RADBOUD UNIVERSITY

COMPASS

MASTER'S THESIS

Evaluating evaluations

-

Political interest in evaluation reports on the functioning of ZBOs

---

Author:

Joep van der Spoel

s4497929

Supervisor:

Prof. dr. Sandra van Thiel

Second reader:

Dr. Ellen Mastenbroek

28.644 words

October 19, 2015

SUMMARY

ZBOs are semi-autonomous agencies, and have become a ubiquitous form of government in the Netherlands. A large amount of governmental tasks have been transfered to ZBOs in a process that is called *verzelfstandiging* or *autonomization* (Pollitt, Bathgate, Caulfield, Smullen, & Talbot, 2001). It is estimated that currently, about 40 per cent of the total spending of the Dutch government goes to *Zelfstandige Bestuursorganen*, or ZBOs (De Kruijf, 2011).

The shedding of governmental tasks and giving them to organizations that are expected to be able to autonomously decide how to execute them, has led to issues of political control and accountability. The Dutch Kaderwet ZBOs included measures to correct this perceived lack of control. Once every five years, ministers are obligated to evaluate the functioning of the ZBOs falling under their responsibility. However, a large amount of ZBOs is not evaluated. Strikingly, members of parliament show little interest in this fact. Moreover, members of parliament show little interest in the results of evaluations that did take place.

This thesis endeavors to provide insight in this political disinterest in evaluations of ZBOs. It investigates the reasons why some ZBOs are evaluated and some are not, it tries to explain the political motives to start an evaluation and it analyzes the (lack of) political interest in evaluation reports on ZBOs. The research question is therefore: *when, how and why do Dutch politicians decide to evaluate ZBOs?*

Theory and methods

First, the position of ZBOs in modern government is theoretically explained. ZBOs are semi-autonomous organizations that have come into existence as part of the wave of decentralization, autonomization and marketization of government functions that began in the 1980s in the Western world(Pollitt et al., 2001). A large amount of governmental tasks were given to these organizations, with the goal of creating a leaner, more efficient central government that focuses on its core tasks of policy making. The (semi-) autonomous organizations could carry out tasks of service delivery and policy implemen-

tation and had autonomy in performing these tasks as effective and efficient as possible. ZBOs are such relatively autonomous organizations, placed at a distance from the central government. However, their autonomy has limits. The central government can still scrutinize and control ZBOs, for instance by setting goals and standards on output with performance indicators. It is even argued that, paradoxically, a more autonomous position for ZBOs is coupled with tighter control by their parent ministries (Smullen, Van Thiel, & Pollitt, 2001).

Public Choice theory gives an explanation for the rise in use of (semi-)autonomous agencies such as ZBOs. Its core assumption is that both politicians and bureaucrats have self-interested goals at heart and that they act strategically to reach these goals. By placing tasks in the hands of autonomous organizations, politicians may gain strategic advantages that make it more likely for them to get (re-)elected. For instance, politicians can use a separate organization to avoid being blamed for bad performance, as they had distances themselves from the tasks of that organization. Furthermore, bureaucrats may have strategic motives: by placing the uninteresting tasks of policy implementation in the hands of other organizations, they could focus on making policy, which is regarded as more rewarding and status-increasing (James, 2003).

This theoretical perspective has consequences for evaluations of ZBOs. They can be used politically, to shift blame away from certain actors and onto others, (Bovens, 't Hart, & Kuipers, 2008), or to manage crises (Resodihardjo, 2006). Therefore, the content of the evaluation becomes the highly contested subject of a political game. Stakeholders try to influence what research body executes the evaluation, who takes place in this body, the scope and the goal of the research question, its methods, the resources it can use (Bovens et al., 2008). When the report is published, stakeholders may use its conclusions to criticize or defend policy choices.

Thus, in sum, the perspective of Public Choice informed the expectations of this thesis: there are political and strategic motives that make politicians decide (not to) evaluate a ZBO, and there are political and strategic reasons for why evaluation reports receive so little attention.

The content of a selection of evaluation reports and the resulting parliamentary debates on these reports were analyzed on three themes. The reports were first analyzed on how they *measured and evaluated the performance of ZBOs*. Second, the *decision-making process* leading up to the evaluation was analyzed. Which actors influenced the decision to evaluate a ZBO and what were the reasons to start an evaluation? Third, the *political effects of a report* when it entered the political arena was studied. How did

politicians react to a report, and what were topics that were important to members of parliament and ministers?

Results and conclusions

Measuring performance

Reports looked similar in structure and content. Reports devoted much attention to the internal processes of the ZBO, its background, organizational structure and environment, its relation to the parent ministry. Thus, reports interpreted and operationalized 'effectiveness and efficiency' in a broad way. The methods that were used were almost always qualitative: interviews and documents were the means to gather data. Quantitative data, benchmarks and baseline measurements were rarely used. The research committee was most often independent. Often, a consultancy agency was asked to do the evaluation, but sometimes a committee of experts was named, presumably by the client. It was however stressed in the reports that these committees were independent. Several reports were done by peer review committees. These peer review committees often evaluated a self-evaluation that was done by the ZBO in advance.

Decision-making process: initiative and reasons to evaluate

Most often, the parent ministry initiated the evaluation although a few reports that were sent to parliament in compliance with art. 39 of the Kaderwet ZBOs were started by the ZBO. When ZBOs initiated the evaluation, it was usually a self-evaluation combined with a visitation from a peer-review committee. Such reports often focused more on organization structure and culture than on performance on output. So, when ZBOs initiate an evaluation, they seem to be able influence the content of the report that is sent to parliament.

Reasons to evaluate varied. In most cases, art. 39 of the Kaderwet ZBOs was mentioned in the report or in the accompanying letters to parliament. Other reasons that were mentioned were the recent occurrence of incidents in which the ZBO played a part, a direct request from members of parliament to start an evaluation.

Political effects of the report

It was confirmed that political interest in evaluation reports on ZBOs is low. In order to get attention from members of parliament, reports have to jump through a few hoops. First and most importantly, reports have to be created. More than 50 per cent of the ZBOs is not evaluated at all. Second, reports must be sent to parliament, accompanied

by a ministerial letter. Most reports that were analyzed were sent to parliament. Third, reports have to be placed on the agenda of general parliamentary debates. Two-thirds of the reports were placed on such an agenda. Fourth, reports have to be actually discussed in that debate. Most reports were mentioned briefly in debates. About half of the reports that were placed on the agenda, were discussed in relative depth. On 8 reports, members of parliament sent written questions to the minister. When reports were discussed, political interest was focused on a large variety of topics, including the ZBOs performance, accountability and transparency, the ZBO's organizational struture, control mechanisms available to the minister, and topics unrelated to the ZBO and the report.

Conclusion

In conclusion, both reports as members of parliament show interest in a large variety of topics concerning the ZBO. Next to performance, the internal organization and structure, the ZBO's environment, its relation to the ministry, accountability, transparency and other topics were discussed. This could mean that members of parliament are not necessarily interested in whether the ZBO performs well, but rather that they are interested in whether the ZBO operates in a fair, just and reasonable way.

This shows that indeed the process of *autonomization* has not reduced the desire for strong mechanisms of accountability and control. However, reports on effectiveness and efficiency do not seem to be a very effective way to hold ZBOs accountable for the fairness and reasonableness of its operations. Current reports seem to be split between truly assessing performance on output and describing the ZBO's organizational processes.

PREFACE

My house mate once said to me in a very philosophical manner: "you should not let time be a constraint to a good scientific process". I have certainly taken my time to write this thesis. The start of this process was slow, a little frustrating and certainly not easy. Now that it is finished, I am quite happy with the result, though, primarily because I feel I have learned so much. I have of course gained some knowledge on ZBOs and evaluation reports, and I have probably improved my English writing skills, but far more importantly, I have learned to plan, manage and organize my own research project. In this sense, this thesis has given me valuable skills that I can most certainly apply in my future endeavors, which will most likely involve jobs and professional careers.

I would like to thank prof. dr. Sandra van Thiel, who has given me valuable advice and guidance during the process of writing this thesis. I always came back from supervision meetings with more energy and motivation than I had before. Many thanks also to the tutoring group I was part of, consisting of Arend de Boer, Timo Lagarde en Rosan Koolen. They have provided me with useful suggestions, and I hope that my suggestions to them have helped them a little, too. I am also very grateful to Peter van Goch, information specialist for the Dutch house of representatives. He invested much time to collect parliamentary records I would otherwise not have found. This has helped me tremendously.

CONTENTS

1. INTRODUCTION

Since the 1980s, privatization and marketization have resulted in the establishment of many autonomous administrative authorities - *Zelfstandige Bestuursorganen* (ZBOs). A large amount of policy programs and governmental tasks have been transferred to these organizations (Pollitt et al., 2001). The goal was to sever the ties between policy making and policy implementation, and to make service provision more effective and efficient.

Inevitably, however, this diminished political responsibility. Political leaders lost influence over the functioning of the ZBO to their bureaucratic subordinates. As a consequence, in times of crisis or unexpected circumstances, crucial decisions may have to be made by the bureaucratic leadership, of which political leaders can have only after-the-fact knowledge. An information gap between the officials of the independent organization and the political leaders might occur, causing principal-agent problems. In short, delegation of tasks to ZBOs may have caused the political leadership to lose the capacity to effectively control the actions of the ZBO (Bovens, 't Hart, & Van Twist, 2007).

The Dutch "Kaderwet ZBO's" of 2007 included measures to correct this perceived lack of political control over ZBOs. It obligates ministers to evaluate the ZBOs falling under their responsibility, once every five years. Strikingly, this evaluation does not always happen, and even more strikingly, there is little political interest in the evaluations that are carried out (Eerste Kamer, 2012).

1.1   Research question and objective

This raises many questions. Why are some ZBOs evaluated and others not? What are reasons to start, or to delay an evaluation? How do politicians react to and use an evaluation report in political debate? In this thesis, I will investigate this apparent lack of political interest in evaluating ZBOs. The goal is to determine the functions of an evaluation report for ministers and members of parliament. To reach this goal, the main question of this thesis is formulated as follows:

When, how and why do Dutch politicians decide to evaluate ZBOs?

The subquestions are:

1. *What are ZBOs?*

2. *What are functions of an evaluation for politicians?*

3. *What are the characteristics of an evaluation report of a ZBO?*

4. *What does the decision-making process leading up to the evaluation look like?*

5. *What is the impact of evaluation reports on political debates?*

6. *To what extent does a decision to evaluate reflect strategic/political motives for the use of evaluations?*

1.2   Societal and scientific relevance

ZBOs in the Netherlands have become a ubiquitous form of organizing government. However, it is unclear exactly how many ZBOs exist, and how many people they employ. The Dutch ZBO Registry lists 109 ZBOs and clusters of ZBOs on its website. [1] When these clusters are split into seperate ZBOs, a number of around 600 individual ZBOs can be counted (De Kruijf, 2011). It is estimated that around 80 billion euros in program costs are made, which is about 40 per cent of the total spending of the Dutch national government (De Kruijf, 2011). These numbers vary: ZBOs are abolished, new ZBOs are created, budgets are cut and increased. Although the numbers may vary, it is clear that ZBOs are a large part of the current way of organizing government. This makes them important subjects for research. It is important to know how ZBOs operate, how effective they are, whether there are negative side effects to this from of government and, if so, how these can be mitigated.

Although ZBOs in the Netherlands and semi-autonomous agencies throughout the Western world have become ubiquitous, research on them has only recently picked up speed. There is still much that we do not know. First, there is unclarity about the definition of types of agencies. Researchers are still mapping and categorizing the amount and types of agencies that exist in certain countries, and how much autonomy these agencies really have (Bouckaert & Peters, 2004). Second, the creation of ZBOs is expected to make government more effective and efficient. However, not

---

[1] https://almanak.zboregister.overheid.nl. Retrieved 11 October 2015

much research exists on the intended and *un*intended consequences of agencification (**james˙executive˙2003; james˙structural˙2011**). The existence of organizations at a distance from the government has consequences for how control and coordination between that organization and the parent ministry are structured. Research on these structures is scarce. Third, the performance of agencies has not been researched extensively. It is difficult to measure performance of ZBOs, partly because their tasks are often not easily expressed in measurable indicators, and because benchmarks do not always exist. Some attempts have been made, but these provide mixed results (Pollit & Dan, 2013). Some have negative conclusions (e.g. Overman & Van Thiel, 2015), others are more positive.

In this thesis I present research that can contribute to the field of research on agencies and ZBOs. Specifically it contributes to research on the *consequences* of agencification. It explores structures for political control of ZBOs. In doing so, its conclusions can shed light on the relation between politicians and ZBOs, and ultimately on the functioning of ZBOs in general.

1.3 Strategy

The first subquestion will be answered in the theoretical framework and provides the necessary background information that helps to fully understand the subject of ZBOs. The second subquestion is a theoretical question, too. I will use public choice theory and theory on the politics of evaluation to answer this question. The third, fourth, fifth and sixth subquestions form the core of the research. To answer the third subquestion I will investigate the content of evaluation reports of ZBOs. Among other indicators, I will look for variations in the methods used, the scope and goal of the research question employed, the formal reasons to execute the evaluation and the composition of the research question. The fourth subquestion focuses on the decision-making process. Who decides to start an evaluation, and for what reasons? Who influences the content of the evaluation? The fifth subquestion is focused on political interest after the report is published. The amount and the topics of attention of politicians will be studied. The the sixth subquestion binds the theoretical expectations and the empiric data together. It is analyzed whether strategic and political motives are visible in the reports and parliamentary debates.

1.4   Theory and methods

I will start with a discussion on the nature of ZBOs and their origins. After this, I will explain that evaluating a ZBO is a complex task, partly because performance is often difficult to measure. This complexity creates opportunities for stakeholders to contest the content of an evaluation report, or to use its conclusions to push for change in policy. I will show that carrying out an evaluation ought to be viewed as a *political* endeavor. This perspective leads to expectations regarding political interest in evaluation reports of ZBOs. In this chapter I provide a theoretical answer to the first and the second subquestion.

It is difficult to investigate the processes surrounding a decision *not* to evaluate the performance of an ZBO, because I could never know in what non-existing cases a conscious decision not to evaluate has been made. Therefore, it is more fruitful to start the research with an investigation into ZBOs that *were* evaluated on their performance. A content analysis of the many evaluation reports that exist is the first step. With this, I can answer the third subquestion, which concerns the content of the reports. I can search for patterns in type of research question, constellation of the research committees, scope and method, among other things. Next, I will analyze political debate on the reports. Most reports are sent to parliament with a letter from the minister responding to the conclusions and recommendations of the report. Moreover, reports may be placed on the agenda of general political debates. Third, some reports are discussed in the form of written questions by members of parliament. The content of these documents is analyzed as well. With this data, I can answer the fourth, fifth and sixth subquestions.

1.5   Outline

In chapter 2, I present the theoretical perspectives that inform this research. In chapter 3, the operationalization of the theoretical concepts presented in chapter two are described. Furthermore I will discuss how data is selected and I will discuss the methods used to collect and analyze empirical material. In chapter 4, I will describe and discuss the results of the analysis of the empirical data. Dhapter 5 is the concluding chapter, in which an answer to the research question is given. I will discuss and reflect on this thesis, and I will present recommendations on what stakeholders can do with the information presented in this thesis.

2. THEORETICAL FRAMEWORK

Agency research can be split up in three strands (James & Van Thiel, 2011). First, many researchers have attempted to define and categorize types of agencies. What are they and how do they differ from each other? Second, researchers have attempted to explain the rise of agencies. Third, researchers have attempted to analyze the consequences of their ubiquity. How do they actually work in practice? Do they live up to their promise of effectiveness and efficiency? How can their performance be evaluated?

In this chapter, I will broadly follow the structure of these three research strands. First, I will delve into the nature of ZBOs. What are they and how are they distinct from other agencies and other forms of government? Second, I will describe the two primary theoretical explanations for the existence of ZBOs and their rise to prominence since the 1980s. Some researchers see the rise of ZBOs as part of a global trend (e.g. Pollitt & Talbot, 2004). Others use an explanation grounded in economic theories of rational choice (e.g. Dunleavy, 1991; James, 2003). Third, I will discuss the consequences for the existence of ZBOs. I will make two arguments in this section. First, I will argue that assessing the performance of ZBOs is very difficult. ZBOs often carry out complex tasks that are not easily expressed in measurable outcomes. Their goals are ambiguous and multiple (Van Thiel & Leeuw, 2002; Noordegraaf & Abma, 2003). Researchers have come up with strategies of evaluation to try to account for these difficulties (e.g. Pawson & Tilley, 1997; Rossi, Lipsey, & Freeman, 1999). Second, I will argue that due to their complexity, the evaluation of ZBOs becomes susceptible to political games. Policy evaluation can be used politically, for instance to shift blame away from certain actors and to others (Bovens et al., 2008; Cohn, 1997) or to manage crises (Resodihardjo, 2006). Not only the results, but also the way the evaluation is executed becomes the highly contested subject of a political game. Stakeholders try to influence which research body executes the evaluation, who takes place in this body, the scope and the goal of the research question, its methods and its resources (Bovens et al., 2008).

In the third section, I will come back to the empirical notion from the beginning. The lack of interest in evaluations of ZBOs among Dutch politicians can be explained

from the perspective of public choice. Rationally acting politicians may only show
interest in the evaluation of ZBOs if they can benefit politically from it. Finally, I
will conclude by formulating a theoretical answer to the first two subquestions. Based
on these answers, I can formulate expectations regarding the empirical reality we will
encounter.

I will formulate expectations on what I will find in the empirical material. These
expectations are sometimes formal hypotheses in which a manipulation of the indepen-
dent variable leads to a change in the depended variable, via a certain mechanism (Van
Thiel, 2010), but sometimes they are formulated more liberally, without a mechanism or
an effect. However, all expectations are "testable propositions" (Silverman, 2011, p.53)
that can - and will - be examined with empirical material. For the sake of readability and
clarity, all are called expectations and are highlighted in the same manner throughout
the text.

2.1   What is a ZBO?

Although ZBOs have existed for decades or even centuries (Van Thiel, 2004; Bouckaert
& Peters, 2004), the creation of many ZBOs happened as a part of the administrative
reforms of the 1980s and 1990s that took place in many Western countries. These re-
forms also influenced Dutch policy choices. Under the name of New Public Management
(NPM), many central state activities were transferred to various semi-autonomous or-
ganizations. These organizations were expected to be more efficient at service delivery.
Moreover, it was expected that the transferral of tasks to these organizations would lib-
erate the central government, so that it could focus on policy making instead of policy
implementation (Hood, 1991). In the Netherlands, many semi-autonomous organiza-
tions were created in a process called *verzelfstandiging*, or 'autonomization' (Van Thiel,
2004).

There are many organizations in many different countries that are similar to
the Dutch ZBOs. These organizations are called "agencies" (Pollitt & Talbot, 2004)
or "Quangos" (Greve, Flinders, & Van Thiel, 1999). All have in common that they
carry out public tasks, and exist at the national level, "at arm's length" of the central
government (Van Thiel, 2004; Pollitt et al., 2001). By public tasks, I mean service
provision, regulation, adjudication and certification (Talbot, 2004, p. 5). Because ZBO
is an administrative term and not a theoretical concept, the term is difficult to define
more specifically. Any more specific criterion added to the broad definition above, is

contested. For instance, one could add that ZBOs are financed with state budget. This is only partly true. ZBOs are often funded with state budget, but they can have other means of funding.[1]

Because of this, it seems best to start from the assumption that ZBOs exist somewhere on a continuum of organizations that all operate at the national level, that all exist at a distance from the central government, and that carry out public tasks (Greve et al., 1999; Van Thiel, 2004). These organizations vary in the extent of their autonomy and under what control mechanisms they fall (Pollitt & Talbot, 2004). They may fall under direct ministerial control, but they may also be subjected to market regulation as the main control mechanism (Greve et al., 1999; Van Thiel, 2004). On this continuum, we find many different organizations with many different tasks, including ZBOs.

Where can we place ZBOs on this continuum? First, according to the Dutch Kaderwet ZBOs (2007), ZBOs do not fall hierarchically under a minister. This sets them apart from, for instance the Dutch contract agencies, which enjoy less autonomy from the ministry. This autonomy applies primarily to managerial tasks. In general, the minister is not responsible for managerial issues, but remains accountable for policy choices (Van Thiel, 2004). Because of this decreased ministerial responsibility, the minister can not fully control the production process of the ZBO - that is, what goes on inside the organization of the ZBO is not visible to the minister. By setting budgets and goals, by imposing performance standards that are measured with detailed indicators, the minister maintains control over the output of the ZBO. Thus, formally, ZBOs have managerial autonomy, but no policy making capacities. This is in line with one of the assumptions behind the NPM reforms, namely that agencies gain autonomy in the way they operate, but refrain from interfering with policy making. It is argued, however, that in practice, ZBOs can have considerable influence in the policy making cycle, through their bureaucratic activities (Bach, Niklasson, & Painter, 2012). They can delay implementation of undesired policies, or sustain policies that they prefer, they can influence policy choices through lobbying activities, and policy makers ask for advice from ZBOs on new policies.

In sum, ZBOs are one of many variants of semi-autonomous organizations that

---

[1] An example of this is the Dutch RDW, a ZBO concerned with registering and approving motorized vehicles and driver's licenses. They own a test track which they can rent out to private parties. In 2012, a bill concerning unfair competition between the government and private companies was adopted. The bill codified what activities were and were not allowed. The renting out of the test track by the RDW was deemed unfair competition, and abolished (See: RDW, 2014).

exist. They are fully functioning independent organizations and have large *managerial* autonomy. They do not hierarchically fall under a minister. Formally, policy making belongs to the domain of the minister, although de facto, ZBOs have some policy making capacities. Through budget and goal setting, the minister determines the output of the ZBO.

2.2   The rise of ZBOs

NPM reforms were a driving force behind the surge in the number of ZBOs in the 1980s and 1990s (Pollitt et al., 2001). NPM is an umbrella term for a "set of broadly similar administrative doctrines which dominated the bureaucratic reform agenda in many of the OECD group of countries from the late 1970s" (Hood, 1991, pp. 3-4). One of the main goals of these reforms is the creation of a leaner, more efficient government that ought to employ business-like management styles and that ought to focus on its core tasks. Connected to this is an emphasis on output - rather than on processes - which should be assessed with measurable performance indicators, goals and targets (Hood, 1991). The transfer of tasks towards (semi-)privatized organizations (agencies) and away from the central government fits the objectives of NPM. By placing a task in the hands of a separate organization, the central government loses control over what happens inside the organization, but it can more easily set budgets, goals and targets for the output of the production process. It is argued that the central government might ultimately gain more control even though the execution of a task is placed at a larger distance. Performance indicators, targets and budgets, are strong instruments for the central government to strictly control the autonomized organization (Smullen et al., 2001).

Many justifications by politicians and bureaucrats for the creation of agencies and specifically ZBOs exist (Van Thiel, 2004; James, 2003). These lay-man explanations are often a reflection of NPM ideas (James & Van Thiel, 2011), and can be placed in three categories (Overman & Van Thiel, 2015). Economical expectations are an increase in efficiency and better quality of policy implementation, service provision closer to the citizens and more possibilities for competition (which, in turn, is assumed to make service delivery better and more efficient). Political effects are a smaller government, less political concern for policy implementation so that politicians can focus on policy making, and more transparency and accountability. Organizational effects are to operate more like a business, neutrality and independence from the (unstable) political climate

to and to make expert responsible for implementation, instead of politicians. In short, agencificiation is expected to contribute to efficient, but less costly service delivery.

There are two strands of *scientific* explanations for the rise of agencies (James & Van Thiel, 2011). The first one explains the rise of agencies as the result of a global trend in which agencies are the preferred way of service delivery (e.g. Pollitt et al., 2001; Pollitt & Talbot, 2004). Based in institutionalism, it focuses on contextual factors such as traditions, acceptable practices, norms and values. The second strand takes a rationalist approach. It focuses on the benefits gained by creating an agency for politicians, bureaucrats and other parties involved (e.g. James, 2003; Dunleavy, 1991). The strands seek an answer to slightly different questions. The first strand asks: "why did similar processes happen in so many different countries?". The second strand asks: ""Why are agencies the preferred policy choice for so many individual politicians and bureaucrats?"

2.2.1 The institutionalist approach: a global trend

The main principle of institutionalism is that organizations are shaped by their environmental context (Selznick, 1996). Ideas, practices, norms and values are imprinted in the organization, creating a paradigm (Hall, 1993). Furthermore, the organizational structure itself - the decision-making procedures, formal and informal rules - shape and limit acceptable patterns of behavior for actors within that organization (Bannink & Resodihardjo, 2006). Because of these two phenomena, institutions limit and shape policy choices available to politicians and bureaucrats.

Importantly, institutions make policy change difficult. *Path dependency*, for instance, is a concept that is used to argue that the policy choices in the past determine the options currently available to politicians [2] (Kuipers, 2009). Change is not impossible, of course. Change is explained through various mechanisms. One important concept that explains institutional change, is the idea of institutional crises (Alink, Boin, & 't Hart, 2001). When the current institutional structures suffer a significant loss of legitimacy in the eyes of major political actors, change is possible or even necessary (Suchman, 1995).

The main question that researchers attempt to answer with institutionalism is why, during the 1980s and 1990s, so many agencies were created in so many different

---

[2] An often cited example is the QWERTY-keyboard. In the days of the typewriter, the layout was designed to minimize the chance that the individual hammers got stuck. Nowadays, this is no longer necessary, but hardly anybody is interested in switching to a different layout, because they have become used to the QWERTY layout.

countries. One answer is found in *isomorphism*, which uses the concept of *institutional legitimacy* (James & Van Thiel, 2011) to explain why institutions slowly start looking alike. Institutional legitimacy is "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (Suchman, 1995, p. 574). Organizations are dependent on their legitimacy for their survival. A lack of legitimacy means that the organization's reason of existence is called into question. The organization's actions are no longer considered normal, appropriate and desirable. This may ultimately lead to the organization's disintegration. One of the strategies to maintain legitimacy is to *conform* to demands from the environment (Suchman, 1995). These demands can be formally imposed standards or informal norms and ideals that already exist and are espoused or practiced by other actors and organizations in the environment. Organizations are pressured to conform to these standards. Slowly but surely, organizations start to converge on certain practices, norms and ideas about how to operate. In this case, creating agencies became the idea that many national governments copied and the norm to which many national governments adapted.

In conclusion, institutional answers to questions about the rise of agencies are concerned with how the institutional environment shapes policy outcomes. Stagnation as well as change can be explained with institutional theory. The global trend towards the use of agencies can be explained using the concept of isomorphism. The broad answer is that environmental and contextual factors create incentives for decision-makers to choose to create agencies.

2.2.2   Rationalist approaches

Rationalist explanations for the rise of agencies are grounded in economic theories of behavior in a relation between the state and agencies. Their relationship is characterized as a "cascade" (James & Van Thiel, 2011, p. 214) of principal-agent relationships. Politicians as principals delegate tasks to bureaucrats, who in their turn delegate these tasks to agencies. Principal-agent theory assumes that both the principal and the agent act rationally and in self-interest. The agent generally has an information advantage, and will use this to further his interests (Moe, 1984). This leads to two problems: *adverse selection* and *moral hazard* (Shepsle, 2006; James & Van Thiel, 2011). Adverse selection concerns hidden information about the characteristics of the agent that the principal cannot know in advance. The principal is not sure whether the agent will support

their policy choices or whether the agent will be competent enough to implement them. Moral hazard is associated with hidden behavior. The principal cannot always directly observe the agent's behavior. The agent might act differently when unmonitored, and hide their true behavior. A solution to these problems is sought in all kinds of contractual arrangements (Moe, 1984) or in oversight (Lupia & McCubbins, 1994), but information asymmetry can never be fully eliminated.

So why do politicians choose to create agencies? From a rationalist perspective, the core assumption is that all actors have self-interested goals at heart. This implies that these actors are not only concerned with making good policy, but also with strategic behavior that is aimed towards advancing their own interests. Politicians ultimately want to get (re-)elected (Dunleavy, 1991), while bureaucrats, operating under the politician, want to maximize their power, income, prestige and job security (James, 2003). One explanation emphasizes that bureaucrats have a strong influence in the decision to create agencies. Another explanation sees delegation of responsibilities as a strategy for politicians to gain an electoral advantage.

Bureau-shaping

The bureau-shaping perspective of agencification focuses on the behavior of bureaucrats (James, 2003). These bureaucrats are agents of the politicians, but principals of the agency at the same time. They too have self-interested goals at heart. Dunleavy (1991, pp. 184-149) summarizes Down's (1967) model of the general motivations that bureaucrats have: instrumental motives are to maximize their power, income, prestige and job security, while minimizing the effort they have to put in. Broader, more altruistic motivations are loyalty to their immediate coworkers, their bureau, their government or their country, commitment to their specific tasks or program, pride in doing their job well, or a desire to act in what they believe to be the public interest. Dunleavy (1991, pp. 165-167) criticizes these broader motives, calling them "excessively loose". These motives are no longer "rational" and therefore, cause a loss of explanatory power. These broad motives can explain any type of behavior and a model that incorporates these motives loses the ability to predict behavior.

Rationally acting bureaucrats have many strategies to improve the "utility" of their work - that is, the amount of power, income, prestige and job security it delivers. Examples of these strategies are, among others, promotion, workload reduction, or budget-maximization (Dunleavy, 1991, p.175). Bureau-shaping, too, is one of these strategies, and agencification can be explained as the result of these bureau-shaping

strategies (James, 2003). The core premise is that bureaucrats want to maximize the budgets available to them, and want to do 'policy work' that requires creativity, innovation and close relations to political power, as opposed to 'managerial work', which involves "routine implementation of procedures, hands-on monitoring of junior staff, and working in a large, extended hierarchy at the point of delivery, remote from political power sources" (James, 2003, p. 25).

The transfer of tasks to agencies leads to a reduction of the budget and tasks of the department previously responsible. But the tasks that are generally transferred are 'managerial tasks'. Furthermore, the agency is placed at a larger distance from the central state. The department previously responsible for these tasks will become a small and elite group, gaining a focus on policy work, and remaining closer to the central political power sources. Senior officials who expect to stay in the core department, will support and pursue agencification. This also explains the idea of separating policy making and policy implementation, where implementation is left to organizations that are placed at a distance from the central government.

James (2003) thus argues that agencification was a bureaucratic idea. Politicians may have had the broader idea of a more efficient government, but bureaucrats translated these goals into policies of agencification. Bureaucrats were confronted with budget constraints, and managed to come up with a strategy that minimized their loss of budgets and policy work.

Delegation as a credit-claiming and blame avoidance strategy

Another public choice explanation for the creation of agencies, is the concept of *blame avoidance* (Weaver, 1986; Hood, 2002; Cohn, 1997). As mentioned, the core premise of public choice is that the primary concern of all politicians is securing (re-)election (Dunleavy, 1991). The implication of this is that politicians are not only concerned with making good policy, but also with strategic behavior that is aimed towards gaining an electoral advantage. *Credit-claiming* and *blame avoidance* are two major manifestations of this type of behavior (Weaver, 1986; Hood, 2002).

Credit-claiming and blame avoidance are two sides of the same coin. According to Weaver (1986), blame avoidance is the more important strategy, due to humans' risk- and loss-aversive nature. When faced with a choice between a risky, but potentially highly beneficial option and a safe, always slightly beneficial option, people pick the safer option (risk aversion). When faced with a choice between a small but certain loss and an option that might either cause no loss or a severe loss, people tend to choose for

the latter (loss aversion).[3] In other words, people dislike taking risks if they are faced with a potential gain, but they do take these risks when faced with a potential loss. In other words still, people dislike losing more than they like winning. They prefer avoiding a loss to obtaining a benefit.[4]

There are at least two consequences for politics. First, politicians perceive that getting blamed for a negative outcome of a policy choice is worse than receiving credit for a positive outcome of a policy choice. Second, voters will remember more vividly a suffered loss than a gained benefit. As a consequence, the cost of getting blamed for a bad result is higher than the benefit of receiving credit for a good result.

As a result, most research on strategic political behavior is focused on blame avoidance (Hood, 2002). Below, I will show how blame avoidance and credit-taking strategies play a part in the creation and evaluation of ZBOs.

Delegation of responsibilities for the implementation of policy programs is an important blame-avoidance strategy (Hood, 2002; Cohn, 1997). Through delegation of implementation, a politician can remove direct responsibility for a policy program, and place it in the hands of an agent. When a policy program fails, the politician can place the blame on the agent and promise improvement. The politician himself is not immediately in danger, although he does accept responsibility for improving the situation. Cohn (1997) argues that the wave of privatizations in the United states of Reagan and the United Kingdom of Thatcher was a way for politicians to take unpopular measures without being blamed for them. He argues that in a traditional hierarchical bureaucracy, the minister is ultimately responsible for actions of his employees. Through delegation strategies such as privatization, the minister can place that responsibility on those below him.

Hood (2002) argues that the effectiveness of delegation to avoid blame depends on the relationship between voters and the relationship between the agent and the politicians. Voters and the agent on which the blame is placed may decide not to accept the blame-shifting move. Voters may decide not to believe that the blame lies with the agent, or they may simply want the politicians head for other reasons. The agent can decide not to accept the blame and employ strategies to blame either another actor entirely, or try to place the blame back on the politician.

Delegation can also be seen as a credit-taking strategy. Majone (2001) argues

---

[3] These effects have even been observed in capuchin monkeys, which implies that they are ancient, evolutionary traits, embedded deeply in our brain (see: Chen, Lakshminarayanan, & Santos, 2006)

[4] For more information, see: Tversky and Kahneman (1992)

that delegation can lead to an increase in credibility for a politician. He shows to the electorate that he is willing to renounce some of his powers if it increases the effectiveness of government. Because politicians are only in office for a limited time, they have an incentive to pursue short term policy goals. Delegation is a politician's attempt to ensure that his policy preferences are carried out, even if he is no longer there. He places tasks in the hands of an organization that is specifically created to carry out the tasks that he wants it to carry out, in the way he wants it to carry them out. This organization must necessarily be independent from the politician, otherwise his successor reneges on the policy plan. Furthermore, because the politician is only in office for a short time, he cannot be blamed for long-term negative effects of his delegation efforts. In sum, politicians gain short term credibility by showing that they are committed to long term policy goals. They renounce some of their powers and place them in an independent organization. In doing so, they can prevent their successors from taking a different course of action.

2.2.3   Conclusion

In conclusion, the rise of ZBOs is explained by a variety of concepts. A global trend that normalizes creating agencies as a form of government, combined with the interests of rational individuals, be they politicians or bureaucrats, has lead to a boom in the creation of ZBOs. For politicians, agencification was the preferred policy choice because politicians believed it helped them get (re)-elected or reach their policy goals. For bureaucrats, agencification was an instrument that could be used in bureau-shaping activities. This happened in an environment in which delegating responsibilities was a normal policy choice that was seen throughout the Western world.

2.3   Do ZBOs work? Assessing the consequences of ZBOs

In the previous two sections, I have discussed literature on how to define ZBOs and introduced the most prominent theoretical explanations for their rise to prominence since the 1980s. I have done this because an understanding of what ZBOs are and why they were created clarifies the nature of the position ZBOs have in modern government, and how politicians deal with them. Now, we turn to the third research topic: how does the creation of ZBOs affect the functioning of modern government, and how can these consequences be assessed? How can we explain political (dis-)interest in the evaluation of ZBOs.

This section starts with the notion that whether ZBOs are a 'good' form of government is difficult to assess. I will argue that this is at least partly caused by the fact that assessing and measuring public sector performance is difficult. Performance of ZBOs forms no exception. This not only has consequences for scientific research on ZBO performance, but also on the evaluation reports on individual ZBOs that are submitted to the Dutch parliament. I will then argue that such evaluation reports are political in nature. This has consequences for the political effects these reports have and for the reasons for starting an evaluation.

2.3.1 Evaluating performance: easier said than done

As we have seen, the creation of ZBOs is expected to achieve a myriad of goals, which are often connected to NPM ideas. The establishment of ZBOs is expected to have economic, political and organizational effects (Overman & Van Thiel, 2015). From an economic perspective, efficiency, effectiveness, lower costs, more customer-friendly service delivery and competitiveness are expected. Political effects are a smaller government, a separation of policy implementation from policy making and more transparency. Organizational effects are a more professional, business-like organization, autonomy leading to a stronger identity and motivation for employees, and neutral, unbiased service delivery (Pollitt et al., 2001; Van Thiel, 2004; Overman & Van Thiel, 2015). In short, the transfer of tasks to ZBOs is expected to lead to better performance. But do ZBOs live up to these expectations?

Empirical evidence that answer this question is scarce and provides mixed, or even negative results (Overman & Van Thiel, 2015; Pollitt et al., 2001). Overman and Van Thiel (2015) find a negative connection between public sector output and agencification. James (2003) asserts that the creation of executive agencies in the UK did not lead to much improvement in public sector performance, because improving public sector performance was secondary to bureau-shaping as a goal of agencification.

One reason that empirical evidence is so scarce, is that it is difficult to assess the performance of ZBOs. Part of the body of NPM ideas is that organizations can determine themselves how best to execute the tasks they are asked to carry out. This means that, for the central government, the focus is placed on output and outcomes, not on input and throughput (Heinrich, 2002). By focusing on output, the central government can leave the input and processes to the ZBO, while at the same time still being able to influence the direction of the organization by setting goals for outputs and outcomes.

The performance on these goals for output and outcomes is measured through the direct and explicit use of indicators and targets (Boyne, 2010). With indicators one can attempt to translate performance into objective and measurable results of an organization's actions. Targets are meant to set expectations on the score an organization should achieve on these indicators (Boyne, 2010). Performance is thus understood as the scores of an organization on a set of indicators for output and outcomes.

This way of measuring performance is criticized (Noordegraaf & Abma, 2003). Performance of public organizations on outputs and outcomes is difficult, if not impossible to properly measure. ZBOs often carry out complex tasks that are not easily expressed in measurable outcomes. For instance, regulatory ZBOs such as the Dutch Financial Markets Authority (AFM) should detect rule violations. Does this mean that the AFM performs well if it detects more rule violations than last year? I do not know the answer, because this number does not tell me much. It could mean that the AFM has become better at detecting rule violations, but it could also mean that it has failed in prevention strategies, leading to an increase in the total amount of rule violations. If the AFM becomes very good at prevention, it will detect less rule violations, simply because there are less to detect. This is a positive result, but it leads to a *negative effect*. It is very difficult to express the effect of prevention activities in measurable numbers on an indicator. Moreover, this example shows that tasks and goals may be mutually exclusive. Good performance on one indicator may cause a decline in another. Furthermore, the link between the activities of a ZBO and its outcomes is very difficult to make. For instance, the goals of the Dutch *Commissariaat voor de Media* are to protect the independency, pluriformity and accessibility of the media supply, thereby contributing to freedom of speech (report Evaluatie Commissariaat voor de Media 2007-2011, 2013). Its tasks are primarily regulatory in nature, it provides permits, controls if television and radio broadcasters follow the rules regarding the amount of commercials. There is no direct causal link between its tasks and its goals.

For these reasons, I expect that evaluation reports will acknowledge such difficulties in assessing performance on output and will pay attention to internal organizational processes to complement results on output and outcomes. Evaluators may either shy away altogether from evaluating ZBOs that have tasks that are difficult to measure, or they may focus much attention on the processes in that organization if its tasks are difficult to measure. Pollitt and Dan (2011) observed that many evaluation reports on NPM reforms - of which agencification is one - did not report on effects, only on processes. This is the reasoning behind the first two expectations.

Expectation 1: ZBOs that have tasks that are easy to express in measurable output indicators, are more likely to be evaluated than ZBOs of which it is difficult to measure performance.

Expectation 2: when ZBOs that have tasks that are difficult to measure, are evaluated, there is little focus on output and more focus on internal processes.

Another reason for the difficulties in measuring performance is that comparison can be difficult. ZBOs are often monopolists (Kickert, 2001), which means that their performance cannot be compared with other organizations that provide identical services. There have been attempts to compare agencies over time and across countries, in order to see whether there are effects of agencification in general (e.g. Overman & Van Thiel, 2015; Pollitt & Dan, 2011). The problem with this type of research is that it is very hard to establish causation. For instance, the article by Overman and Van Thiel (2015) compares the level of service delivery across countries. Some of these countries have created agencies in an attempt to improve service delivery, whereas other countries have not. The authors find a negative effect of agencification on service delivery: countries that did create agencies score worse than those that did not. However, concluding from this that worsened service delivery is *caused* by agencification would be a *post hoc ergo propter hoc* fallacy. There may have been many other interfering variables at play. Perhaps the countries that chose for agencification did so in order to mitigate external effects that would have caused an even worse decrease in service delivery if these agencies were not created. Just because a decrease in service delivery happened *after* a period of agencification, does not mean it was *caused* by it.

So, measuring performance would be easier if causation can be established through the comparison of the performance of the ZBO to similar organizations (benchmarking), or through comparison to a baseline measurement of how well the ZBO's tasks were executed before it was created. Therefore, I expect that ZBOs are more likely to be evaluated if a benchmark or a baseline measurement is present.

Expectation3: ZBOs are more likely to be evaluated if there is access to a benchmark or a baseline measurement with which the ZBO's performance can be compared.

A third difficulty is that policy goals have to be translated into measurable indicators, in order to measure performance. But those objectives are often multiple, complex and

diffuse, because they are the result of a political process (Weiss, 1993). During the processes of persuasion, negotiation and bargaining, many promises and deals regarding the expected effects of a policy program may be made, in the hopes of gaining support for the policy choice. As a result, a policy program has multiple, diffuse, complex and unrealistically grand objectives, to the point where it is expected to soo solve every and any problem in the world. Weiss (1993, p. 96) gives an example: "Public housing will not just provide decent living space; it will improve health, enhance marital stability, reduce crime, and lead to improved school performance". We have seen that agencies, too, are expected to reach many different goals, such as improved efficiency, service delivery closer to the people, smaller government, separation of policy making and implementation et cetera (Van Thiel, 2004; Overman & Van Thiel, 2015). This leads to two problems. First, how do we translate a goal such as "bring government closer to the people" goals into measurable indicators that we can use to evaluate a single agency? Goals such as this are hard to express in quantitative data. Second, how do we know which goals are the most important ones? Does a policy program fail if it cannot live up to all inflated expectations?

On the basis of this assumption, a fourth expectation is formulated:

Expectation 4: ZBOs are less likely to be evaluated if their goals are multiple, diffuse and complex.

In sum, it is difficult to assess whether ZBOs are a good way of government. Attempts have been made, but these provide mixed results. ZBO performance is difficult to express in measurable indicators, because some effects are obfuscated or negative and because policy objectives are complex, multiple and potentially mutually exclusive. Comparison of performance is difficult because of a lack of comparable data, and because of difficulties in establishing causation. This is expected to have influence on how the performance of ZBOs is evaluated. When performance is difficult to express in measureable indicators, evaluations are not executed or they focus on internal processes. Benchmarks and baseline measurements make it easier to measure performance, which is expected to positively influence the likelihood that a ZBO is evaluated.

2.3.2   Evaluations are political in nature

Evaluation research systematically investigates the effectiveness of social programs (Rossi et al., 1999, p. 4). Ideally, evaluation research is a rational, systematical, objective in-

vestigation of the effectiveness of social policies, that can truthfully and apolitically determine whether a social program works or not (Rossi et al., 1999; Pawson & Tilley, 1997). However, in reality, evaluation research takes place in a political context, which a policy analyst cannot ignore (Weiss, 1993; Pawson & Tilley, 1997; Bovens et al., 2008; Rossi et al., 1999).

How does politics influence evaluation research? First, the subject of the evaluation - that is, a policy or program - is the result of a political decision. Moreover, the decision to evaluate a ZBO may be political in nature. Because of this, it will be defended or attacked by stakeholders, and the evaluation report will be used in this political struggle. Thus, secondly, the evaluation report enters the political arena. Here, it plays a part in agenda-setting and problem definition. In this sense, policy evaluation is not only the end, but also the beginning of the policy cycle (Bovens et al., 2008). Third, an evaluation report inherently makes judgments. Implicitly or explicitly, it defines problems, challenges the nature and legitimacy of certain programs, their goals and their strategies, while it makes other programs more difficult to challenge by supporting them.

In the remainder of this section, I will discuss the political nature of evaluations by describing what might happen when an evaluation report enters the political arena and by discussing potential political reasons to start an evaluation. I will then briefly highlight that several actors may be able to influence the start of an evaluation of a ZBO.

Political effects of evaluation reports

There are several political effects an evaluation report can have when it enters the political arena. First, stakeholders may add more or less importance to certain goals on which the ZBO is evaluated. This influences the impact of a report in the political arena. Second, I will argue that evaluation reports make political judgments. The stronger these judgments are, the more political interest they will garner. Third, I will argue that being evaluated may be threatening to the ZBO's legitimacy.

We have already seen that an agency is expected to reach goals that are multiple, complex, mutually exclusive and that overestimate the capabilities of the agency. It is expected are that this leads to improved efficiency, separation of policy making and implementation, service delivery closer to the people, smaller government and more (Van Thiel, 2004; Overman & Van Thiel, 2015). These are the goals espoused by politicians

and used in the political decision-making process, but actors may also have private, self-interested goals. Bureau-shaping, employed by bureaucrats, is an example of such goals. Blame avoidance through delegation (Weaver, 1986; Hood, 2002) can be another one.

Given this large and diffuse amount of goals, different actors may focus on different goals. There can be disagreement on which goals are the important ones. It is even possible that certain goals of agency creation have not surfaced in the public political discourse. This has consequences for the evaluation report. Generally, a ZBO is evaluated against official, formally stated goals. Thus, when an evaluation report enters the political arena, its conclusion may be unimportant for certain stakeholders who feel that the ZBO was not tested against the goals that were important to them. The ZBO may be evaluated against less important or even completely meaningless criteria. Then, the evaluation loses its political importance and may easily be discarded or contested (Weiss, 1993). Stakeholders can always argue that a ZBO scores well on certain goals, although it is evaluated negatively on other goals. Of course, in the eyes of the stakeholder, the goals the agency does reach are the "actual" or more important goals of the agency. This gives leeway to policy makers. They can argue against the conclusions of an evaluation, or even disregard them. They can also use the evaluation's conclusions to defend a ZBO.

> *Expectation 5: if evaluation reports address only official goals, they lose relevance for some stakeholders.*

In asking the question: "how effective is this ZBO in meeting its goals?" lies an implicit assumption that reaching them is desirable. Moreover, in this question lies the assumption that the ZBO as a form of organization could potentially and realistically reach these goals. The ZBO may need some tweaks, or even a large reorganization, but in principle, it is to be expected that the ZBO as a form of organization has a good chance of reaching the goals (Weiss, 1993). Hence, the evaluation criteria and the resulting evaluation report are inherently judgmental in nature. They accept certain assumptions and reject others. I expect that evaluation reports will vary in 'seriousness' of judgments: some will be broadly positive and determine that only slight tweaks in the organization are needed to optimize the ZBO's functioning, others will be negative and propose more serious reforms. I expect that political interest rises when reports propose larger measures. However, there is a limit to this. I expect that evaluation

reports do not question the fundamental reasons behind the creation of ZBOs. Reports will not conclude that the primary tasks of the ZBO are undesirable or that the ZBO as an organizational form should be discontinued and that its tasks should be placed in, for instance, a departmental unit within the ministry.

Expectation 6: If evaluation reports make strong political judgments, political interest is heightened.

But:

Expectation 7: Evaluation reports do not question the fundamental ideas behind the creation of ZBO, lowering political interest.

Since some ZBOs are evaluated as often as they should be, whereas others are not (Eerste Kamer, 2012), the very decision to evaluate one ZBO and not the other, can be a political statement - even though all ZBOs are formally obligated to be evaluated once in five years. To be evaluated may have political consequences, regardless of the outcomes of that evaluation. A ZBO that is not being evaluated is safe from criticism, whereas a ZBO that is evaluated is suspect: there must be doubts about the functioning of this ZBO, otherwise it would not be subjected to an evaluation (Weiss, 1993). Therefore, being evaluated may be damaging to the legitimacy of the ZBO. However, the argument that the ZBO is simply being evaluated due to legal requirements might mitigate this effect. The ZBO can say that nothing is wrong and that the evaluation is a formal procedure. This argument may be less convincing if it is widely known that this procedure is easily avoided. Because ZBOs are often not evaluated, in spite of the legal requirements, it is reasonable to expect that stakeholders may still regard a ZBO that is being evaluated as suspect. They may not believe that nothing is wrong if they see other ZBOs *not* being evaluated when they are legally supposed to. This argument may also work the other way. A positive evaluation may be a legitimizing force for the ZBO. It can function as evidence for the ZBO's claims that it is doing its work well. If this argument holds, it may be the ZBO itself that wants an evaluation to happen, because it can benefit politically from it. In sum, if being evaluated can be threatening to a ZBO, that ZBO will try to prevent the evaluation from being done. It will be defensive or uncooperative. If the evaluation can be used by the ZBO in its defense, the ZBO may try to initiate an evaluation. In other words:

Expectation 8: The evaluation of a ZBO is less likely to occur when to be evaluated is threatening to a ZBO's legitimacy.

<div align="center">

Reasons for starting an evaluation

</div>

Before I discuss what reasons for starting an evaluation exist, I will need to determine *who* might be able to start an evaluation. This is important because different actors may have different reasons for initiating an evaluation. Who took the initiative for an evaluation may indicate the reasons for starting the evaluation.

In the previous sections, I have mentioned several potential initiators of evaluations. For instance, I have argued that ZBOs will try to prevent evaluations from being executed, if this may cause a threat to their legitimacy. Here, I assumed that ZBOs are in a position to influence whether and when it is evaluated. Furthermore, I have noted that there are legal requirements that obligate ministers to execute an evaluation. Thus, both ministers and ZBOs can be expected to be able to influence the start of an evaluation, but it is also possible that the start of an evaluation was determined by the requirements stipulated in article 39 of the Kaderwet ZBOs and that neither the ZBO nor the minister could influence the initiation of an evaluation. I expect that there is less political interest in reports that were started simply because of legal requirements.

Expectation 9: a ZBO can evaluated on no initiative of stakeholders, but as part of an automatic process, as prescribed by legal requirements.

Expectation 10: ZBOs can be evaluated on the initiative of their political principals.

Expectation 11: a ZBO can be evaluated on the initiative of the ZBO itself.

The establishment of a ZBO is a political decision - meaning that it is the result of bargaining and compromises between political actors. There are supporters and opponents, who have their reputation attached to this decision. People depend on the ZBO for their job or their career, because they have publicly supported the decision for that program, or because they work for or are a client of the ZBO implementing it. They have an interest in the survival or the abolishment of the ZBO, and are willing to invest time and resources to make this happen (Weiss, 1993). Evaluation reports can be used in this political struggle, both in defense and in attack of the ZBO. Thus, a first reason for starting an evaluation is that it may aid in this political struggle. The need to build or

maintain support for - or against - the ZBO, may be an important reason to initiate an evaluation. This reasoning informs the following expectation. If stakeholders expect to be able to use the conclusions of an evaluation in their political struggle, an evaluation is more likely to be initiated.

> *Expectation 12: ZBOs are more likely to be evaluated if stakeholders can expect to use the conclusions of an evaluation report to create or maintain support for or opposition against the ZBO.*

Much literature exists on how policies are changed by entrepreneurial politicians or governments (Alink et al., 2001; Goldfinch & 't Hart, 2003; Bannink & Resodihardjo, 2006, e.g.). Evaluation reports can be used to create a window of opportunity that allows these politicians to push for change. Usually, this literature is concerned with large-scale reforms, changing the way the government works. It is very difficult to achieve these large-scale reforms, because of institutional constraints and resisting stakeholders (Kuipers, 2009; Bannink & Resodihardjo, 2006). Therefore, the politician that wants change, may have to use strategies to overcome these constraints. One strategy is to invoke a sense of crisis (Alink et al., 2001), in the hopes of convincing other actors that incremental changes are no longer sufficient. Actors that do not want change, may respond with strategies to contain crises (Boin & 't Hart, 2003). One of these strategies is to install an evaluation committee that is tasked with determining what went wrong, who was responsible for it and what to do next (Resodihardjo, 2006). The politician hopes that he can buy time: the evaluation is expected to take several months, by the time the report is finished, political, public and media interest may have dwindled. Furthermore, by swiftly initiating an evaluation, they can show that they are in control of the situation. They show that they are responding to the crisis by remaining calm and determined (Resodihardjo, 2006).

Resodihardjo (2006) mentions that this strategy can backfire. When after a few months the report enters the political and public domain, it may renew attention and create new momentum for change. It can be used by policy entrepreneurs to create new momentum in the crisis-situation the report was supposed to resolve. Those that want reform, can use the report to fan the remaining smoldering embers of the crisis-situation in an attempt to rekindle the fire of their cause.

In sum, an evaluation report can be used in the political struggle for or against a policy change, especially when attention is high and the debate is fierce. The report

can be used in several ways. An entrepreneurial politician (Goldfinch & 't Hart, 2003) can use a (critical) evaluation report to change policies by adding fuel to the debate and invoke a sense of crisis. Second, a politician can attempt to maintain the current policy by using the report to buy time, diminish attention and defuse the debate.

> *Expectation 13: ZBOs are more likely to be evaluated when an incident/crisis situation has occurred.*
>
> *H15a: ZBOs are evaluated as part of a crisis-management strategy to maintain stability.*
>
> *H15b: The evaluation report is used by stakeholders to create a window of opportunity for change.*

Another reason politicians may want to initiate an evaluation is to avoid blame. Policy makers may attempt to delegate political judgments to the experts in the evaluation committee, as part of a blame avoidance strategy (Hood, 2002). When a politician has to make a choice of which it is difficult to assess the consequences, he can decide to ask experts and advisors to weigh the consequences for him. When the choice turns out to be wrong, the politician can say he was simply following the advise of experts. The policy maker 'passes the buck' and forces others to make a choice that has potential negative political consequences (Weaver, 1986).

This strategy, however, can backfire, Hood (2002) argues. First, experts may disagree with each other, forcing the politician to make the final choice. Then, the politician can still be blamed for making the wrong choice or taking the wrong advice. Second, an evaluation committee can also decide to present the choice on a sliding scale of risks. A first option may have a minor negative effect on a certain policy outcome, but a slightly positive effect on another. A second option may have a stronger negative effect, but also a stronger positive effect. Here, the politician is still forced to make the final choice: how strong a negative effect is he willing to accept in exchange for a positive effect? Third, evaluation committees can protect themselves by recommending very cautious measures that can not do much harm, or by recommending huge policy overhauls and reorganizations to solve a minor problem, so that the politician still has to determine whether solving the problem is worth the potential (political) costs.

> *Expectation 14: ZBOs are more likely to be evaluated if political actors expect to be able to place responsibility for decisions in the hands of the evaluation committee.*

Weiss (1993) points out that another reason to start an evaluation might be newness. As mentioned, some ZBOs have existed for decades, which may give them an uncontested institutional position. This is an argument from the institutionalist tradition. Ancient institutions do not change quickly, as they have become an uncontested, normal form of organization. Other ZBOs were only recently created, which may be a reason to evaluate whether it performs as expected. They are young institutions, that may not yet have solidified their position Tradition - or lack thereof - may determine whether some ZBOs are evaluated, and others not (Weiss, 1993).

Expectation 15: younger ZBOs are more likely to be evaluated than older ones.

2.4 Conclusion

Building on a description of ZBOs as a form of government and their rise to prominence since the 1980s, I have come to fifteen expectations, which can be placed in four categories. First, I have argued that measuring the performance of ZBOs is difficult and that reports may have a focus on the processes of a ZBO, lack access to benchmarks and baseline measurements and use evaluation criteria that are irrelevant for certain stakeholders. These characteristics are expected to influence whether a ZBO is evaluated regularly or not. Second, I have argued that reports are political in nature. They can be used by stakeholders in political debates and can be threatening to the ZBOs legitimacy. Furthermore, the stronger the conclusions and recommendations of the report, the higher political interest rises. On the other hand, when conclusions are irrelevant for stakeholders, political interest is lowered. Linked to this are expectations regarding political and strategic reasons to start an evaluation. Which actor initiates this evaluation may be useful in determining political reasons for starting an evaluation. ZBOs or politicians may initiate an evaluation to protect or attack certain interest. Evaluations may be initiated to prevent or cause a crisis-like situation, to avoid blame for difficult decisions. Furthermore, institutional arrangements may be a reason to start an evaluation. ZBOs that were only recently installed, may not yet have the same uncontested institutional position as older ZBOs, and may therefore be more likely to be evaluated.

Below, I list the fifteen expectations. In the next chapter, I will operationalize these expectations into measurable indicators.

On measuring performance:

1 ZBOs that have tasks that are easy to express in measurable output indicators, are more likely to be evaluated than ZBOs of which it is difficult to measure performance.

2 When ZBOs that have tasks that are difficult to measure, are evaluated, there is little focus on output and more focus on internal processes.

3 ZBOs are more likely to be evaluated if there is access to a benchmark or a baseline measurement with which the ZBO's performance can be compared.

4 ZBOs are less likely to be evaluated if goals are multiple, diffuse and complex.

On the political nature of evaluation report:

5 If evaluation reports address only official policy goals, they lose relevance for some stakeholders.

6 If evaluation reports make strong political judgments, political interest is heightened.

7 The evaluation reports do not question the fundamental ideas behind the creation of the ZBO, which lowers political interest.

8 The evaluation of a ZBO is less likely to occur when being evaluated is threatening to a ZBO's legitimacy.

On initiating an evaluation of a ZBO:

9 The decision to evaluate a ZBO is made automatically, because of legal requirements.

10 ZBOs are evaluated on the initiative of their political principals.

11 A ZBO is evaluated on the initiative of the ZBO itself.

On reasons for deciding to evaluate a ZBO:

12 ZBOs are more likely to be evaluated if stakeholders can expect to use the conclusions of an evaluation report to create or maintain support for or opposition against the ZBO.

13 ZBOs are more likely to be evaluated when an incident/crisis situation has occurred.

   (a) ZBOs are evaluated as part of a crisis-management strategy to maintain stability.

(b) The evaluation report is used by stakeholders to create a window of opportunity for change.

14 ZBOs are more likely to be evaluated if political actors expect to be able to place responsibility for decisions in the hands of the evaluation committee.

15 Younger ZBOs are more likely to be evaluated than older ZBOs.

3. METHODS

In this chapter, the research design will be explained. I will first operationalize the theoretical concepts into measurable, observable indicators. After this, I will consider data collection and case selection. The chapter ends with an elaboration of the measures taken to ensure the validity and reliability of this research project.

3.1 Operationalization of theoretical concepts

The purpose of this thesis is to analyze political interest in evaluation reports of ZBOs. In the theoretical framework, I have formulated 15 hypotheses, which I have categorized in four groups. First, the complexity of the evaluation process may influence political interest. Second, the expected effects the evaluation report has when it enters the political arena may affect political interest. Third, the initiator of the evaluation may be relevant. Last, the reasons for deciding to evaluate may cause difference in eventual political interest. In this section, I will operationalize these concepts by determining observable indicators for each concept. The following table shows the operationalization of the theoretical concepts. An explanation follows below the table.

Table 1: Operationalization of theoretical concepts

Variable (from expectations)	Indicators	Expectation on variable	*Reports*	*Debate*
Complexity in measuring performance				
Difficulty in expressing tasks in measurable indicators	1) There is direct acknowledgement in the report that output of a ZBO is difficult to express in measurable indicators. 2) Presence and availability of quantitative indicators for performance 3) Compare tasks of ZBOs with how often it has been evaluated	1) ZBOs carrying out complex tasks are less often evaluated. 2) Quantitative data is not often available and used. 3) Regulatory ZBOs have tasks that are notoriously difficult to express in output indicators, whereas executive ZBOs have tasks that are relatively easy to measure.	X	
Focus on processes	1) Direct acknowledgement of a focus on processes in report, indicators for processes used in report. 2) The topics of evaluation are the processes that lead to certain output/outcomes. 3) The conclusions and recommendations are aimed at internal processes.	ZBOs carrying out complex tasks are evaluated on processes. There is less political interest in processes, as these are managerial issues.	X	
Presence of benchmark or baseline measurement	1) Benchmark or baseline measurement is used in evaluation report. 2) Prominence of such measurements in report.	If a ZBO can be compared with other organizations or with its past performance, political impact of report rises.	X	
Multiplicity, diffuseness and complexity of goals	1) ZBO is evaluated against a variety of official goals (which may be mutually exclusive). 2) In political debate, goals are mentioned on which the ZBO was not evaluated in the report.	If goals are multiple, diffuse and complex, then an evaluation report receives less political interest.	X	X
(Expected) political effects of evaluation				
Judgmental nature of evaluation report	1) Topics of conclusions and recommendations in evaluation report. 2) Harshness of conclusions and far-reachingness of recommendations. 3) Reactions from politicians on these recommendations.	1) Topics may be minor, small tweaks to organizational processes, or major policy overhauls that require action from the ministry. Major changes lead to more political interest. Reports will not question fundamental nature of ZBOs as a form of organization. 2) Harsher judgments and stronger recommendations in reports lead to more political interest. 3) Stakeholders are more interested in large changes	X	X

Use of reports by stakeholders in political debate	1) The report is mentioned in political debate. 2) Its recommendations are discussed and studied	Reports are used by opponents and proponents alike for political purposes.		X
The nature of goals on which ZBO is evaluated	1) What are the goals against which the ZBO is evaluated? 2) Who formulated these goals? 3) Do they satisfy the expectations of stakeholders?	If evaluated only on official goals, the report loses relevance for stakeholders.	X	X
Threat of being evaluated	ZBO under evaluation is regarded as suspicious by stakeholders.	Being evaluated is damaging to legitimacy, regardless of outcomes or occasion of evaluation. The reason for the evaluation can amplify or mitigate this effect.		X
Initiative				
On whose initiative is decided for the evaluation of a ZBO?	1) An evaluation can be initiated by the political principal of the ZBO or by the ZBO itself. 2) Automatic/legal requirements may be a direct cause to start an evaluation. Neither ZBO or political principal had a choice in the matter.	1) If the initiator is the political principal of ZBO, there may be more political interest than if the ZBO itself initiated the evaluation. 2) If the direct cause is a simple legal reason, political interest is lowered.	X	X
Reasons for starting evaluation				
Age of ZBOs	Compare age with frequency of evaluations.	Young ZBOs are more often evaluated.	X	
Incidents/ crisis-situation	1) Incidents are mentioned in report. 2) Incidents are a topic of evaluation. 3) Research question refers to specific topics/issues.	ZBOs are more often evaluated in times of crisis, as part of a crisis-management strategy or as part of attempt to create an opportunity for reform.	X	X
Avoidance of difficult, unpopular decisions by politicians	1) Recommendations are controversial, which is expressed in report or during political debate after publication. 2) Recommendations show signs of evasion strategies by committee: they are formulated weakly, are presented in a list form, from which the political principal must choose, or propose draconian measures.	Evaluation happens when politicians want to place difficult decision in hands of evaluation committee.	X	X

3.1.1   Complexity in measuring performance

The basic assumption that guides expectation 1 through 4, is that political interest is low because making clear, simple, intuitive, causal relations is difficult in an evaluation report. ZBOs may carry out tasks of which the results are difficult to express in measurable indicators. This means that it is not easy to assess whether a ZBO is doing well and why it is doing well. In short, *measuring performance is complex.* This may lower political interest in evaluations in ZBOs. There are several observable indicators to this theoretical concept. First, it may be directly acknowledged in evaluation reports. A report may read: "Although these indicators do not paint the full picture, they may give a broad sense of the ZBO's performance".

Second, I expect a strong *focus on the internal, organizational processes* of the ZBO. Many pages in the report will be dedicated to how the ZBO is organized, what rules and regulations apply, and 'how things are done'. These processes may be linked to output or outcomes indicators, but a causal relation will be made only with caution. Conclusions will paint a picture of how the ZBO operates and recommendations are made to improve these internal processes. Furthermore, the research topics, the main research question and its sub-questions reflect the focus of attention.

I have discussed that one of the primary reasons to create ZBOs is to place tasks at a distance from the government, so that 'politicians can form policy, and managers can manage'. Therefore, a report that focuses on processes is of less interest to the ZBO's political principal. The principal is expected to be less interested in the managerial issues addressed in the report. In sum, a report strongly focused on processes is of less political interest than a report focused on output and outcomes.

A third observable indicator is the presence or absence of *baseline measurements and benchmarks.* If a ZBO can be easily compared with other organizations or with past performance, a judgment on output and outcomes is easier to make. This heightens the political interest in the report. As mentioned, ZBOs are often monopolists, so comparison with other organizations is not often available. Performance of ZBOs may be compared with that of international organizations, or with previous evaluations.

To measure whether multiplicity and diffuseness of the ZBO's goals and tasks has an influence on political interest, I will determine on what goals ZBOs are evaluated and whether reports that evaluate on multiple, different goals receive more attention in political debate.

3.1.2 Expected political effects

To measure the *judgmental nature* of an evaluation report, I look at the topics of research and its conclusions and recommendations. Furthermore, I attempt to estimate the gravity of the report's conclusions and recommendations. Both the content of the conclusions and recommendations and their gravity are expected to influence political interest in and attention for the report. I expect that very critical reports in which large-scale recommendations are done will receive more political interest than reports that present mild conclusions and relatively minor recommendations.

The *nature of the goals on which a ZBO is evaluated* may also influence political interest in evaluation reports. Indicators are the content of research questions and topics, a directly stated research goal or reason to evaluate. Furthermore, who determined these goals may be indicative of their nature. Were they set independently by the research committee, or did the ZBO or its political principal have influence in determining the topics of research and the goals of the evaluation?

The use of reports in parliamentary debate is measured by analyzing how the report is discussed in parliament. By comparing the types of questions that are asked with the topics presented in the evaluation, I estimate whether political debate is about the content of the report or about topics that are not (or only indirectly) related to the content of the evaluation report. For instance, the publishing of a report may be discussed on its content and to check what the minister is going to do in response to the report, or the report may be used as a reason to discuss a politically salient issue that is only slightly related to the report.

Being evaluated may be *threatening* to the legitimacy of the ZBO. Whether this is true, is influenced by the reason of starting an evaluation and by the goals the evaluation is supposed to reach.

3.1.3 Initiative

The fact that not all ZBOs are evaluated in a timely manner, must mean that a conscious decision is made to evaluate a ZBO. Not every ZBO is evaluated simply because of legal stipulations. This decision can come from the *political principal* of the ZBO or from *the ZBO itself.* Hypotheses 5 through 7 formulate these expectations. The evaluation reports mention their principal initiator and is explicit on *legal requirements for evaluation,* if there are any. Furthermore, in accompanying letters to parliament, the initiator is mentioned as well. The expectation is that there is less political interest in the

evaluation report when a ZBO is simply evaluated because of legal requirements. If the ZBO itself is the initiator, political interest will be low as well, because this means that politicians were not interested enough to initiate the evaluation themselves. Political interest will be highest when the political principal has initiated an evaluation.

3.1.4   Reasons for initiating an evaluation

The reasons for initiating an evaluation are an important variable. First, ZBOs that were established only recently are expected to be evaluated more consistently than older ZBOs. To observe this effect, the age of ZBOs is compared with how frequently they have been evaluated. Second, an incident or crisis-like situation may be a direct reason to start an evaluation. In this case, the goal of the evaluation is to find out how the incident could have happened. If an incident or crisis-like situation was the direct reason for starting an evaluation, it is mentioned in the introduction of the report, in its accompanying letter to parliament, or in both. Political interest in such reports is expected to be higher, as the incident makes the evaluation more urgent. Third, the evaluation report may function as an instrument to back up a politician's unpopular decision. If this is the case, I expect to see signs of attempts to evade being used as an instrument by the research committee. Their recommendations will be formulated in such a way that the principal cannot hold the evaluation committee responsible for making a certain decision. Recommendations will be presented on a sliding scale, so that the principal still has to choose, or they will be weakly formulated (e.g. "the committee recommends the principal or the ZBO to look into a certain option to see if it is a valid option"), so that the principal still has to make the decisions.

3.1.5   Political interest

Political interest is observable in parliamentary debate regarding the evaluation report. First, the amount of attention given to an evaluation report in parliament is an indicator. If a report is not mentioned in parliamentary debate, or placed on the agenda but not discussed, political interest is very low. If questions regarding the report or the functioning of the ZBO are asked, political interest is higher. Second, the topics of the questions regarding an evaluation report from members of parliament to the minister is an indicator with which to assess political interest in evaluation reports. If questions are directly related to the content of the report, for example to clarify a conclusion or to check what action the responsible minister will take to follow up on that conclusion, political interest is high. If a report is used as a reason to address an unrelated political

topic, interest is low.

3.2   Case Selection and description

As of August 1st, 2015, there are, according to the registry of ZBOs [1], 88 ZBOs and 21 clusters of ZBOs, totaling 109. A cluster of ZBOs consists of organizations that all carry out the same task, which requires public authority. For instance, many car repair shops have a permit to control whether cars are fit for the road. They carry out this public task, and in doing so, they have authoritative decision-making power. They are therefore part of the ZBO cluster *Erkeninghouders Algemene Periodieke Keuring* or 'licensed periodic motor vehicle inspector'. I have found no evaluation reports on clusters of ZBOs. Evaluation reports on clusters would have been important to add to the analysis, as members of these clusters can be easily compared with others. This could have helped in testing hypotheses of benchmarking and monopolists.

This leaves us with 88 ZBOs. 45 of them have been evaluated at some point in time. Of the other 43, no evaluation reports exist.[2] Of these 45, four are currently in the process of being discontinued or merged with other ZBOs. These are the *Stimuleringsfonds Culturele Mediaproducties, ZON-MW, KNAW and the Participatiefonds.* Given their imminent termination or loss of ZBO-status, they will not be part of this research. This leaves us with 41 ZBOs.

Of the *Waarderingskamer*, only a non-public evaluation report exists. I have removed this ZBO from the analysis.

This leaves 40 ZBOs that *a)* exist now and will still exist in the near future, *b)* have been evaluated and *c)* have had their evaluation report made public. Of these ZBOs, the most recent reports will be analyzed. For a list of the ZBOs, their full names, their abbreviations and a very short description of their main task, see appendix A.

Of one ZBO, namely the *Huurcommissie*, I know an evaluation was done in 2014, because that report was mentioned in the annual report of 2014. However, I could not locate this report. I have taken the previous evaluation report, from 2008. It must be noted that back then, the *Huurcommissie* did not exist in its current form. Multiple local

---

[1] zboregister.overheid.nl

[2] The LBIO was evaluated in 2002, during an evaluation of the then new *regelingen inning kinderalimentatie*. Because this report only evaluated the effects of this specific arrangement, it is very different in nature than the other evaluation reports, which focused specifically on the ZBO itself. I have therefore removed this report from the analysis. I have also decided to count the LBIO as 'not evaluated before'.

huurcommissies (that is plural) existed. A general evaluation of how they worked was executed. Later, these local committees were merged into one national *Huurcommissie*.

Six culture funds, namely the *Mondriaan Fonds, Stichting Fonds voor Cultuurparticipatie, Stichting Nederlands Fonds voor de Film, Stichting Nederlands Fonds voor Podiumkunsten, Stichting Nederlands Letterenfonds* and the *Stichting Stimuleringsfonds Creatieve Industrie* were evaluated in one report. This means there are 35 reports for 40 ZBOs.

The AFM and the DNB were evaluated separately, but a third report on how they cooperated is part of this evaluation. Thus, three reports were written on the two ZBOs. All three are used in the analysis, which gives us 36 reports.

3.3   Data collection

I have collected 36 evaluation reports. For 31 reports, an accompanying letter from the parent minister to the parliament existed. These were analyzed as well. No accompanying letter existed in the cases of the COA, CEA, culture Funds, RvA and SVB.

Where existing, I have collected and analyzed records of parliamentary debate following the publishing of the report. I have found two types of documents. First, 22 reports were placed on the agenda of general parliamentary debates (*algemene overleggen*), while 14 were not. These were the reports on the AFM and DNB (3), COA, the CCMO, the NRGD, the CEA, the Kadaster, KvK, the culture funds, the RvA, SVB, TNO-AGE and Verispect.

Second, I have found 8 records of questions asked by members of parliament to the minister, and answers by that minister. The direct reason for these questions is the report. In the cases of the CBR, NZa (*Nederlandse Zorgautoriteit*, or Dutch Healthcare Authority), CAK, RDW, Staatsbosbeheer (National Forest Service), Stimuleringsfonds voor de Pers, ZiN (Zorginstituut Nederland) and TNO-AGE. For a list of the analyzed documents per ZBO, see appendix B.

3.4   Method of analysis

Based on the theoretical expectations, a coding lay-out was created. This lay-out could not perfectly code every piece of data. Unexpected data occurred. Therefore, I maintained the possibility to create new codes for pieces of data I had not anticipated. For instance, the report on the KvK read: "A comparison with other organizations cannot be done well, given the largely unique tasks of the [Chamber of Commerce]" (p.24). The

coding scheme provided me with a label: "benchmarks or baseline measurement used", but this did not exactly fit the content of the piece of data. I therefore created the code "benchmark considered, but not applied". In sum, I have developed a hybrid coding scheme that fit the operationalizations, but also left room to add codes if necessary. For the coding scheme, see Appendix C. I have used the software *Atlas.ti* to conduct the analysis. With this program, I could and systematically order and categorize the documents. I have labeled every document on their type (report, letter to parliament, records of general parliamentary debates and records of questions and answers) and on the ZBO. This way, I could not only determine whether a certain code occurred in one piece of data, but also in whether it appeared in all texts of one type. For instance, I could analyze data from all texts on, for instance, the AFM, but I could also analyze all parliamentary records for one code or group of codes.

3.5 Validity and reliability

One caveat to this research design is that it can only say something about evaluations that have taken place. These evaluations have all overcome the hurdles that are forces that stop evaluations from being executed. Furthermore, it is possible that political considerations were obfuscated and not mentioned in political debate or in the report. The legal requirements may have functioned as a scapegoat or an excuse to finally start an evaluation. I cannot measure these obfuscated effects, since I used only publicly available documents. Every word in these documents has been carefully considered, so that their effect is exactly and predictably what the writer intended. Can I as researcher trust these documents to speak the truth? I would argue that this does not really matter. Whatever the documents say, it is their words that directly influence the (political) reality (Silverman, 2011). Parliamentary debate - or its absence - is based on these documents, so therefore, they are a valid source of data.

Another benefit is that documents are a *naturally occuring* source of data (Silverman, 2011), meaning that the way that they came to exist was not in any way influenced by interference of the researcher, as opposed to data collected from interviews. Interview data is always provoked by questions asked by the researcher. The information is colored by the questions the researcher asks, or by how well the researcher and the interviewee could get along. This benefits the reproducibility of this research project.

4. RESULTS

In this chapter, the results of the analysis are presented. The goal of this chapter is to present and interpret the empirical material. Conclusions and answers to the research question are given in chapter 5. Earlier, 15 hypotheses were placed in four groups. This chapter will follow the structure of these four groups. Before I discuss these groups, I will present some general findings on the content of the evaluation reports. These findings do not fall under a certain hypothesis, but add relevant background information on what these evaluation reports look like and how they present data, conclusions and recommendations.

4.1 Introduction

To introduce the results I will first present some general findings. I will describe the cases with data on the structure of reports, the tasks of ZBOs, the research question and the timing of the publishing of evaluation reports. These findings will provide contextual information and are useful for understanding the circumstances under which reports are written and published, while not directly providing an answer to the research question.

4.1.1 Different ZBOs, different tasks

Van Thiel and Yesilkagit (2014) categorize the different tasks semi-autonomous agencies have. I will briefly describe the main tasks of the selected ZBOs, based on these categories.

Most ZBOs of which reports were analyzed are regulatory bodies. These ZBOs supervise markets to ensure that transactions take place according to the rules. Examples are the NZa (*Nederlandse Zorgautoriteit*) and the AFM (*Autoriteit Financile Markten*).

A second category includes ZBOs tasked with accreditation, testing and certification of products and services. For instance, the RDW (*Rijksdienst Wegverkeer*) tests whether motorized vehicles are fit for the road. This category includes the BKD, CBR, CEA, CvTE, RvA, RDW, NVAO and NIWO.

Another large category of tasks are executive organizations. These ZBOs carry out tasks on behalf of the central government, such as paying out welfare benefits or handing out permits. Typical examples are the UWV (*Uitvoeringsinstituut WerknemersVerzekeringen*), or the SVB (*Sociale VerzeringsBank*).

A fourth category includes ZBOs that transfer money. They hand out subsidies and support the activities of other organizations. As such, they stimulate activities by other organizations or people. For example, The NWO (*Nederlandse Organisatie voor Wetenschappelijk Onderzoek*) aims to stimulate and promote the role of scientific knowledge in society. It does so by subsidizing scientific research. Next to the NWO, the culture funds and the SvJ (*Stimuleringsfonds voor de Journalistiek*) fall in this category.

Other tasks are registration and record keeping, advisory and education. Four ZBOs have registration and record keeping as the primary task. Three have advisory tasks. They collect, possess and publish specific information on certain policy fields, and they advice the central government or other organizations on certain issues. For the CBS (*Centraal Bureau voor de Statistiek*), the OVV (*Onderzoeksraad voor Veiligheid*) and TNO (*Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek*) information collection and consultancy is the key activity. One ZBO, namely the *Politieacademie* is an educational organization. It provides education and training for police officers.

Table 2: Tasks of ZBOs (N=40).

Primary task	Amount	ZBO
Regulation	10	AFM, DNB, BFT, CCMO, CvM, CTGB, Huurcommissies, NZa, Verispect, ZiN
Certification	8	BKD, CBR, CEA, CvTE, RDW, RvA, NVAO, NIWO
Executive	6	CAK, COA, LVNL, Staatsbosbeheer, SVB, UWV
Registration	4	Kadaster, KvK, NRGD, RvP
Subsidies	8	Culture funds (6), NWO, SvJ
Advisory	3	CBS, OVV, TNO-AGE
Education	1	Politieacademie

4.1.2 Research committee

Of the 36 evaluation reports, 18 were carried out by an independent consultation organization such as PWC or Berenschot. In 9 cases, the evaluation was carried out by

an independent committee that was not from one consultancy organization. How the members of these committees were recruited was not mentioned in the report. Generally, such committees consisted of seniors from the field or professors from universities. In the reports it was often stressed that these committees were able to operate independently from ministry, ZBO or other stakeholders. In the case of the BFT, such a committee asked a consultancy organization to perform part of a study. This sub-report was then used by the committee to assess the functioning of the BFT. The committee remained responsible for the conclusions.

In four cases, a committee from within the ministry was appointed to carry out the evaluation. In three of these cases, the ministry appointed a consultancy organization to create a partial report. For instance, the DNB and the AFM were evaluated by the ministry, but a separate report on their cooperation was written by a consultancy organization. Five reports were written by a peer-review committee, of which three were accompanied by a self-evaluation by the ZBO. These committees seem to be able to carry out their work independently. The *politieacademie* was evaluated by a ministerial inspectorate.

Table 3: Research committee (N=36).

Research Committee	Amount	on ZBOs
Consultancy agency	15	AFM (report on individual functioning), DNB (report on individual functioning), CAK, CBR, CEA, Kadaster, KvK, NIWO, NRGD, NZa, RDW, SvJ, TNO-AGE, Verispect, ZiN
Independent committee	9	BFT, CBS, CvdM, CTGB, CvTE, NWO, OVV, Staatsbosbeheer, UWV
Both independent committee and consultancy organization	1	BFT
Ministry	1	BKD
Both Ministry and consultancy organization	2	AFM and DNB (report on cooperation), LVNL
Peer review committee	5	COA, Culture Funds, NVAO, RvA , SVB
Self-evaluation	1	CCMO
Unknown	2	Huurcommissies, RvP

4.1.3   Timeliness of evaluation

Of all the 109 ZBOs and clusters of ZBOs that exist now,[1] 45 have been evaluated at some point in time. Of these, only four have not been evaluated in the past five years. In other words, four ZBOs have not been evaluated when they should have been, even though they have been evaluated at some point in time. These are: BFT, CAK, COA and BKD. The BFT was evaluated in 2009, CAK in 2008, BKD in 2004 and COA in 2002. Six ZBOs, namely AFM, DNB, Kadaster, KvK, SVB and UWV were evaluated last in 2010. At this point, it is unknown whether new evaluation reports on these ZBOs will be sent to parliament before the end of 2015. Therefore, evaluations on these six ZBOs are not are not counted as "late". The fact that these ZBOs were evaluated less than five years ago may indicate that ZBOs are evaluated regularly, if they are evaluated

---

[1] According to the website zboregister.overheid.nl

at all. After all, only a few ZBOs that were evaluated were not evaluated in the past five years. However, it might be possible that there is a large gap in years between the most current reports and the previous reports in the cycle. I is possible that recently, more effort was made to start evaluations more often.

4.1.4 Conclusion

In conclusion, the reports that were analyzed considered a variety of different types of ZBOs. Research committtees are generally stated to be independent, even though it is not always clear how members of the committees were selected. Third, most reports are less than five years old, which means that *if* a ZBO was evaluated, it was evaluated recently.

4.2   Measuring performance

The reports were analyzed on how they collected, measured and presented data, what topics they investigated and what research questions they answered.

4.2.1   Research topics

Twenty-four reports explicitly mentioned effectiveness and efficiency (*Doelmatigheid en doeltreffendheid*) in their research question. The reason for this is that in art. 39 of the Kaderwet ZBOs uses these exact phrases are used. This means that for 12 reports, a different research question existed. Two ZBOs were assessed on their compliance with international criteria, namely the NVAO and the Raad voor Accreditatie. AFM and DNB were evaluated specifically on how they cooperated. KvK and UWV were judged on how they cooperated with other organizations in the environment.

Four reports did not have a specific research question. These were the reports on the six culture funds, BKD, NWO and SVB. To illustrate this, I will shortly describe the report on the SVB. This was a visitation report, based on a self-evaluation by the SVB. This self-evaluation was structured around four themes, namely quality, price/performance, responsive action and participation, and transparency. These themes were based on the Handvest Publieke Verantwoording.[2] The main question implicitly was how the SVB thinks it performs on these themes and whether the visitation committee agrees with these assessments, but this was never stated explicitly.

---

[2] The Handvest Publieke Verantwoording is an initiative of several public organizations to improve the structures for public accountability and transparency in (semi-)public organizations. Several ZBOs have joined this initiative, including CAK, RDW, ZiN, Kadaster, SVB and Staatsbosbeheer. Members can ask for a group of peers to investigate how well they are doing on public accountability and transparency. A few reports have been sent to parliament as an evaluation report as meant in art. 39 of the Kaderwet ZBO's

Table 4: Main research questions. N=36.

Topic of research question	Amount	ZBOs
Effectivity and efficiency	24	AFM (report on individual functioning), DNB (report on individual functioning), BFT, CAK, CBS, CCMO, CEA, COA, CvdM, CvTE, Huur-commissies, Kadaster, KvK, LVNL, NIWO, NRGD, NZa, OVV, RvP, Staatsbosbeheer, SvJ, TNO-AGE, Verispect, ZiN
Compliance with international standards	2	RvA, NVAO
Cooperation between two ZBOs	2	DNB and AFM (report on cooperation), KvK
Incident	1	CBR
Development of ZBO since autonomization	1	RDW
Internal management	1	UWV
Public accountability	1	Staatsbosbeheer
Quality of education	1	Politieacademie
Relation between ministry and ZBO	1	Staatsbosbeheer
Validity and fairness of decision-making process	1	CTGB
No stand-alone research question	3	Culture funds, BKD, NWO, SVB

Sometimes, more than one research question exists. For instance, Staatsbosbeheer is evaluated on three topics: effectiveness and efficiency, public accountability and the

relation with its parent ministry.

The main question is often split up in several research topics. These are written down in subquestions, or in operationalizations of the main question.

Important research topics were the ZBOs' relation with their parent ministry and with their environment. Most reports at least mentioned the ZBO's environment and parent ministry, but ten reports very extensively analyzed the ZBO's relation with the parent ministry, and 7 investigated the ZBO's relation with the environment. The reports on BKD, COA, LVNL, NIWO, NWO, NZa, RDW, SvJ, SVB and ZiN evaluated the ZBO's relation with the ministry. The reports on BKD, NIWO, NZa, RDW, SVB, TNO-AGE and Zorginstituut Nederland in particular were explicitly evaluated on their relation with their ministry.

Operational topics were also important. While only 6 ZBOs were evaluated specifically on their organizational functioning, most reports started with a thorough discussion of the organization of the ZBO, its position and its operational management. Finances was only three times explicitly mentioned as a topic (CBS, LBIO, LVNL), but here too, most reports discussed budgets and operational costs of the ZBO in the ZBO's description.

Reports seem to be rather similar in structure and content. The reports often use the same approach to discuss the functioning of the ZBO. The main question generally reflects what the Kaderwet ZBOs prescribes: how effective and efficient is the ZBO in carrying out its tasks? This question is often answered by first describing in detail the formal tasks and position of the ZBO. Its legal basis is described, its tasks and the instruments it has to perform these tasks. The ways it cooperates and coordinates with other actors, including the ministry and other stakeholders is often analyzed as well in the first chapter. Finances as a topic sometimes gets its own chapter, sometimes it shares its chapter with other operational management topics. Another often appearing topic is 'Governance', which is understood as the system of monitoring, accountability and control: who monitors who, on what grounds. The ZBO being a ZBO has consequences for these monitoring relations, so this topic is considered in this chapter as well.

Sometimes, performance, effectiveness and efficiency get their own chapter. For instance, the LVNL report starts with the position of the LVNL and its environment, followed by a chapter on finances and operations. Then, a chapter is devoted to performance. The LVNL is compared to other (international) organizations, its productivity is evaluated against targets. The chapter evaluates the LVNL on output targets and

indicators. The CvE report has a chapter on efficiency and effectiveness, too. Some evaluation reports, such as the CCMO report and the NRGD report, are almost fully focused on organizational processes. The NRGD report's research question directly mentions effectiveness and efficiency, which shows that the research question is interpreted in different ways in different reports.

In sum, a table of contents of a typical evaluation report would look roughly as follows:

- Chapter 1: Introduction, often includes brief description of research methods.

- Chapter 2: Description of the ZBO's tasks, organizational structure and position, environment and background.

- Chapter 3 through 5: Body of the report. Findings and analysis of the functioning of the ZBO. Sometimes, one chapter titled 'Findings' is used, sometimes multiple chapters cover different topics such as finances, governance, quality of service, "effectiveness and efficiency" and internal management.

- Chapter 6: Conclusions and recommendations.

4.2.2   Research methods

33 of the 36 reports gathered data from interviews and documentation. Two of the other three reports (RvP and the culture funds) did not mention how data was collected. The report on the CCMO used a survey among employees. Eight other reports mentioned the use of surveys among clients and customers, as a supplement to interviews and document analysis. These were: AFM and DNB, BFT, Huurcommissie, NRGD, Politieacademie, ZiN.

Seven reports (CTGB, COA, Cultuurfondsen, KvK, NVAO, Raad voor Accreditatie and SVB) used a visitation by a committee of peers as the primary method. Seven reports employed a self-evaluation (BKD, COA, CCMO, NVAO, NWO, SVB, Verispect). COA, NVAO and SVB did a self-evaluation prior to such a peer-review. The reports on BKD and NWO were written by an external commmittee that primarily based itself on the self-evaluation. For the report on Verispect, the self-evaluation played a somewhat smaller role, as other methods such as interviews were used as well. Thus, only the report on the CCMO consisted only of a self-evaluation.

In total, seven reports mention an attempt to use benchmarks and baseline measurements as a way to gain insight in the performance of the ZBO. The reports on

the LVNL, the Huurcommissie and the Commissariaat voor de Media explicitly and expansively use benchmarks. 13 reports made references to previous evaluation reports although the thoroughness of the comparison with these varied. For instance, the report on the RDW dedicated an entire chapter to the recommendations of the previous evaluation, and discussed how well these were implemented. Other reports used them less thoroughly.

Some reports mention the reason for not using benchmarks or baseline measurements. In these cases, the reason is always that comparable organizations did not exist, given the unique tasks and position of the ZBO. The report on the CvTE, for instance, mentions that a benchmark with two other ZBOs was attempted. It turned out that no other organizations that were comparable in both tasks and organizational structure existed, which caused the results to be not insightful (CvTE report, p. 17).

The reports on the Commissariaat voor de Media, the Huurcommissie, the OVV, the Politieacademie, and on TNO-AGE mentioned the use of a case study to complement data gathered from interviews, documents or surveys.

Table 5: Methods of evaluation

Method used	Amount	ZBO
Documents and interviews	33	AFM, DNB, BFT, BKD, CAK, CBR, CBS, CEA, COA, CvM, CTGB, CvTE, Huurcommissies, Kadaster, KvK, LVNL, NIWO, NRGD, NVAO, NWO, NZa, OVV, Politieacademie, RvA, RDW, Staatsbosbeheer, SvJ, SVB, TNO-AGE, UWV, Verispect, ZiN
Surveys among clients	9	AFM, DNB, CCMO, Huurcommissies, KvK, NRGD, Politieacademie, ZiN.
Case-study	5	CvdM, Huurcommissies, OVV, Politieacademie, TNO-AGE
Benchmarks and Baseline measurements	7	CvTE, Huurcommissies, Kadaster, RDW, SVB, SvJ, ZiN
Peer-review	6	CTGB, Cultural funds, KvK, NVAO, RvA, SVB
Self-evaluation	7	BKD, COA, CCMO, NVAO, NWO, SVB, Verispect

4.2.3 Multiplicity of goals

Broad categories of tasks of ZBOs are monitoring and regulation, service or product delivery, registry keeping, information collection and advising people, businesses, governments. Some ZBOs can perform these tasks for commercial purposes. ZBOs are evaluated on their ability to perform the tasks the legislator created them for. It is certainly true that some ZBOs have a very broad purpose, such as the *Commissariaat voor de Media* which ultimate goals are the protection of the independency, pluriformity and accessibility of the media supply, thereby contributing to freedom of speech (report Evaluatie Commissariaat voor de Media 2007-2011, 2013, p. 5). These broader goals (outcomes) of the legislator are of course very difficult to evaluate. Whether the efforts of the ZBO contribute to such broader goals is therefore a question that is not often answered. However, the performance of ZBOs on their tasks are evaluated. There does not seem to be much variance in the multiplicity, diffuseness and complexity between goals of ZBOs. Most have difficult to reach broad outcome goals set by the legislator, but measurable performance goals on output of tasks.

4.2.4 Conclusion

In sum, evaluation reports on ZBOs are relatively similar in several respects. Applied research methods are very often interviews and a study of documentation. In a few cases, surveys among employees or clients or an attempt to use benchmarks or baseline measurements are done. Results from such data play a secondary role. The core method is almost always qualitative interviews and documentation analysis.

The research topics show similarities as well. Effectiveness and efficiency was often mentioned in the research question, although the operationalization could vary. Other important topics were the ZBO's relation to the environment and specifically to the ministry. Much attention for internal processes was found, which is striking, because one would expect that the ministry, which is the principal of the ZBO, is more interested in input and output - value for money, so to say. An evaluation of the internal organization is of interest to the ZBO itself, surely.

In the theoretical framework, I expressed the expectation that ZBOs with more diffuse and complex functioning and goals are less likely to be evaluated. There is no evidence to support this. The link between "difficult to evaluate because of not easily measured tasks or diffuse and complex goals" and "likeliness to be evaluated regularly" cannot be made. ZBOs of all kinds are evaluated; some have regulatory tasks, such as

the AFM or the BFT, but some have executive tasks, such as the UWV.

So while the link between likeliness to be evaluated regularly and the nature of tasks and cannot be made, some important conclusions can be drawn. First, ZBOs are evaluated on many different topics, but the starting point is often the same: the effectiveness and efficiency of a ZBO. Furthermore, there is much attention for processes. The internal organization of the ZBO is of interest; transparency of decision-making procedures and an open relation with the ministry and the environment are topics of relatively large importance. Third, quantitative data plays a secondary role. The research methods were very often exclusively qualitative: interviews and a study of documentation.

4.3   Initiative

The client of the evaluation is almost always the ministry under which the ZBO falls. In 24 of the reports, the ministry was the client. Eight reports were carried out at the request of the ZBO. These were the SVB, the RvA, the LVNL, the Huurcommissies, the CTGB, the COA and the CEA. The report on the CCMO too, was initiated by the CCMO itself, but it was carried out by itself as well. Four reports never clearly stated who the client was: Politieacademie, Staatsbosbeheer, RvP and ZiN. The NVAO was evaluated to check if it was compliant with standards from an international umbrella organization, the European Association for Quality Assurance (ENQA). The evaluation was initiated by ENQA, which appointed the research panel.

Of the five cases (COA, CEA, culture funds, RvA and SVB) that were not sent to parliament, three were initiated by the ZBO itself and only one was initiated by the ministry. In the report of the RvA, no initiator was mentioned. Furthermore, four were peer review reports. The report on the CEA was done by a consultancy organization.

Table 6: Initiator of the evaluation report (N=36)

Client	Amount	ZBOs
Parent ministry	24	AFM, DNB, BFT, BKD, CAK, CBR, CBS, Commissariaat voor de Media, CvTE, Cultuurfondsen, Kadaster, KvK, LVNL, NIWO, NRGD, NVAO, NWO, NZa, OVV, RDW, Stimuleringsfonds voor de pers, TNO-AGE, UWV, Verispect
ZBO itself	7	SVB, Raad voor Accreditatie, Huurcommissie, CTGB, COA, CEA, CCMO.
Unclear	4	Politieacademie, Raad voor Plantenrassen, Zorginstituut Nederland
International umbrella organization	1	NVAO

4.3.1 Evaluation on the initiative of the ZBO itself

Seven ZBOs were evaluated on their own initiative. These were the CCMO, the COA, the CEA, the CTGB, the Huurcommissie, the Raad voor Plantenrassen and the SVB. SVB, COA, and CCMO chose for a self-evaluation. Both the self-evaluations of the SVB and the COA were complemented by a peer review committee. The CCMO was not scrutinized by a peer review committee. The CTGB and the RvA requested an international committee to carry out a visitation. The reports made no mention of a self-evaluation prior to these. The CEA was evaluated by an external consultancy organization, while it was made clear that the CEA itself initiated this evaluation. For the Huurcommissies and the RvP it was unclear who carried out the evaluation.

When the ZBO itself initiates an evaluation, the report is often not sent to parliament by the minister, although the reports are made public. Here, the CCMO, the Huurcommissie and the RvP are the exception. The reports were sent to parliament, with the accompanying letters all mentioning the obligation to evaluate of article 39 of the Kaderwet ZBOs. Of all the reports, six were not sent to parliament. These were: CEA, COA, CTGB, culture funds, RvA and SVB. Four of these were initiated by the ZBO itself. The two others were the Culture funds and the RvA. These two are both visitation reports. Four of the six that were not sent to parliament, are peer review reports, namely the COA, CTGB, Culture funds and SVB.

In sum, evaluations initiated by ZBOs were generally self-evaluations and visitation reports, and were generally not sent to parliament. Only when the initiating ZBO mentions legal requirements as a reason to initiate the evaluation, the report was sent to parliament. There seem to be no other differences between the reports that explain why one report was sent to parliament to fulfill legal requirements, and others were not.

4.3.2   Conclusion

In conclusion, a great majority of the reports were initiated by the ministry. The other reports were initiated by the ZBOs themselves. Most of the reports that mentioned legal requirements such as art. 39 of the Kaderwet ZBOs were initiated by the ministry, although there are a few exceptions. The fact that so few reports that were initiated by the ZBO itself were sent to parliament is an indication that political interest in these reports is lower than when reports were initiated by the ministry. In the section on political effects, the connection between political interest, legal requirements and initiative is discussed in more detail.

4.4   Reasons for starting an evaluation

There were several direct reasons to start an evaluation. The most common were legal requirements. Simply the fact that a law prescribed ZBOs to be evaluated spurred the evaluation. Other common reasons for an evaluation were a request by members of parliament, and incidents that occurred. I will first describe data on these three reasons. After this, I will describe a few other reasons that spurred an evaluation report to be done. These reasons were not predicted or hypothesized beforehand. Of course, reports are not limited to one reason. In the last subsection, I will describe what reasons often co-occurred.

4.4.1   Article 39 of the Kaderwet ZBOs and other legal requirements

Most ZBOs are evaluated because of legal requirements. These requirements are explicitly mentioned in the introduction of the report and in the accompanying letter to parliament. Article 39 of the Kaderwet ZBO is not the only legal basis on which ZBOs are evaluated. For instance, the AFM and the DNB were evaluated based on article 1:44 of the *Wet Financieel Toezicht*, a law on financial control. The Kadaster was formally evaluated based on article 32 of the *Organisatiewet Kadaster*. Such articles have the same content as article 39 of the Kaderwet ZBOs, prescribing a periodical evaluation

of the *effectiveness and efficiency* (*'doelmatigheid'* and *'doeltreffendheid'*) of the functioning of the ZBO. Both articles were canceled in 2013, presumably because they had become redundant when the Kaderwet ZBOs was made applicable to these ZBOs. There are more instances in which this has occurred. The AFM and the DNB were evaluated on the basis of art. 1:44 in the *Wet Financieel Toezicht*. This article, too, has now expired and is replaced by the article in the Kaderwet. It is reasonable to expect that new evaluations will be executed on the basis of art. 39 of the Kaderwet ZBOs as well.

In total, twenty-two reports mentioned legal requirements as the primary reason to start an evaluation. Of these, 12 mentioned article 39 of the Kaderwet ZBOs. 11 mentioned similar articles in other laws. Twenty-three reports mentioned reasons other than these legal requirements. What were they?

4.4.2 Requested by members of parliament

A direct reason for evaluation was a ministerial promise to parliament to execute an evaluation. Such a promise was made in five cases, namely the CAK, KvK, NRGD, NVAO and Staatsbosbeheer. Sometimes, the reason members of parliament requested such an evaluation, seemed to be part of a protocol. A new bill is proposed and a member of parliament asks that it is evaluated after a few years, because they want to know whether it had reached the proposed purpose. However, for example in the case of Staatsbosbeheer, a few members of parliament requested an evaluation because they were unsatisfied with the way Staasbosbeheer worked and had many questions and irritations. In other words, because certain incidents had come to the attention of members of parliament.

4.4.3 Incidents

The BKD, CAK, CBR and UWV were directly evaluated as a reaction to incidents that had occurred. In the reports of AFM, DNB, NZa, incidents were mentioned, but it was made explicit that the report was not concerned with these incidents. For instance, the reports and the minsterial letter to parliament on the AFM and DNB said that these evaluation reports "should be explicitly viewed as regular and periodical" (my translation, JvdS) evaluation reports, executed due to the legal requirements that AFM and DNB be evaluated on their effectiveness and efficiency. The report on Staatsbosbeheer lies somewhere in between; although incidents had occurred and an evaluation was requested because of these, Staatsbosbeheer was also evaluated due to legal requirements.

Often, a separate report is concerned with investigating the causes and conse-
quences of incidents. An example is the regular, periodical report for the Politieacademie
and a coexisting report on specific incidents and problems at the Politieacademie, known
as the Vogelzang report.[3] The regular report did not refer to the Vogelzang report.

4.4.4   Other reasons for starting an evaluation

Two ZBOs, the COA and the SVB were evaluated in order to adhere to the *Hand-
vest Publiek Verantwoorden*. This is a 'movement' (for lack of a better word) of ZBOs
and other agency-like organizations that publicly announce their commitment to trans-
parency and accountability. Organizations that are part of this movement are expected
to let a visitation committee judge their organization on how transparent and account-
able 'open to the public' they are. In the cases of the COA and the SVB, no other
evaluation reports exist. Staatsbosbeheer or RDW have been evaluated on the criteria
used in the Handvest, but they have also been evaluated by other means.

Four ZBOs (KvK, NVAO, NZa and BFT) were evaluated specifically to evaluate
certain changes in laws or policies. These changes resulted in a new position or new tasks
for the ZBO. Such evaluation reports were primarily focused on assessing whether the
consequences of the changes were as expected, and whether unexpected consequences
could be resolved. CEA was never evaluated before, which was explicitly mentioned as
a reason to start an evaluation. Two ZBOS were evaluated in order to gain or maintain
membership of an international umbrella organization. These are the NVAO and the
Raad voor Accreditatie.

4.4.5   Initiative and reasons to evaluate

There are some links to be made between the initiator and the primary reasons to
evaluate. Furthermore, there are some reasons to evaluate that often appeared together
in one case. As discussed in the section on initiative, most ZBOs were initiated by the
parent ministry. In this section, it was shown that most reports were started because of
legal requirements. Most reports that were initiated by the parent ministry were sent
to parliament for the reason of legal requirements. Furthermore, most reports that were

---

[3] The Vogelzang report was a report published in 2013 and discussed internal managerial problems.
There was disagreement and distrust between the works council and the board of directors and there
was frustration on the direction of the organization. These frustrations received attention from the
media and from parliament. The Vogelzang committee was requested to make proposals on how to
solve these issues. One proposal was to change the board of directors.

started due to legal requirements, were executed by independent committees. However, a few exceptions exist.

First, two reports mentioned legal requirements, but were initiated by the ZBO itself (Huurcommissies and LVNL). Second, not all reports that were sent to parliament on the basis of legal requirements were executed by independent committees. One was executed by the ministry (BKD), one by the ZBO itself (CCMO). The report on the CEA was not sent to parliament, even though it was executed by an independent consultancy organization. The report on the culture funds was initiated by the ministry, but not sent to parliament. Third, as mentioned in the section on initiative, three peer review reports were sent to parliament in order to adhere to legal requirements (CCMO, the Huurcommissie and the RvP).

These exception indicate that reports do not necessarily have to adhere to strict requirements before they are sent to parliament. There is freedom in which reports may be sent to parliament. Although most are initiated by the ministry and are executed by independent organization, some are initiated by the ZBO itself. Furthermore, a few peer review reports were sent to parliament with legal requirements as the reason. On the other hand, only one case that was initiated by the ministry was *not* sent to parliament. This was the case of the culture funds.

4.4.6   Conclusion

I hypothesized that three reasons might be important in the decision to evaluate. First, younger ZBOs may be evaluated more often. Second ZBOs might be evaluated so that politicians can use them as an argument when making difficult decisions. Third, incidents could be a reason.

In this section, I have discussed the mentioned reasons to start an evaluation. Most often, legal requirements were the primary reason. Other reasons to start an evaluation were requests by members of parliament and the occurrence of incidents in the near past. Minor reasons were public accountability as prescribed in the *Handvest Publieke Verantwoording* and major policy overhauls affecting the position and tasks of ZBOs.

With these reasons alone, I can not provide an answer to the hypothesis that politicians start evaluations to avoid politically controversial decisions. I will need to describe where political interest lies first. Therefore, I will postpone an answer to this hypothesis to the conclusion of the next section. For the other two hypothesis, namely that age and incidents are a reason to evaluate, I can provide an answer. However, for

both hypotheses I will refer to political interest in the reasons.

No connection was found between age of ZBOs and being evaluated. The reason for evaluation was never that a ZBO was new. Sometimes, new tasks or policies were a reason for evaluation. Most ZBOs that were evaluated, were created between 1992 and 2014. A few, such as the DNB (1814), were older. The distribution of age of ZBOs that were evaluated does not seem to differ from the distribution of age of all ZBOs

So age in and of itself does not seem to be a reason for ZBOs to be evaluated. There are, however, some indications for a connection between *political interest* in the age of ZBOs and whether they are evaluated, although the the evidence is scarce. The interest in the case of ZiN seems to be motivated by its age. The ZBO was only recently created, and politicians seemed to be interested if its creation had the expected effects.

Incidents do seem to be an important reason to start an evaluation. I come to this conclusion because incidents were sometimes mentioned as a direct reason to evaluate, but mainly because political interest in incidents is found to be high. When incidents had occurred, reports and the ZBO were more often discussed and scrutinized.

4.5   Political effects of evaluation reports

In this section, I will first discuss the topics of recommendations and conclusions in the reports. Then, I will discuss how politicians generally react to these reports.

4.5.1   On the reports' recommendations and evaluative nature

A large range of topics for recommendations and judgments is presented in reports. In seven reports, no formal recommendations were given. These were the CTGB, KvK, CvdM, Politieacademie, RvP, RvA, SVB. I will name several topics that are relevant to this research project. I will address the most often occurring topics, but also the missing topics.

First, the communication and coordination with the environment seems to be an important topic that the research committees want to address. In ten cases, there was much attention for the relationship between the ZBO and its ministry. Furthermore, in fourteen cases, there was attention for the relations between the ZBO and organizations in its environment. This is striking, because the research topics and questions did not often explicitly address the relationship between the ZBO and its surroundings, including the ministry. For instance, ZiN was evaluated on effectiveness and efficiency. Its six core tasks each got their own chapter, but the first chapter described the tasks of the ZBO

and its relation to the environment. However, the recommendations are sorted on four major topics. One on ICT, one on organizational management, one on transparency and coordination with the environment and one on the coordination with the parent ministry. This shows that often, the role of ZBOs in relation to their environment and specifically the ministry is important, needs improvement and is potentially unexpectedly an issue. It must be noted that sometimes, the relation with the ministry and the environment *are* mentioned in the research question. For example the report on the RDW specifically mentions the relation with the ministry in the research question, and the report on the KvK operationalizes 'effectiveness' and 'efficiency' partly into the KvK's relation with its parent ministry.

Related to this are topics such as transparency and public accountability. Much of attention is given to these topics. The two reports that were done because of the Handvest Publieke Verantwoording naturally paid much attention to transparency, but also the OVV or the NVAO reports, for instance, carry many recommendations to improve the transparency of decision-making processes and procedures. Here, transparency towards those affected by the decision is meant, but transparency can also mean openness to the environment. Staatsbosbeheer, for example is recommended to open up to the public so that it can get feedback on its services and products from multiple stakeholders in society. All in all, ten reports advised explicitly on issues of transparency.

Ten reports mentioned costs. These were Staatsbosbeheer, RDW, Nza, Kadaster, BKD, BFT, AFM and DNB, CBR, LVNL and ZiN. However, the financial position is often merely described, while explicit recommendations to cut costs or to raise budgets were not often presented. It seems that researchers are careful while discussing costs, and take political realities into account. There are generally no specific guidelines on whether to cut costs or raise budgets, and on how to do this. For instance, the report on the RDW discusses thoroughly how the ZBO should manage its budgets and its reserves, but it accepts the budgets as they are. Furthermore, costs are often used as a measure for performance. If the ZBO stays within its budgets, it is doing well, but if budgets are exceeded, something should be done about it. This is interesting, because members of parliament are very much interested in costs, as we shall see in section 4.5.2.

Some topics of recommendations were scarce. ICT, for instance, was very much neglected. Only the reports on the CBR, ZiN and the NZa considered ICT. This is striking, considering recent attention for large ICT projects on which government is accused of having no control over.[4] Only the report of Staatsbosbeheer evaluated the organization of the ZBO as a form of government. Here, alternatives were presented. The conclusion was that no alternative was better. Even the most critical reports did not question the reason for existence of the ZBO as a ZBO. Whether it would be better to turn the ZBO into an agency, for example, was not discussed.

Table 7: Topics of recommendations (selection)

Topic of recommendation	Amount	ZBOs
Relation ZBO and environment	14	BFT, BKD, CAK, CEA, COA, Cultuurfondsen,CvTE,Kadaster,LVNL, NWO, NZa, Staatsbosbeheer, SvJ, ZiN
Relation ZBO and ministry	10	BKD, CBS, LVNL, NWO, NZa, RDW, Staatsbosbeheer, TNO-AGE, UWV, ZiN
Transparency and accountability	10	CvTE, CTGB, Kadaster, NIWO, NVAO, OVV, RDW, RvA, SvJ, ZiN
Costs	10	Staatsbosbeheer, RDW, Nza, Kadaster, BKD, BFT, AFM and DNB, CBR, LVNL and ZiN
ICT	3	CBR, NZa and ZiN
ZBO as a form of government	1	Staatsbosbeheer
No formal recommendations	7	CTGB, KvK, CvdM, Politieacademie, RvP, RvA, SVB

4.5.2 Political interest in reports

Political interest was measured by analyzing how the ministry responded to the conclusions of the reports in its accompanying letter to parliament. Where available, parliamentary debate was analyzed on the types of topics that politicians were eager to debate or ask questions about.

---

[4] See: *Parlementair onderzoek naar ICT-projecten bij de overheid*, parliamentary document 33 326, nr. 5, https://zoek.officielebekendmakingen.nl/kst-33326-5.html, retrieved 18 October 2015

Ministerial letters to parliament

As mentioned before, thirty of the thirty-six reports were accompanied with a letter from the ministry when they were sent to parliament. Generally, these letters briefly stated the reason for the evaluation, its main goal and who executed the research. Sometimes, a relatively large summary of the main conclusions and recommendations was given, but certainly as often, no summary was given.

Only on three occasions, namely on the report of the Huurcommissies, the KvK and the RvP, no direct response on the conclusions in the report was given. In the case of the RvP, the letter consisted of a simple note to offer the report to the parliament. The letter accompanying the report of the Huurcommissies was larger, and included a short summary of the main conclusions, but gave no political statement on these conclusions. These changes may have already mitigated the issues put forward by the report. The same was observed in the case of the KvK. Issues that were found by the report were not commented on, just mentioned in order to present a solution, which was a large overhaul of the organization of the KvK.

In all the other cases, at least one - but often more than one - political statement was made. The size and 'seriousness' of such statements varied. For instance, the letter accompanying the NIWO report stated very briefly: "I will take the issues to which attention was drawn into consideration in my communication with the NIWO" (kamerbrief NIWO, my translation, JvdS). On the other hand, the letter accompanying the report on the BFT was fourteen pages long, summarized every recommendation and gave a reaction on all. The other reports varied between these extremes.

These political statements were generally positive and compliant. Conclusions were heard, ministers agreed with them and recommendations were implemented. Only in the cases of the BFT, the Kadaster, the NWO and the RDW political statements were made that were *not* in accordance with the conclusions and recommendations of the evaluation report. An example of such a statement is found in the letter on the report of the RDW: "[...][T]he [...] mentioned recommendations will be acted upon, excepting one recommendation on the efficiency. Acting upon this recommendation would, as [the research committtee] itself indicated in the report, lead to high administrative expenses. This is not in line with the policy of this government" (letter to parliament RDW, p.2, my translation, JvdS).

Sometimes, ministers drift off-topic: In the cases of CBS, Kadaster, NRGD and Verispect, ministers took actions that were neither recommended nor advised against

by the evaluation committee. In the case of the CBR, the minister mentioned incidents that had not been mentioned in the evaluation report. These incidents were that the Works Council of the CBR suspected fraud in the organization. The minister mentions these suspicions in her letter to parliament while such incidents were never mentioned in the evaluation report.

Discussion of reports in AOs

In twenty-two cases, the ministerial letters were placed on the agenda of general debates in parliament (*algemene overleggen*, or AOs). On these agendas, other documents such as letters and reports were placed as well. For fourteen reports, no placement on any agenda was found. These were the reports of the AFM and DNB (3), COA, the CCMO, the NRGD, the CEA, the Kadaster, KvK, the culture funds, the RvA, SVB, TNO-AGE and Verispect. Of these reports, the ones on the CCMO, NRGD, CEA, Culture funds, RvA, TNO-AGE, Verispect and OVV are from 2014 or 2015. It is possible that these reports will be placed on an agenda in the future. For TNO-AGE, questions regarding the report were asked, which might make it more likely that it will be placed on an agenda of an AO in the near future. On the other hand, no ministerial letter accompanied the reports on the culture funds, RvA and CEA, which may make it less likely they will ever be placed on the agenda of an AO.

I have analyzed these AOs on the following criteria: 1) whether the report is mentioned in the debate, 2) whether the report is discussed (e.g. questions on its content are asked), 3) whether the ZBO is mentioned, 4) whether its tasks are discussed and 5) whether the ZBOs functioning and performance on these tasks is assessed. The results are summarized in the following table. Below, a more detailed explanation follows.

Table 8: Attention given to reports and ZBOs in AOs (N=22)

Criterion	Amount	ZBOs
The ZBO is mentioned	19	NVAO, CAK, CBR, CTGB, BFT, LVNL, CvdM, NWO, Huurcommissies, NZa, CvTE, ZiN, Politieacademie, OVV, UWV, SvJ, RDW, NIWO, Staatsbosbeheer
The report is directly mentioned	16	NVAO, CAK, CBR, CTGB, BFT, CvdM, NWO, Huurcommissies, NZa, CvTE, ZiN, OVV, UWV, SvJ, NIWO, Staatsbosbeheer
The report is discussed	13	CAK, CBR, BFT, CTGB, CvdM, NWO, NZa, CvTE, ZiN, OVV, UWV, NIWO, Staatsbosbeheer
Tasks the ZBO implement are discussed	12	NVAO, CBR, CAK, CTGB, BFT, CvdM, NWO, NZa, CvTE, ZiN, UWV, SvJ, Staatsbosbeheer
The ZBO's performance is discussed and assessed	8	CAK, CBR, CTGB, BFT, NZa, ZiN, UWV, Staatsbosbeheer
Neither ZBO nor report is mentioned at all	3	BKD, CBS, RvP

Of the twenty-two reports that were placed on the agenda, most were at least mentioned. Only in three AOs, neither report nor ZBO was mentioned. These were the BKD, the CBS and the Raad voor Plantenrassen. Three more reports remained unmentioned, while the ZBO it concerned *was* mentioned: RDW, Politieacademie and LVNL.

However, many reports were only briefly referred to. For instance, the NVAO and its report were referenced, but they were not thoroughly discussed. The NVAO was mentioned as the executor of government policy, but whether they did this well or badly was not discussed. The evaluation report was mentioned twice, but only briefly. No questions or remarks on the content of the report were asked or given.

Another example of low interest is the report on the LVNL. In a debate on aviation in general, the report was placed on the agenda, but never mentioned in the debate. The LVNL itself was mentioned, but only as the writer of other reports, or as the executor of tasks. How well the LVNL performs these tasks was not discussed in this debate. In

both cases, the agenda of the debate was on a much broader topic than just the report on the ZBO or the ZBO's performance. The LNVL was mentioned in a debate on aviation in general and the NVAO was mentioned in a debate on the entire accreditation system.

Sometimes, reports were mentioned simply as a starting point for making general party-political statements on a broader issue. For instance, during the parliamentary debate on Staatsbosbeheer, one member criticized the validity and reliability of the report and used that critique to discredit the validity of the recommendation not to change the ZBO-status of Staatsbosbeheer. The member of parliament then proceeded to make the case that the ZBO should become a governmental service.

On the other hand, some reports were discussed in a debate specifically on the ZBO. These were CBR, NZa, Staatsbosbeheer and ZiN. A good example is the debate on ZiN. In this debate, the functioning of the ZBO and the report were extensively discussed. Members of parliament discussed the functioning of the ZBO, how this could be improved, how the ZBO related to other organizations. The recommendations and conclusions of the report were discussed to clarify them, to ask the minister what he was going to do to implement them, or to discuss them on their merits. Moreover, the research design was discussed. Questions such as what the survey looked like, who were interviewed, whether they were representative and why certain cases were chosen were asked. The cases of the CBR, NZa and Staatsbosbeheer were likely discussed due to incidents that had occurred. It is unclear why ZiN received its own AO, but it might be because it is a very new and young ZBO.

In the debates on CTGB, Staatsbosbeheer, Stimuleringsfonds voor de Pers, UWV, and especially ZiN, the research design of the report was criticized. For instance, the conclusions of the report on Staatsbosbeheer were considered invalid by one member of parliament, because the list of interviewees was too small and consisted of primarily managers, which invalidated the representativeness of the report. In the AO on the CTGB, different reports were compared. Their conclusions were compared with eachother, and questions regarding them were asked.

The ZBOs that were discussed most thoroughly were BFT, CAK, CBR, CTGB, NZa, Staatsbosbeheer, UWV and ZiN. In these cases, members of parliament asked questions on the content of the report and on the tasks and performance of the ZBOs. For instance, in the debate on the UWV, the ZBO's future existence was discussed. The report was mentioned, and critically reviewed. Some of its recommendations were met with criticism from members of parliament. Some members of parliament were critical of the report's short length; they felt it was not comprehensive enough.

The extent to which a report is discussed does not seem to be determined by characteristics of the report. For instance, the reports on the CTGB, Staatsbosbeheer and UWV were among the most thoroughly discussed reports, while one was a peer review report (CTGB), one was a relatively small report on a limited topic (UWV) and one was a very large report on many topics (Staatbosbeheer).

There does seem to be a link between incidents surrounding the ZBO and the amount of attention given to the ZBO and the report. There are three arguments that support this notion. First, in five of the eight reports that were most thoroughly discussed, incidents were mentioned. Second, only one case (AFM and DNB) in which incidents were mentioned, was not discussed thoroughly in an AO. The other cases were all among the most extensively discussed cases in AOs. Third, a document with questions from members of parliament existed in five of the six cases in which incidents were mentioned, either in the report or in communication among politicians. Here too, the AFM and DNB are the exception. In three cases, questions were asked, but not mention of an incident was found. These were the cases of the TNO, the Stimuleringsfonds voor de Pers and the RDW. It must be noted that these connections are made on the basis of only a few cases, so the link may be based on coincidence. Robust conclusions on the connection between incidents and political attention cannot be drawn from these observations.

In sum, most reports are placed on the agenda of AOs. Often, these AOs are not focused entirely on the ZBO, but on a larger topic such as a newly introduced bill or a policy sector. When placed on the agenda, the report or the ZBO it concerns are likely to get mentioned. However, most reports and ZBOs are only briefly mentioned. In a few cases, the report and the performance of the ZBO were discussed more thoroughly, with questions regarding the organization and functioning of the ZBO and witch questions regarding the recommendations and conclusions of the report, and even its research design.

Questions from members of parliament

In eight cases, records of questions that were asked by members of parliament regarding the report were found. These cases are the CAK, CBR NZa, RDW, Staatsbosbeheer, Stimuleringsfonds voor de Pers, TNO-AGE and ZiN. In general, several dozens of questions were asked on a single ZBO. The topics of these questions were very diverse. Questions were asked on broader, general policy goals, but also on the organizational

processes of the ZBO and their management of human resources. Some questions were not or only slightly related to the ZBO and the report. I will mention the most common topics and give a few examples of these topics.

First, there was much parliamentary interest in ministerial control of the ZBO. Questions on how the ministry and the ZBO had divided their tasks, how the activities of the ZBO were monitored and how communication went between the ZBO and the minster were often asked. The case of the NZa provides an example. Many questions were asked regarding the relation between the ministry and the ZBO. The reason for this interest seemed to vary. Sometimes, members seemed to want the minister to control the ZBO more tightly, but sometimes they wanted to know whether the ZBO could still operate independently from the minster, without too much political interference.

In relation to this, issues of transparency were often questioned. Members of parliament were particularly interested in transparency towards themselves. They wanted to stay informed and often requested for certain documents to be made public. They seemed rather suspicious of newly installed boards of directors, and asked whether these people were independent and competent. For instance, a question regarding the CBR read: "What is the basis of your trust in the new members of the board of directors and the board of supervisors?" (List of questions CBR, my translation, JvdS).

A third topic of interest was costs. It seems reasonable to expect that parliament is interested in costs. It is almost always relevant to ask about money and it does not require much specific knowledge to ask: "how much did it cost?" or "how do you control the costs of this program?". For instance, in the discussion of the report on Staatsbosbeheer, one recommendation was to improve the position of the Board of Supervision. One question was: "... Will this be associated with higher costs of the Board of Supervision?" (List of questions Staatsbosbeheer, p. 4, my translation, JvdS).

Fourth, at least some attention was given to managerial issues. Questions such as how does the ZBO work, how is it structured, what are its strategies, how are human resources managed, how does the ZBO relate to its environment and to other ZBOs were asked.

Last, many questions were not directly related to the conclusions of the report. For instance, one question regarding the Stimuleringsfonds voor de Pers was about whether there was budget to stimulate magazines aimed at ethnic or cultural minorities, and whether these magazines were published in other languages than the Dutch language. This was not an issue that was given any attention in the report or by other members of parliament. Other unrelated questions were asked on incidents that were

not evaluated in the report and other reports on the same ZBO.

In conclusion, political interest was very varied. Questions were asked and statements were made on many different topics. A few topics that received much attention were costs, transparency and the ministerial relation with the ZBO. It seems that political interest was primarily motivated by the most salient topics of the report, combined with the most salient topics of the day. There are some signs of political use of reports: questions on unrelated topics were asked, but the data are simply not sufficient to come to decisive conclusions.

4.5.3   Conclusion

Overall, the most important political effect of evaluation reports was that their recommendations and conclusions were often *said to be* accepted and implemented without much debate. This can be considered a large effect. However, when considering the debates that reports sparked and the use of reports in political debate, the impact is small. Although most reports were placed on the agenda of a general political debate, a substantial amount of reports never made such an agenda. Moreover, most of the reports that *were* placed on an agenda received little attention. It seems that a politician's attention is piqued by topics that are already more salient. It seems that a report is discussed when its subject is already of interest. A report does not seem to be able to steer political interest much. In this sense, reports have limited political impact.

Earlier, I formulated four expectations regarding the political effects of reports. First, I expected that politicians might have different interests regarding the ZBO than the report discusses and that they are more interested in reports if the correspond with their own political agenda. Second, I expected that the more drastic measures a report proposed, the more interest a report would garner. Third, however, reports were not expected to question the fundamental ideas behind the creation of the ZBO. A fourth expected political effect was that evaluation reports were threatening to a ZBO's legitimacy. Furthermore, I can now answer whether politicians evaluate ZBOs in order to avoid being blamed for unpopular decisions.

Whether ZBOs are evaluated to avoid difficult decisions is hard to answer, but there are signs that the idea can be rejected. Reports are almost never controversial. In fact, recommendations are very often said to be accepted and implemented without much debate. Members of parliament do not often discuss the conclusions of the reports, and more often spur the minister to implement its recommendations. This is a sign that

ministers do not make controversial decisions based on the reports. In some cases it did happen that politicians were critical of the conclusions and recommendations of the report, and attacked them. The minister then had to defend the report. This happened not often, though. Therefore, whether the hypothesis should be rejected is unsure, but there is stronger evidence against it than for it.

Connected to this is the expectation that politicians are more interested in the report if it corresponds to their own goals and that they use the report in order to advance their own interests. This has been observed in a few cases, the most clear example being the case of Staatsbosbeheer. Some politicians mention the report as an excuse to start a party-ideological argument. To this end, they were observed to question the reliability and validity of the report. Here, too, evidence is scarce, and therefore not enough to provide a conclusive answer.

As mentioned, political interest does not seem to be influenced much by evaluation reports. While it is hard to assess from the formal, distant language how big a recommendation really is, reports did not seem to propose large policy overhauls, or large reorganizations to be done. Reports did accept the politically determined parameters within which they had room to do recommendations. Therefore, it was hard to measure whether stronger recommendations sparked more political interest. I can say, however, that the fundamental ideas behind the creation of a ZBO were not questioned in the reports. Because reports did not question the fundamental *reason of existence* of ZBOs, and because political interest is generally cold to luke-warm, I can say that evaluations do not seem to be very threatening for ZBOs.

4.6   Conclusion

In one sentence, the main conclusion to be drawn from this chapter is that reports were rather similar in content and received cold to luke-warm political attention.

I should start by re-emphasizing that most ZBOs were not evaluated, that many evaluation reports were not placed on agendas of general political debates, and that attention was low even when they were placed on such an agenda. Thus, when I talk about political impact of reports, or political reactions towards reports, I talk about a select minority that jumped through four hoops: 1) being evaluated, 2) being presented to parliament, accompanied by a ministerial reaction, 3) being placed on the agenda of a parliamentary debate, and 4) being mentioned and discussed in that debate. Here, the smallest hoop is being evaluated in the first place. The second smallest hoop is getting

attention from members of parliament in general debates, while getting a ministerial reaction is to be expected. Being placed on an agenda is also fairly common.

On the flip-side, most reports were sent to parliament with a ministerial letter. Moreover, most conclusions were accepted and most recommendations were said to be implemented. In this sense, evaluation reports have a very real impact on how a ZBO operates.

The content of the reports did not vary much. The methods were qualitative in nature, while a minority of reports used benchmarks, baseline measurements or surveys as secondary data. In fact, even the structure of the reports was fairly similar. Most evaluations were done in the context of the legal requirement to evaluate on "effectiveness and efficiency", but the interpretation of this was fairly broad, with many reports extensively discussing the relation of the ZBO with the environment, organizational processes and human resource management, public accountability and transparency, while more narrow interpretations such as costs, 'value for money' and customer satisfaction were secondary research topics.

The most common reasons to start an evaluation were, in order, legal requirements, a request by members of parliament and incidents. Most often, the ministry made the decision to evaluate, but ZBOs could start an evaluation too. In a few cases, reports initiated by ZBOs were sent to parliament as well.

I will revisit the expectations once more. First, I expected that difficulty in measuring the performance of ZBOs would influence political interest in reports. Second, I expected that several actor could take initiative, which could influence the political effects of an evaluation report. Third, I expected that the reasons to start an evaluation could be political in nature. Fourth, I expected that there were several characteristics of reports that influenced the political effects and impact of evaluation reports.

Table 9: Expectations and brief summary of results

Expectation	*Results*
Complexity in measuring performance	
Difficulty in expressing tasks in measurable indicators	"Effectiveness and efficiency" was interpreted in different ways. Research topics varied. Often, the ZBO's relation with its environment and parent ministry, costs, and operational management were included in the evaluation. A few reports directly acknowledged that performance was difficult to measure.
Focus on processes	Reports focus on numerous topics, including internal processes, the ZBOs relation to the environment and with the parent ministry
Presence of benchmarks and baseline measurements	Benchmarks and baseline measurements were not often used. The most often used methods were interviews and studies of documents.
Multiplicity, diffuseness and complexity of goals	The link between "difficult to evaluate because goals are diffuse, complex and not easily measured" and "likeliness to be evaluated regularly" could not be made.
Initiative	
ZBOs can be evaluated on the initiative of ministries, ZBOs or automatically	ZBOs were most often evaluated on the initiative of the ministry. Often, legal requirements were the reason to start an evaluation. ZBOs sometimes initiated an evaluation. Such evaluations were often done by peer-review commmittees, instead of independent committees appointed by the ministry.
Reasons for starting evaluations	
Age of ZBOs	No link between the age of ZBOs and the likeliness of being evaluated was found.
Incidents	Incidents were found to increase the likelihood that ZBOs are evaluated. Often, a separate report on the incidents existed next to a periodical report that was written to meet the legal requirements.
Avoidance of being blamed for difficult, unpopular decisions by politicians	Because conclusions were generally accepted and recommendations were said to be implemented swiftly and because political interest was most often low, it is unlikely that politicians start evaluations to avoid blame for difficult decisions.
Political effects	

Reports are judgmental in nature	Reports offered a wide variety of conclusions and recommendations on a wide variety of topics. The report's conclusions were often accepted, and their recommendations said to be implemented swiftly. Reports did not question the fundamental reasons to create ZBOs.
Use of reports by stake-holders in political debate	Reports were often sent to parliament and placed on the agendas of general political debates. In these debates, they were not often extensively discussed. In a few cases, written questions regarding the report and the ZBO were asked. These questions addressed a wide variety of topics and included questions that were not related to the report or the ZBO. This implies that the publishing of a report can be used as am excuse to talk about what interests members of parliament.
The nature of goals on which ZBO is evaluated	Because political interest in reports was generally cold to luke-warm, we can conclude that political interest did not correspond to the goals on which the ZBO was evaluated. It was not often made clear in the reports who formulated the goals of evaluation, although some reports made explicit that they used evaluation criteria formulated by the client, which was often the parent ministry.
Threat of being evaluated	Because political interest in reports was often cold to luke-warm, because reports did not often propose large reorganizations or policy overhauls and because ZBOs themselves initiated evaluations, it can be expected that evaluation reports are not very threatening to the ZBO's legitimacy.

5. CONCLUSION

In this final chapter, an answer to the research question is formulated and discussed I will end with a few recommendations addressed at the writers of the evaluation reports, ZBOs, ministries and members of parliament.

5.1 Answering the research question

I will answer the main research question, which was *when, how and why Dutch politicians decide to evaluate ZBOs?* by following the structure of the three elements - when, how and why. The when-question matches the sub-question *what does the decision-making process leading up to the evaluation look like?* The how-question mirrors the sub-question *what are the characteristics of an evaluation report of a ZBO?* The why-question corresponds to the sub-questions *what are functions of an evaluation for politicians?*, *what is the impact of evaluation reports on political debates?* and *to what extent does a decision to evaluate reflect strategic/political motives for the use of evaluations?* In doing so, I will answer the second, third, fourth, fifth and sixth subquestions.

5.1.1 When?

Most importantly, Dutch politicians do *not* decide to evaluate the functioning of ZBOs. Most often, namely, ZBOs are not evaluated. But in the cases in which an evaluation did happen, what triggered it? Most often, it is the Dutch Kaderwet ZBOs, which reads in art. 39: "the minister sends a report for the purpose of evaluating the effectiveness and efficiency of the functioning of a ZBO" (my translation, JvdS). Other direct reasons for evaluation are a request by members of parliament, the initiative of the ZBO itself and the occurrence of incidents, but these triggers are scarce.

   The minister does initiate most evaluations, and is the client of most. As the client, the minister can instruct the evaluation committee on what to evaluate and what questions to answer. In some cases, the ZBO initiates the evaluation. Most of these evaluations are not sent to parliament, although some are.

5.1.2   How?

The evaluation committee is generally an independent research committee of experts or peers. Often, an external consultancy organization is requested to carry out the evaluation. Sometimes, the ZBO performs a self-evaluation, which is then judged by an independent committee. Sometimes, civil servants from the parent ministry carry out the evaluation, but this is not common practice.

Almost always interviews and a study of documentation are used as the primary method of collecting data. The research question often has the literal words 'effectiveness' and 'efficiency' in the phrasing. Their definition, however, is interpreted broadly. Thus, many topics are studied. Research is not limited to a narrow comparison of output and input, or 'value for money', but covers a variety of topics such as the ZBO's relation to its environment and parent ministry, human resource management, transparency of decision-making procedures and public accountability.

5.1.3   Why?

Can we discern any strategic and political motives from how these reports are then discussed in parliament? Most reports are offered to parliament by the parent minister with a letter in which the report's conclusions and recommendations are briefly summarized. Importantly, these conclusions are very often accepted without comment by the minister, who promises that the recommendations will be implemented swiftly. The majority of the reports are then placed on an agenda of general parliamentary debate, although a substantial amount of reports is not. In these debates, reports are rarely discussed in a meaningful way. The general debates are often on a larger topic such as a policy area or a new bill, and not specifically on the ZBO. Interest in reports and in the ZBOs they discuss rises when incidents play a part.

Members of parliament have the opportunity to ask the ministry questions in response to the report. They use this opportunity rarely. When they do, the topics of their interest varies enormously and is not necessarily related to the content of the report. Thus, it seems that political interest can only be influenced in a small way by the content of the reports and is more dependent on other factors, such as, for instance, media attention. Reports that consider incidents do catch more attention from politicians. When incidents surround the ZBO, but theses incidents are not discussed in the report, the ZBO still receives more attention.

When it comes to political, strategic use of evaluation reports, not much evidence

was found. In a few cases, reports were used to advance party-ideological ideas, but this was rare. Members of parliament did sometimes have questions on the design and the validity and reliability of the report, but then again, they had questions on every conceivable topic. So no real evidence of strategies of blame avoidance was found.

In sum, it is hard to assess why ministers initiate an evaluation when looking at political debate, partly because there is not much debate to look at. Of course, instrumental functions of an evaluation report should not be overlooked. Politicians may simply want to know how a ZBO is performing, or they may want the reports as an instrument to control the ZBO. Furthermore, it seems that sometimes, ZBOs are evaluated simply because the minister is required to send an evaluation report to parliament. For instance, the report on the CCMO was an already existing report that was sent to parliament, just to fulfill the requirements of the Kaderwet ZBOs. In such cases, the minister has no political reasons to start an evaluation.

I can, however, say something about the political impact of evaluation reports. Simply because their conclusions and recommendations are often accepted implemented, the impact of evaluation reports is large. Furthermore, reports are more often discussed when incidents have occurred in the near past. Political saliency of the ZBO in general, is a predictor for political interest in the report. Incidents seem to be most strongly connected to political interest in reports, although many reports are written without apparent incidents present.

5.2   Discussion

This thesis started with an observation: there is little political interest in the functioning of many ZBOs, while these represent such a large part of government activity. The expectation was that politicians and perhaps the ZBO itself can use evaluation reports to influence this political interest and that this might be the reason that some ZBOs are evaluated and some are not. While more evidence is found in support of the observation - reports generally received little attention in parliament - little evidence was found in support or in opposition of political use of reports by politicians. It seems that political interest is not easily influenced by the content of the report. When political attention is already high, it focuses on a wide variety of topics that may or may not be related to the ZBO or the report. In the following paragraphs I will first present one important consequence of this conclusion. Then, I will point out a few characteristics of the research design and the theoretical assumptions of this thesis that may have influenced

its conclusions. I will also present a few suggestions for further research.

First, while political interest was generally found to be low, *when* political interest was present, it was very much directed at ministerial control. Both the reports and the parliaments are very interested in the relationship between the ZBO and the parent ministry. How the ministry controls the ZBO and transparency in that relationship are large topics of interest. Moreover, many reports place much attention on organizational processes. Politicians show interest in these topics, too. This seems paradoxical, because when the parent ministry wants to know about the 'effectiveness' and 'efficiency' in the functioning of the ZBO - which it consciously placed at a distance from itself in order to encourage independence of the ZBO and in order to focus on policy making - one would expect it is primarily interested in 'value for money', that is, input and output. However, in practice, ministries and members of parliament are interested in internal processes, transparency, accountability and control. Here, it seems that performance is no longer the primary concern - fairness and legitimacy is. "Our way of doing things is reasonable, rational and fair, so therefore we are doing well", is what politicians and ZBOs want to be able to say. This is essentially what evaluation committees try to evaluate, rather than effectiveness and efficiency. This largely defeats the purpose of giving public tasks to independent organizations such as ZBOs. It lends credibility to the argument by Smullen et al. (2001) and Bach et al. (2012) that giving governmental tasks to separate organizations does not necessarily lead to a loss of control by the central government and to more autonomy for the separate organization, and may in fact increase ministerial control over such organizations.

Second, this thesis used the assumption from public choice theory that politicians - and bureaucrats - act in strategic and self-interested ways. Put bluntly, when presented with a choice, they choose the option that makes it most likely that they get (re-)elected. Because of this perspective, this thesis focused on *political* and *strategic* motives that could explain why there is so little political interest in evaluation reports. It did not focus on instrumental and functional reasons to evaluate. Simply the idea that politicians want to know about the performance of a ZBO, may be perfectly valid. However, the theoretical perspective underlying this thesis prohibits such expectations. Therefore, complementary research could be done on the instrumental impact of evaluation reports. How do ZBOs work on the implementation of the report's recommendations? What changes in policies and work processes can be seen after a report is published? Such questions were not asked in this thesis, as the focused was placed on political and strategic use of reports.

last, only public documentation and records were used to gain insight in potential political and strategic use of reports. This data may have obfuscated 'real' political strategies and considerations, because it is very difficult to unearth these 'true' motives with data that contains only carefully chosen words and phrases. With these documents, a full picture on some expectation could not be painted. For instance, little data could be gathered on the expectation that evaluations were threatening to the legitimacy of ZBOs. From interviews with writers of reports, political influence on the research project could have been measured more accurately, and through this, political and strategic use could have been analyzed more thoroughly. Furthermore, combined with interviews with managers of the ZBOs, the political impact of such reports on ZBOs could have been investigated. These interviewees may have been a little more open about political and strategic use of report. A research project such as this may answer questions on the influence of reports on the legitimacy of ZBOs. Critical evaluation reports may be dangerous to the integrity of a ZBO, while positive evaluations may be contributing to the ZBOs legitimacy and existence. In sum, based on the findings in this thesis, a few cases might be selected for a qualitative case study to thoroughly explore reasons to evaluate, political and strategic motives and the mechanisms in which they all might influence political interest in evaluation reports.

5.3 Recommendations

I conclude with a few recommendations, addressed to several actors invested in evaluation reports on the ZBOs' functioning. These recommendations are based on the previous section on the conclusion and discussion.

To ministries

1. *Use art. 39 of the Kaderwet ZBOS as an opportunity to evaluate the performance of ZBOs.* Some reports were sent to parliament with no apparent reason other than to fulfill legal requirements. If such a requirement exists, it seems a worthy excuse to initiate an evaluation to gain insight in how the ZBO is functioning. Then, the ministry can determine what exactly it wants the evaluation committee to evaluate. This has functional benefits, because the ministry can make sure that the committee researches what the ministry wants to know. It also may have strategic and political benefits, as the ministry can influence the timing and content of the report when it initiates the evaluation.

2. Related to this, *develop a robust way to evaluate output of the activities of the ZBO* and create benchmarks and baseline measurements. Use the ZBO's results on such outputs to decide whether this contributes to the ministry's policy goals. Currently, the ministry assigns a research committee to evaluate the ZBO on 'effectiveness' and 'efficiency'. Differences were observed in how research committees interpreted and operationalized these terms. With standardized evaluation criteria, a more uniform way of evaluating performance can be developed. This benefits the value and relevance of the reports and it increases the comparability of data on the functioning of ZBOs.

3. In the reports, there is much attention to the reasonableness and fairness of the organizational processes of the ZBO. While this may be important, the report's focus should be effectiveness and efficiency. Therefore, it is recommended that in these reports, more focus is placed on input and output. A different report regarding the reasonableness and fairness of the operations of the ZBO can be written. Then, the ministry can assess ZBOs more thoroughly on performance *and* on reasonableness, fairness and other questions of legitimacy. Alternatively, questions of legitimacy can be left to boards of supervisors and other means of control. The reports can still focus on effectiveness and efficiency.

To members of parliament

1. *Ask for reports on all ZBOs and demand that art. 39 is adhered to.* Members may not be very interested in these reports, but they should at least ask that the reports are made, so that they at least have the opportunity to read them if they want to gain more insight in the functioning of ZBOs.

2. It is in the interest of parliament that a standardized way of evaluation is developed, because it may make it easier for members of parliament to read reports and to compare performance. So, members of parliament might *demand that evaluations are executed and presented in a structured and standardized way.* This makes reports more accessible, and it makes ZBOs more easy to compare. This leads to an improved mechanism for parliamentary control.

3. Members of parliament show that they are interested in a variety of topics other than effectiveness and efficiency. Specifically, they seem to be interested in costs and in accountability and control. Both topics are often addressed in evaluation reports, but this may distract from the performance of the ZBO. Therefore, it

is recommended that either more attention is paid to what the report concludes specifically on the performance of the ZBO, or that extra time is taken to discuss accountability and performance in separate debates.

4. When reports focus too much on the internal processes of the ZBO, members of parliament can *ask why so little attention is given to output and outcomes* and demand that they receive information on how the ZBO executes its tasks. Then, they can discuss whether these tasks contribute to desirable policy goals.

To research committees

1. Evaluation committees can *recommend that quantitative data such as benchmarks and baseline measurements are developed*. Evaluation reports have real political impact, because its conclusions are often agreed with, and its recommendations are often promised to be implemented swiftly. This power can be used by the research committee, for instance to recommend that quantitative data on the performance should be created and collected, so that new evaluation committees may use that data to create a better assessment of the ZBO's functioning and performance.

2. Research committees can *ask for specific clarification on how the client expects 'effectiveness' and 'efficiency' to be operationalized*. Currently, these concepts are interpreted in a variety of ways, which could decrease the value and political impact of the report.

To ZBOs

1. *If they have the chance, ZBOs should initiate an evaluation themselves*. This way, ZBOs can have more influence in the evaluation process. This thesis has shown that in a few cases, reports that were initiated by the ZBO itself were sent to parliament in order to meet the requirements of art. 39 of the Kaderwet ZBOs. Ideally, ZBOs could anticipate this and have a report ready when that five-year term approaches, so that it can offer that report to the ministry when the time comes. The benefit is that the ZBO has more influence in determining what is evaluated exactly and how it is evaluated.

BIBLIOGRAPHY

Alink, F., Boin, A., & 't Hart, P. (2001). Institutional crises and reforms in policy sectors: the case of asylum policy in europe. *Journal of European Public Policy*, *8*(2), 286–306.

Bach, T., Niklasson, B., & Painter, M. (2012). The role of agencies in policy-making. *Policy and Society*, *31*(3), 183–193.

Bannink, D. & Resodihardjo, S. (2006). The myths of reform. In *Reform in europe: breaking the barriers in government*. Aldershot, UK: Ashgate Publishing.

Boin, A. & 't Hart, P. (2003). Public leadership in times of crisis: mission impossible? *Public Administration Review*, *63*(5), 544–553.

Bouckaert, G. & Peters, G. B. (2004). What is available and what is missing in the study of quangos? In C. Pollitt & C. Talbot (Eds.), *Unbundled government: a critical analysis of the global trend to agencies, quangos and contractualization* (pp. 22–49). London: Routledge.

Bovens, M., 't Hart, P., & Kuipers, S. (2008). The politics of policy evaluation. In R. E. Goodin, M. Moran, & M. Rein (Eds.), *The oxford handbook of public policy* (pp. 319–330). Oxford: Oxford University Press.

Bovens, M., 't Hart, P., & Van Twist, M. (2007). *Openbaar bestuur: beleid, organisatie en politiek*. Alphen aan den Rijn: Kluwer.

Boyne, G. (2010). Performance management: does it work? In R. Walker, G. Boyne, & G. Brewer (Eds.), *Public management and performance: research directions* (pp. 207–226). Cambridge: Cambridge University Press.

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How basic are behavioral biases? evidence from capuchin monkey trading behavior. *Journal of Political Economy*, *114*(3), 517–537.

Cohn, D. (1997). Creating crises and avoiding blame the politics of public service reform and the new public management in great britain and the united states. *Administration & Society*, *29*(5), 584–616.

De Kruijf, J. A. M. (2011). *Autonomy and control of public bodies: legal and economic autonomy and parliamentary control on service provision by zbos* (Dissertation, University of Twente, Enschede).

Dunleavy, P. (1991). *Democracy, bureaucracy and public choice: economic approaches in political science.* Hemel Hempstead: Harvester Wheatsheaf.

Eerste Kamer. (2012). *Verbinding verbroken?* C, B. Kamerstukken vergaderjaar 2012-2013.

Goldfinch, S. & 't Hart, P. (2003). Leadership and institutional reform: engineering macroeconomic policy change in australia. *Governance: An International Journal of Policy, Administration, and Institutions*, *16*(2), 235–270.

Greve, C., Flinders, M., & Van Thiel, S. (1999). Quangos - what's in a name? defining quangos from a comparative perspective. *Governance*, *12*(2), 129–146.

Hall, P. A. (1993). Policy paradigms, social learning, and the state: the case of economic policymaking in britain. *Comparative politics*, *25*(3), 275–296.

Heinrich, C. J. (2002). Outcomes–based performance management in the public sector: implications for government accountability and effectiveness. *Public Administration Review*, *62*(6), 712–725.

Hood, C. (1991). A public management for all seasons? *Public administration*, *69*(1), 3–19.

Hood, C. (2002). The risk game and the blame game. *Government and opposition*, *37*(01), 15–37.

James, O. (2003). *The executive agency revolution in whitehall: public interest versus bureau-shaping perspectives.* Basingstoke: Palgrave MacMillan.

James, O. & Van Thiel, S. (2011). Structural devolution and agencification. In T. Christensen & P. Laegreid (Eds.), *Ashgate research companion to new public management* (pp. 209–222).

Kickert, W. J. M. (2001). Public management of hybrid organizations: governance of quasi-autonomous executive agencies. *International Public Management Journal*, *4*(2), 135–150.

Kuipers, S. (2009). Paths of the past or the road ahead? path dependency and policy change in two continental european welfare states. *Journal of Comparative Policy Analysis*, *11*(2), 163–180.

Lupia, A. & McCubbins, M. D. (1994). Learning from oversight: fire alarms and police patrols reconstructed. *Journal of Law, Economics, & Organization*, 96–125.

Majone, G. (2001). Two logics of delegation agency and fiduciary relations in eu governance. *European Union Politics*, *2*(1), 103–122.

Moe, T. M. (1984). The new economics of organization. *American journal of political science*, *28*(4), 739–777.

Noordegraaf, M. & Abma, T. (2003). Management by measurement? public management practices amidst ambiguity. *Public Administration*, *81*(4), 853–871.

Overman, S. & Van Thiel, S. (2015). Agencification and public sector performance: a systematic comparison in 20 countries. *Public Management Review*, (ahead-of-print), 1–25.

Pawson, R. & Tilley, N. (1997). *Realistic evaluation.* London: SAGE.

Pollit, C. & Dan, S. (2013). Searching for impacts in performance-oriented management reform: a review of the european literature. *Public Performance and Management Review*, *37*(1), 7–32.

Pollitt, C., Bathgate, K., Caulfield, J., Smullen, A., & Talbot, C. (2001). Agency fever? analysis of an international policy fashion. *Journal of Comparative Policy Analysis: Research and Practice*, *3*(3), 271–290.

Pollitt, C. & Dan, S. (2011). *The impacts of the new public management in europe: a meta-analysis.* COCOPS working paper No. 3: www.cocops.eu. Retrieved from www.cocops.eu

Pollitt, C. & Talbot, C. (2004). *Unbundled government: a critical analysis of the global trend to agencies, quangos and contractualisation.* London: Routledge.

RDW. (2014). *Jaarverslag 2014.*

Resodihardjo, S. (2006). Wielding a double-edged sword: the use of inquiries at times of crisis. *Journal of Contingencies and Crisis Management*, *14*(4), 199–206.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (1999). *Evaluation: a systematic approach.* United States: SAGE Publications.

Selznick, P. (1996). Institutionalism "old" and" new". *Administrative Science Quarterly*, *41*(2), 270–277.

Shepsle, K. A. (2006). Rational choice institutionalism. In R. Rhodes, S. Binder, & B. Rockman (Eds.), *The oxford handbook of political institutions* (pp. 23–38). Oxford: Oxford University Press.

Silverman, D. (2011). *Interpreting qualitative data: a guide to the principles of qualitative research.* London: Sage Publications.

Smullen, A., Van Thiel, S., & Pollitt, C. (2001). Agentschappen en de verzelfstandigingsparadox. *Beleid en Maatschappij*, *28*(4), 190–201.

Suchman, M. C. (1995). Managing legitimacy: strategic and institutional approaches. *Academy of Management Review, 20*(3), 571–610.

Talbot, C. (2004). The agency idea: sometimes old, sometimes new, sometimes borrowed, sometimes untrue. In C. Pollitt & C. Talbot (Eds.), *Unbundled government: a critical analysis of the global trend to agencies, quangos and contractualization* (pp. 3–21). London: Routledge.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and uncertainty, 5*(4), 297–323.

Van Thiel, S. (2004). Quangos in dutch government. In C. Pollitt & C. Talbot (Eds.), *Unbundled government: a critical analysis of the global trend to agencies, quangos and contractualisation*. London: Routledge.

Van Thiel, S. (2004). Trends in the public sector. why politicians prefer quasi-autonomous organizations. *Journal of Theoretical Politics, 16*(2), 175–201.

Van Thiel, S. (2010). *Bestuurskundig onderzoek: een methodologische inleiding*. Bussum: Coutinho.

Van Thiel, S. & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public Performance & Management Review, 25*(3), 267–281.

Van Thiel, S. & Yesilkagit, K. (2014). Does task matter? the effect of task on the establishment, autonomy and control of semi-autonomous agencies. *International Review of Administrative Sciences, 80*(2), 318–340.

Weaver, R. K. (1986). The politics of blame avoidance. *Journal of public policy, 6*(4), 371–398.

Weiss, C. H. (1993). Where politics and evaluation research meet. *Evaluation Practice, 14*(1), 93–106.

APPENDIX

. A. LIST OF ZBOS, THEIR ABBREVIATIONS AND MAIN TASK

Abbreviation	Full name in Dutch	English translation	Primary task
AFM	Autoriteit Financile Markten	Financial Markets Authority	Regulation of Dutch financial markets.
BFT	Bureau Financieel Toezicht	Financial Supervision Office	Regulation of Dutch financial markets.
BKD	Stichting Bloembollenkeuringsdienst	Inspectorate of Flower bulbs	Certification and inspection of the quality of bulb plants in the Netherlands
CAK	Centraal Administratie Kantoor	Central Administration Office	Executive: registration of information in Dutch Healthcare
CBR	Centraal Bureau Rijvaardigheidsbewijzen	Central bureau for testing driver's abilities	Executive: test the driving skills of the general public and professionals
CBS	Directeur-generaal van de Statistiek	Statistics Netherlands	publish statistics to be used by policy makers, researchers, business and the general public

CCMO	Centrale Commissie Mensgebonden Onderzoek	Central committee on Research Involving Human Subjects	Regulatory
CEA	Commissie Eindtermen Accountantsopleiding	Committee for the attainment targets of the accountanta education program	Certification of accountants.
COA	Centraal Orgaan opvang asielzoekers	Central bureau for the reception of asylum seekers	Executive. Responsible for the reception and registration of asylum seekers.
CTGB	College voor de toelating van gewasbeschermingsmiddelen en biociden	Board for the Authorization of Plant Protection Products and Biocides	Regulation. Controls the use of pesticides in Dutch markets.
CvdM	Commissariaat voor de Media	Dutch Media Authority	Regulation. Enforcing compliance with the Dutch Media bill.
CvTE	College voor Toetsen en Examens	Board of Examinations	Certification and examination of students in education
DNB	De Nederlandsche Bank	Dutch Central Bank	As central bank, it makes monetary policies, controls inflation, circulates banknotes and coins and regulates financial organizations.

Huurcommissie	Huurcommissie	National Rent Tribunal	Regulation, solve disputes on rent, maintenance e.a. between tenants and landlords
Kadaster	Dienst voor het kadaster en de openbare registers	The Netherlands' Cadastre, Land Registry and Mapping Agency	Registration. Maps land property parcels and borders
KvK	Kamer van Koophandel	Chamber of Commerce	Registration and supporting businesses
LVNL	Luchtverkeersleiding Nederland	Netherlands Air Traffic Control	Executive. Controls air traffic within Dutch borders and on airports.
NIWO	Nationale en Internationale Wegvervoer Organisatie	Organization for national and international road transport	registration and certification of transport companies
NRGD	Nederlands Register Gerechtelijk Deskundigen	Netherlands Register of Court Experts	Registration of experts that are allowed to provide information and research in court.
NVAO	Nederlands-Vlaamse Accreditatierganisatie	Dutch-Flemish Organization for Accreditation	certification and accreditation of educational organizations
NWO	Nederlandse organisatie voor Wetenschappelijk Onderzoek	National Research Organization	Fund scientific sesearch

NZa	Nederlandse Zorgautoriteit	Dutch Healthcare Authority	Regulation of health care markets
OVV	Onderzoeksraad voor de Veiligheid	Dutch Safety Board	Investigative research into accidents and disasters in a large variety of policy sectors such as public transport, aviation, heavy industry e.a.
Politieacademie	Politieacademie	Police Academy	Education and training of the Dutch police forces
RDW	Rijksdienst Wegverkeer	Road Traffic agency	Registration, testing and controlling the quality and condition of motorized vehicles
RvA	Raad voor Accreditatie	Dutch Accreditation Council	Accreditation and certification of inspection and certification organizations
RvP	Raad voor Plantenrassen	Board for Plant Varieties	Registration of plant varieties
Staatsbosbeheer	Staatsbosbeheer	National forestry agency	Maintaining the Dutch forests and other nature reserves.
SVB	Sociale Verzekeringsbank		Executive: implement social security policies. payments of welfare benefits to Dutch citizens

SvJ	Stimuleringsfonds voor de Journalistiek	Stimulationfund for Journalism	Subsidize and support the quality, independence and diversity Dutch journalism
TNO-AGE	Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek - Adviesgroep voor Economische zaken	Netherlands Organisation for Applied Scientific Research - Advisory group for Economic affairs	As part of TNO, it Advises the ministry of Economic affairs on the use of the Dutch subsoil.
UWV	Uitvoeringsinstituut Werknemersverzekeringen	Employee Insurance Agency	Executive: payment of welfare benefits to Dutch employees
Verispect	Verispect	Verispect	Regulation. Regulate and enforce compliance with the Dutch Metrologiewet.
ZiN	Zorginstituut Nederland	Dutch Healthcare Institute	Regulation in Dutch healthcare: maintaining Dutch basic health insurance and stimulating quality of Dutch healthcare.
Six national governmental culture funds			
Mondriaan Fonds		Mondriaan Fund	Stimulate and subsidize visual art and cultural heritage

Stichting Fonds voor Cultuurparticipatie		Fund for Cultural Participation	Subsidize artistic and cultural activities for Dutch citizens with the aim of improving participation in cultural activities of citizens.
Stichting Nederlands Fonds voor de Film		Dutch Fund for the Film	Subsidize and support Dutch filmmakers.
Stichting Nederlands Fonds voor Podiumkunsten		Dutch Performing Arts Fund	Subsidize and support the professional performing arts, such as music, theater and dance.
Stichting Nederlands Letterenfonds		Dutch Foundation for Literature	Subsidize and support literary writers.
Stichting Stimuleringsfonds Creatieve Industrie		Creative Industries Fund	Subsidize and support architecture and indistrial design.

. B. LIST OF ANALYZED DOCUMENTS PER ZBO

AFM and DNB

- Doelmatigheid en doeltreffendheid AFM - verslag op basis van artikel 1:44, eerste lid, van de Wet op het financieel toezicht. (2010).

- De samenwerking tussen DNB en AFM op basis van de Wft Twin-peaks in de praktijk. (2010).

- Doelmatigheid en doeltreffendheid DNB - verslag op basis van art. 1:44, eerste lid, van de Wet op het financieel toezicht. (2010).

- Sent to parliament: 32466, nr. 1.

BKD

- Rapport van bevindingen betreffende de zelfevaluatie van de Stichting Bloembollenkeuringsdienst. (2004).

- Sent to parliament: 27831 nr. 10

- General Debate: 30300 XIV, nr. 10

BFT

- Toezicht en inzicht een helder denkraam. (2009).

- Sent to parliament: 29911, nr. 33 and 29911, nr. 47.

CAK

- Rapportage Onderzoek CAK. (2008).

- Sent to parliament: 29689, nr. 224.

- General debate: 30597, nr. 98.

- Written questions: 29689, nr. 253.

CBR

- Toekomstgericht onderzoek CBR. (2011).

- Sent to parliament: 29398, nr. 261.

- General debate: 29398 nr. 268.

- Written questions: 29398 nr. 266.

CBS

- Evaluatieonderzoek Centraal Bureau voor de Statistiek 2004-2009. (2011).

- Sent to parliament: 25268, nr. 74.

- General Debate: 33240 XIII, nr. 21.

CCMO

- Zelfevaluatie CCMO 2009-2013 - Derde rapportage taakvervulling CCMO, periode 2009-2013. (2014).

- Sent to parliament: 29963, nr. 13.

CEA

- Commissie Eindtermen Accountantsopleiding. Evaluatierapport. (2015).

COA

- Verslag van de visitatie van het Centraal Orgaan opvang asielzoekers. (2002).

Commissariaat voor de Media

- Rapport Evaluatie Commissariaat voor de Media 2007-2011. (2013).

- Sent to parliament: 33750 VIII, nr. 60.

- General Debate: 33750 VIII, nr. 87.

CTGB

- Report on the International Visitation of the Board for the Authorisation of Plant Protection Products and Biocides. (2013).

- Sent to parliament: 27858, nr. 217.

- General Debate: 27858 nr. 262.

Culture Funds

- Rapport Visitatiecommissie Cultuurfondsen. (2014).

CvTE

- Evaluatie College voor Examens. (2014).

- Sent to parliament: 34000 VIII, nr. 94.

- General Debate: 31289, nr. 249.

Huurcommissies

- Vervolgmeting maatschappelijk rendement van de huurgeschillenbeslechting. (2008).

- Sent to parliament: 28648 nr. 12.

- General Debate: 27926, nr. 128.

Kadaster

- Een driehoeksmeting op het Kadaster: sturing, financiering en kwaliteit bedrijfsvoering. Derde evaluatie Kadaster. (2010).

- Sent to parliament: 32500 XI, nr. 7.

KvK

- Evaluatie van de wetswijziging van 1 januari 2008 van de Wet op de Kamers van Koophandel en fabrieken. (2010).

- Sent to parliament: 32004, nr. 2 and 32004, nr.3.

LVNL

- Wettelijke evaluatie Luchtverkeersleiding Nederland, Periode 2005-2012. (2014).

- Sent to parliament: 31936, nr. 230.

- General Debate: 31936, nr. 257.

NRGD

- Evaluatie van het Nederlands Register Gerechtelijk Deskundigen. (2014).

- Sent to parliament: 29279, nr. 230.

NIWO

- Evaluatie NIWO 2010-2012. (2013).

- Sent to parliament: 33485, nr. 2.

- General Debate: 29398, nr. 408.

NVAO

- Report of the Panel of the external review of NVAO. (2012).

- Sent to parliament: 32210 nr. 24.

- General Debate: 33472, nr. 25.

NWO

- Nieuwe dynamiek, passende governance. (2013).

- Sent to parliament: 29338, nr. 121 and 29338, nr. 126.

- General Debate: 29338, nr. 129.

NZa

- Ordening en toezicht in de zorg. Evaluatie van de Wet Marktordening Gezondheidszorg en de Nederlandse Zorgautoriteit. (2014).

- Sent to parliament: 25268, nr. 87 and 25268 nr. 112.

- General Debate: 25268 nr. 115.

- Written Questions: 25268, nr. 123.

OVV

- Rapport Evaluatie Onderzoeksraad voor Veiligheid. (2014).

- Sent to parliament: 32008, nr. 2.

- General Debate: 29517, nr. 104.

Politieacademie

- De Staat van het Nederlandse politieonderwijs 2011.

- Sent to parliament: 29628 nr. 323.

- General Debate: 29628, nr. 349.

RvA

- EA Report of the Re-evaluation of Raad voor Accreditatie. (2014).

RvP

- Verslag ex artikel 39 Kaderwet zelfstandige bestuursorganen. (2012).

- Sent to parliament: 25268 nr. 75.

- General Debate: 33240 XIII, nr. 20 and 33240 XIII, nr. 22.

RDW

- Derde evaluatie verzelfstandiging RDW. (2011).

- Sent to parliament: 29 398 nr. 265.

- General Debate: 29398, nr. 287.

- Written Questions: 29398, nr. 276.

Staatsbosbeheer

- Ruimte in het bos! Staatsbosbeheer als maatschappelijke organisatie. (2009).

- Sent to parliament: 29 659 nr. 40.

- General Debate: 29 659 nr. 46.

- Written Questions: 29 659 nr. 42.

SVB

- Verslag van de tweede visitatie van de SVB.

SvJ

- Evaluatie van het Stimuleringsfonds voor de Pers 2006-2010. (2011).

- Sent to parliament: 32827 nr. 28.

- General Debate: 32827, nr. 41.

- Written Questions: 32827, nr. 31.

TNO-AGE

- Evaluatie van TNO-AGE. (2014).

- Sent to parliament: 30991, nr. 17.

- Written Questions: 30991, nr. 20.

UWV

- Rapport Commissie Interne Sturing UWV. (2010).

- Sent to parliament: 26448 nr. 448.

- General Debate: 26448, nr. 458.

Verispect

- Evaluatie instellingen onder de Metrologiewet. (2015).

- Sent to parliament: 33159 nr. 2.

ZiN

- Evaluatie Zorginstituut Nederland. (2014).

- Sent to parliament: 25268, nr. 99 and 25268, nr. 104.

- General Debate: 25268, nr. 121.

- Written Questions: 25268, nr. 108.

Incidents

- Incident mentioned, but not evaluated or judged in report
- Incidents evaluated in report

Political interest

- Presence of political interest:
    - The report report is mentioned in political debate
    - The ZBO is mentioned in political debate
- Ministerial reaction to the content of the report
    - political action based on recommendation
    - political action contrary to recommendation
    - political action not based on recommendation
    - political judgement of conclusions report: agreement
    - political judgement of conclusions report: disagreement
    - political judgement of conclusions report: positive
    - political judgment of incident mentioned in report
    - political mention of incidents outside of evaluation report
    - no political action on recommendation
    - no political reaction towards conclusion report

- Topics of political interest

    - cooperation between two ZBOs
    - customer satisfaction
    - Research design used in evaluation
    - financial
    - ministerial implementation of recommendations from the report
    - tasks and functioning of ZBO
    - governance
    - HRM
    - ICT
    - incidents
    - income from ZBO's own activities
    - kaderwet ZBOs
    - merging two ZBO's
    - minister defends reports conclusion
    - ministerial accountability to-

wards Parliament

- ministerial control over ZBO
- monitoring of activities ZBO
- organizational culture of ZBO
- organizational management of ZBO
- organizational position of ZBO
- organizational structure of ZBO
- other report
- performance indicators
- policy
- power of the ZBO
- relation ZBO and environment
- relation ZBO and ministry
- A different report on incidents.
- unrelated to evaluation unelated to ZBO
- transparency

Content of the report

- Methods used in evaluation report

  - visitation / peer-review
  - Benchmark or baseline measurement available
  - Benchmark considered but not applied
  - case study
  - Comparison with other ZBO
  - Comparison with previous evaluations
  - comparison with similar foreign organizations
  - qualitative data, documents and interviews
  - quantitative data: survey

- Self-evaluation
- Direct acknowledgement that tasks are difficult to express in measureable outcomes

- Topics of research questions

  - compliance with international standards
  - cooperation between two ZBOs
  - development of ZBO since autonomization
  - effectivity and efficiency of a ZBO
  - Incident mentioned in research question
  - internal organization
  - public accountability
  - relation ministry and ZBO
  - relation ZBO and environment

- Other topics researched or described

  - accountability for quality
  - adequacy of used instruments
  - administrative position
  - commercial activities
  - compliance with requirements
  - core task of ZBO
  - culture
  - customer satisfaction
  - dependency on other organizations
  - enforcement
  - evaluation of evaluation criteria used for other evaluations of the ZBO
  - financial

– governance

– legal basis of ZBO

– legitimacy

– organizational position as ZBO

– organizational, operational functioning

– outcomes

– personnel

– policy making and implementation

– quality of service delivery

– relation with environment

– relation with ministry

– scientific processes

– transparency

- Topics of judgments

  – incident mentioned, but not evaluated or judged

  – commercial activities

  – coordination between ministry and ZBO

  – coordination between ZBO and stakeholders/environment

  – costs

  – efficiency and effectivity

  – governance

  – ICT

  – indicators were bad and should be developed

  – long term strategy

  – managerial conflicts

  – new tasks

  – organizational culture

  – organizational management

  – organizational position

– organizational structure

– HRM

– transparency

- Topics of recommendations

  – improve accountability

  – change in policy

  – change laws and or regulation

  – Change organizational culture

  – change organizational structure

  – change organizational position

  – change tasks

  – change work processes

  – create competition

  – coordination between ministry and ZBO

  – coordination with environment

  – create better benchmarks

  – create better performance indicators

  – customer satisfaction

  – develop long term strategic agenda

  – finances

  – formalize policy choices

  – governance

  – ICT

  – maintain current organizational position

  – maintain current policy

  – overdue maintenance

  – personnel

  – improve public accountability

  – transparency

- Evaluation criteria: set by whom?

- defined by central government
- Handvest Publieke Verantwoording
- set by international standards

Miscellaneous

- task of the ZBO
  - accreditation, testing and certification
  - advisory
  - commercial activities
  - education
  - fund management
  - information collection
  - registration
  - research
  - stimulation of activities by other organizations
  - Subsidize
  - toezichthouder
  - uitvoeringsinstantie
  - vergunningen verstrekken

- Researcher
  - committee of peers
  - consultancy agency committee
  - committee of international umbrella organization
  - independent committee (other)

- Ministerial committee

- Client of the report
  - international umbrella organization
  - ministry
  - ZBO itself

- Directly mentioned reasons for evaluation
  - Reason for evaluation, directly mentioned
  - automatic/legal requirements
  - legal ground for evaluation: Kaderwet ZBOs art. 39
  - legal ground for evaluation: NOT kaderwet ZBOs art. 39
  - demanded in political debate
  - Handvest Publieke Verantwoording
  - Incident mentioned as reason
  - legislative changes
  - never evaluated before
  - new tasks were given
  - not mentioned
  - on initiative ZBO itself
  - positioning of organization
  - requirement for extension of membership
  - transparency