taro, K.; Hanni, E. 2015. "Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite-Sports Policy: The Case of Estonia." *Administrative Culture*, 24 - 46.

Taro, K. 2015. "In pursuit of efficient public administration: Reform proposals for the public sector." Vetik, R. (Ed.). *Estonian Human Development Report 2014/2015*, 250 - 255.

Raudla, R.; Douglas, J. W.; Taro, K.; Agu, C. 2015. "The Impact of Performance Audit on Public Sector Organizations: The Case of Estonia." *Public Organization Review*, Published online. DOI <http://dx.doi.org/10.1007/s11115-015-0308-0>

Taro, K.; Parrest, N. 2014. "Valikud Eesti haldusorganisatsiooni loomisel." [Options in Creating Estonian Administrative Organisation.] *Juridica*, X, 715 - 731. (in Estonian, summary available in English)

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.

Nõmm, K.; Loom, R.; Paulus, M. 2007. "Erivajadustega laste õppimisvõimalused." [Educational opportunities of children with special needs.] *Haridus* [Education] 1-2, 39 - 43.

Nõmm, K. 2005. "Tervishoiuvaldkonna töötajate kutsekindlus ja tegelik tööhõive." [Job-stability and employment of health professionals] *Eesti Arst* [The Estonian Medical Journal] 5, 369 - 374.

Nõmm, K.; Kuldjärv, M. 2001. "Riiklike teenuste kodanikukesksus ja uus tehnoloogia." [Citizen-Centered Public Services and New Technologies] *Riigikogu Toimetised* [The Journal of the Estonian Parliament] 4, 159 - 165.

Presentations at international conferences

Nõmm, K. "How to Deliver the Message? Using Research Methodology and Fun in Performance Audit." International Seminar in Performance Auditing in Hague, organized by the Netherlands Court of Audit, November 2011.

Nõmm, K. and Randma-Liiv, T. "International Organizations as Performance Management Promoters: the Case of the OECD Recommendations for Estonia." Paper presented at the EGPA conference in Bucharest, September 2011.

Nõmm, K. 2009. "Improving Audit Efficiency and Reliability with Electronic State Registries and Information Systems." Presentation for EUROSAI Training Event: "Experience with the implementation and development of performance audits in

reaction to challenges and opportunities in a changing environment” in Prague, April 2009.

Nõmm, K. 2008. „Experiences in Audit of Social Programmes in Field of Education”. Presentation for VII EUROSAI Congress in Krakow, June 2008.

Theses supervised

- 2015 Aleksi Stepanov “Tulemuspalga rakendamine Ida-Viru kohalikes omavalitsustes.” [The Implementation of Performance Related Pay in Ida-Viru Municipalities] Master’s thesis.
- 2015 Helen Kranich “Halduse efektiivsuse suurendamine avaliku teenus hinna muutmisega - kohtusse pöördumisel nõutavate riigilõivude näide.” [Increasing the Efficiency of Administration Due to Public Service Price Amendment - the Court Fees Example] Master’s thesis.
- 2011 Kristiina Kase “Siseauditi võimalused ja probleemid avaliku sektori organisatsiooni valitsemise hindamiseks Eesti Vabariigi ministeeriumide näitel.” [Opportunities and Problems for Internal Audit to Evaluate Public Sector Organization's Governance: The Case of the Ministries of the Republic of Estonia] Master’s thesis, co-supervised with Evelin Pungas.
- 2009 Jüri Vlassov “Euroopa Liidu Liikmesriikide julgeolekuasutuste ülesannete võrdlev analüüs.” [Comparative Analysis of the Tasks of Security Authorities in EU Member States] Master’s thesis, co-supervised with Erkki Karo.
- 2009 Maris Ode “Kulutuste jagamisel põhineva kõrghariduse finantseerimissüsteemi ja sissetulekutest sõltuvate tagasimaksetega õppelaenusüsteemi rakendamise võimalikkus Eestis.” [Possibilities for Introducing Cost Sharing for Financing Higher Education and Income-Contingent Financing of Student Loans in Estonia] Master’s thesis.

ELULOOKIRJELDUS

1. Isikuandmed

Ees- ja perekonnanimi: Külli Taro (Nõmm)
Sünniaeg ja -koht: 18. jaanuar 1980, Pärnumaa
Kodakondsus: Eesti
E-posti aadress: kylli.taro@gmail.com

2. Hariduskäik

2003 - 2007 Tartu Ülikool, teadusmagistri kraad (MA), avalik haldus ja sotsiaalpoliitika
2001 - 2002 Konstanzi Ülikool (Saksamaa), politoloogia ja avaliku halduse vahetusüliõpilane
1999 - 2003 Tartu Ülikool, avaliku halduse osakond, bakalaureusekraad *cum laude*

3. Keelteoskus

Eesti emakeel
Inglise kõrgtase
Saksa kesktase
Vene madalam kesktase
Soome algtase

4. Olulisem rahvusvaheline täiendusõpe

2009 UNDP Regional Centre for Public Administration Reform ja Läti Riigikantselei, "Restructuring Government to Overcome Crisis"
2007 UK National Audit Office, International Training Course on Financial and Value for Money Auditing

2004 WHO ja Sotsiaalministeerium, „Policy Workshop on Burden of Disease and Cost Effectiveness Analysis“

2002 Helsingi Ülikool, rahvusvaheline suveprogramm “Imperfections of Governance: Challenges Posed by Environment, Economics and Globalisation”

5. Töökogemus

2013 - 2015 SA Eesti Koostöö Kogu, riigivalitsemise programmi juht
Riigihalduse ministri loodud haldusreformi ekspertkomisjoni liige
Riigisekretäri loodud haldusvõimekuse valdkondliku komisjoni liige

2005 - 2012 Riigikontroll, auditijuht
INTOSAI privatiseerimise, majandusliku reguleerimise ja PPP töögrupi liige

2004 - 2005 Riigikontroll, audiitor

2002 - 2004 Riigikogu põhiseaduskomisjon, konsultant

2001 Poliitikauuringute Keskus Praxis, analüütik

2000 - 2001 Justiitsministeerium, haldusmenetluse seaduse ja selle rakenduseaduse väljatöötamise töögrupi liige

6. Kaitstud lõputööd

2007 Teiste riikide näidete kasutamine poliitikakujundamisel ja muudatuste põhjendamisel (*Magister Artium*), Tartu Ülikool

2003 Poliitika ülevõtmine: pearahasüsteem Eesti perearstide rahastamisel (*Baccalaureus Artium*), Tartu Ülikool

7. Teadustegevus

Eelretsenseeritud publikatsioonid

Taro, K.; Hanni, E. 2015. Using a Comparative Method in Performance Audit for Evaluating Effectiveness of the Elite-Sports Policy: The Case of Estonia. *Administrative Culture*, 24 - 46.

Taro, K. 2015. Efektiivse riigihalduse otsinguil: riigipidamise kava. Vetik, R. (Toim.). *Eesti Inimarengu Aruanne 2014/2015*, 250 - 255.

Raudla, R.; Douglas, J. W.; Taro, K.; Agu, C. 2015. The Impact of Performance Audit on Public Sector Organizations: The Case of Estonia. *Public Organization Review*, Avaldatud veebis. DOI <http://dx.doi.org/10.1007/s11115-015-0308-0>

Taro, K.; Parrest, N. 2014. Valikud Eesti haldusorganisatsiooni loomisel. *Juridica*, X, 715 - 731.

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.

Nõmm, K.; Loom, R.; Paulus, M. 2007. Erivajadustega laste õppimisvõimalused. *Haridus* 1-2, 39 - 43.

Nõmm, K. 2005. Tervishoiuvaldkonna töötajate kutsekindlus ja tegelik tööhõive. *Eesti Arst* 5, 369 - 374.

Nõmm, K.; Kuldjärv, M. 2001. Riiklike teenuste kodanikesksus ja uus tehnoloogia. *Riigikogu Toimetised* 4, 159 - 165.

Ettekanded rahvusvahelistel konverentsidel

Nõmm, K. "How to Deliver the Message? Using research methodology and fun in performance audit." International Seminar in Performance Auditing, the Netherlands Court of Audit, november 2011.

Nõmm, K. and Randma-Liiv, T. "International organizations as performance management promoters: the case of the OECD recommendations for Estonia." EGPA konverents Bucharestis, september 2011.

Nõmm, K. 2009. "Improving audit efficiency and reliability with electronic state registries and information systems." EUROSAI Training Event: "Experience with the implementation and development of performance audits in reaction to challenges and opportunities in a changing environment" Praha, aprill 2009.

Nõmm, K. 2008. „Experiences in Audit of Social Programmes in field of Education”. VII EUROSAI Congress. Krakow, juuni 2008.

Juhendatud lõputööd

- 2015 Aleksi Stepanov “Tulemuspalga rakendamine Ida-Viru kohalikes omavalitsustes.” Magistritöö.
- 2015 Helen Kranich “Halduse efektiivsuse suurendamine avaliku teenus hinna muutmisega - kohtusse pöördumisel nõutavate riigilõivude näide.” Magistritöö.
- 2011 Kristiina Kase “Siseauditi võimalused ja probleemid avaliku sektori organisatsiooni valitsemise hindamiseks Eesti Vabariigi ministriumide näitel.” Magistritöö, kaasjuhendaja Evelin Pungas.
- 2009 Jüri Vlassov “Euroopa Liidu Liikmesriikide julgeolekuasutuste ülesannete võrdlev analüüs.” Magistritöö, kaasjuhendaja Erkki Karo.
- 2009 Maris Ode “Kulutuste jagamisel põhineva kõrghariduse finantseerimissüsteemi ja sissetulekutest sõltuvate tagasimaksetega õppelaenusüsteemi rakendamise võimalikkus Eestis.” Magistritöö.

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESES
SERIES I: SOCIAL SCIENCES

6. **Ülle Madise.** Elections, Political Parties, and Legislative Performance in Estonia: Institutional Choices from the Return to Independence to the Rise of E-Democracy. 2007.
7. **Tarvo Kungla.** Patterns of Multi-Level Governance in Europe: The Challenge of the EU's Enlargement. 2007.
8. **Mikk Lõhmus.** Avaliku halduse detsentraliseerimine Eestis. 2008.
9. **Tarmo Kalvet.** Innovation Policy and Development in the ICT Paradigm: Regional and Theoretical Perspectives. 2009.
10. **Thomas Duve.** Die Verschuldung deutscher Gemeinden: Präventionsansätze im Spannungsverhältnis von kommunaler Selbstverwaltung und staatlicher Gesamtverantwortung. 2009.
11. **Veiko Lember.** Contracting-Out Public Services and Public Procurement for Innovation: Revisiting Contracting Limits in Estonia and Beyond. 2009.
12. **Raivo Linnas.** An Integrated Model of Audit, Control and Supervision of the Local Government Sector: The Case of Estonia. 2010.
13. **Simon Lang.** Transformations in European R&D and Regional Policies within the Multi-Level Governance Framework: The Changing Nature of the European Union Ten Years after the Launch of the Lisbon Strategy. 2010.
14. **Erkki Karo.** Governance of Innovation Policy in Catching-up Context: Theoretical Considerations and Case Studies of Central and Eastern European Economies. 2011.
15. **Margit Suurna.** Innovation and High-Technology Policy, Policy-Making and Implementation in Central and Eastern European Countries: The Case of Estonia. 2011.
16. **Marek Tiits.** Technology Foresight and the Catching-up Strategy in Small Countries: The Case of Estonia. 2011.
17. **Vasileios Kostakis.** The Political Economy of Information Production in the Social Web: Towards a "Partner State Approach". 2011.
18. **Jane Järvalt.** Strategic Human Resource Management in the Public Service: Evidence from Estonia and Other Central and Eastern European Countries. 2012.
19. **Robert Krimmer.** The Evolution of E-voting: Why Voting Technology is Used and How It Affects Democracy. 2012.

20. **Külli Sarapuu.** Mapping and Explaining Post-Communist Development of Administrative Structure: The Case of Estonian Public Administration 1990–2010. 2013.
21. **Kristi Joamets.** Gender as an Element of Marriage Capacity in the Context of National and Supranational Law in the European Union. 2014.
22. **Aleksandr Aidarov.** Estonian National Minorities' Cultures: Successes and Failures Policy Goals, Instruments and Organization. 2015.
23. **Riin Savi.** The Impact of Fiscal Crisis on Public Administration: Cutback Management and Changes in Decision-Making. 2015.
24. **Priit Vinkel.** Remote Electronic Voting in Estonia: Legality, Impact and Confidence. 2015.
25. **Külli Taro.** The Attribution Problem in Performance Measurement in the Public Sector: Lessons from Performance Audits in Estonia. 2016.