THE PLURIFORMITY OF POLICY EFFECT EVALUATION

An analysis of the effect evaluations of the municipality of Wierden in relation with its learning ability

---

C.C.G. (Chiel) Mensink
Master Public Administration
(Policy & Governance)

GEMEENTE WIERDEN

Governments will always play a huge part in solving big problems. They set public policy and are uniquely able to provide the resources to make sure solutions reach everyone who needs them.

Bill Gates, January 24th 2010

UNIVERSITY OF TWENTE
2015

THE PLURIFORMITY OF POLICY EFFECT EVALUATION

An analysis of the effect evaluations of the municipality of Wierden in relation with its learning ability

Author: C.C.G. (Chiel) Mensink
                 University of Twente
                 Faculty Behavioural, Management and Social Sciences
                 Public Administration – Policy & Governance
                 S-1621238

Graduation committee
University of Twente   Dr. H.G.M. (Herman) Oosterwijk
                             Dr. M.R.R. (Ringo) Ossewaarde
Municipality of Wierden   Mr. D.C. (Dick) Roessink

Place: Reutum ⁄ Wierden

Date: 22-12-2015

Status: DEFENITIVE

Correspondence address
C.C.G. (Chiel) Mensink
Ikinkstraat 5
7667 PT Reutum
ccgmensink@icloud.com

PREFACE

Before you lies the thesis 'The Pluriformity of Policy Effect Evaluation'. This thesis is written in the context of my master 'Public Administration' in the specialisation 'Policy & Governance' at the University of Twente. With this thesis the master program is completed and therefore results in the replacement of my student-career by a professional-career.

In January 2014 began for me an adventure in which I saw a lot of challenge accompanied by uncertainty. This adventure concerns my graduation period (bachelor) in Toruń, Poland, for the study Urban and Regional Planning at Saxion University of Applied Sciences. This period has brought me a lot. New insights, new knowledge, international experience and especially memorable moments in my passion for the Second World War. After successful completion of this study it became time to leave Toruń and to make the return trip to the beautiful region of Twente. Despite this highly instructive study, I felt not 'ripe enough' for the job market. I wanted more! During the study Urban and Regional Planning I obtained affinity with Public Administration. Hence the choice for the master Public Administration was not a surprise. During my pre-master as well as my master, I have been able to expand my knowledge and to form a different perspective of the world. Scientific education made a critical person of me. Do not assume anything, but look further and think! This was what I was missing in my previous study (bachelor). This master was able to make me 'ripe' for the job market.

The critical view that I developed during my master is also characteristic in this research. My aim was to carry out an external research in order to be able to contribute to a practical problem. Through my family in law I came in contact with the municipality of Wierden. Here I had my first meeting in May 2015 in which the municipality of Wierden presented three options. In the end this had led to the topic 'policy effect evaluation'. In my view this is a very important part of government activities. In times of social uncertainty and complexity it is of great importance to see to what extent governments are effective (acting legitimate). In addition, creating awareness of this uncertainty and complexity is an important step in the right direction. This research plays a role in the reorganisation/organisational development of the municipality of Wierden. I have experienced this period as a learning period. I obtained knowledge into the way a municipal organisation works and have gained many new insights during this process. Also, I have met many new persons. Conducting this research at the municipality of Wierden was very enjoyable. Nevertheless, the research process had also its downs. I spent much time to frame theories and unconsciously I widened the scope of the research losing the necessary focus. This has resulted in what Dr. H.G.M. Oosterwijk called "killing your own babies". Many theories, in which I have spent a lot of time, were unnecessary and could be erased.

I experienced the graduation period as very instructive, but also very enjoyable. There was a lively and open (informal) attitude at the municipality of Wierden. So my thanks go to the municipality of Wierden for their support, guidance and facilitation of the process. In particular I would like to thank Mr. D.C. (Dick) Roessink for his supervision, pleasant cooperation and interest in the research process. In addition, I would like to thank all persons at the municipality of Wierden who directly or indirectly contributed to the data collection of my research. This concerns in particular the persons interviewed. Special thanks also goes to Dr. H.G.M. Oosterwijk (first supervisor) for his knowledge, insights and information and the conveyance herein as well as his support and supervision. The meetings were often held in discussion-form instead of question-answer-form. I experienced this form of supervision as very pleasant and instructive. I would also like to express my gratitude to Dr. M.R.R. (Ringo) Ossewaarde (second supervision) for his help, support and time. Finally, I want to thank my parents because they made me able, by trust and support, to complete my student-career.

Reutum / Wierden, 22 December 2015

C.C.G. Mensink

ABSTRACT

A long time people lived in the idea that society was makeable. Problems in society could be solved with the bureaucracy in which the government monocentrically solves problems. However, during the 70s of the last century there was more awareness that policy evaluations should play an important role in government operations (Hoogerwerf & Herweijer, 2014). Because it does not seem obvious that certain intended effects are achieved, policy evaluation is needed to determine the effect of policy more effectively. Uncertainties concerning effects and the complexity of society is caused by various developments in government and society. Our cognitive ability is limited to understand this complexity resulting in uncertainties and risks. This endorses the importance of policy evaluations focusing on effect measurement. Complex problems are characterised by very little certainty of knowledge and/or little consensus about norms. This leads in practice to situations where the policy can be influenced by external variables. These variables may exert negative influence on the policy, but can also contribute positively. This may lead to situations in which goals are achieved, but that these can hardly be attributed to the policy. Policy evaluations focused on effect measurement can provide insights about this. Despite the fact that the municipality of Wierden acknowledges the need for policy evaluations, it also acknowledges that this forms a weak part of the policy process and that is does not implement many effect evaluations. Therefore this research focuses on the following question:

How does the municipality of Wierden perform policy effect evaluations and how can the learning ability of the organisation be increased based on the effect-measuring performance?

There are several forms of policy evaluations focusing on effect measurement. Based on the theories of Korsten (2013), van de Graaf and Hoppe (1996) and Bekkers (2012) an evaluation ladder is prepared. This ladder ranks different policy effect evaluation forms based on reliable evidence that the evaluations provide. The position on the ladder largely determines what the learning ability of the organisation is. This means that the higher up the ladder, the more reliable the results. Argysis and Schön (1974) identify three feedback loops in which an organisation is able to learn. The first focuses on the adjustment of actions (single-loop learning), the second focuses on changing the assumptions (double-loop learning) and the third focuses on learning to learn (triple-loop learning). This largely determines the learning ability of a government/organisation which is closely related to policy effect evaluations.

This research made use of document analyses and interviews. The document analysis have made clear which policy evaluations exist (de facto), how many policy evaluations focus on effect measurement and how the effectiveness of the policy is made transparent. It also provides insight into the obligatory (imposed) actions concerning policy effect evaluations from legislation and regulations. The interviews have made clear how the municipality of Wierden thinks concerning policy effect evaluation. The interviews also provided explanations about aspects of performances regarding effect evaluations. This had resulted in the de jure and de facto image of the municipality of Wierden on policy effect evaluations. The results show that the municipality of Wierden complies with the mandatory aspects of policy effect evaluations, but that it does not fully comply with the codified ambitions they have set out in ordinances. This concerns in particular the optimisation of goal-orientation, efficiency, effectiveness and quality of product and process as well as writing explicit goals. The results reveal that the municipality of Wierden carried out 16 effect evaluations in the period 2010-2015 (de facto effect evaluations). The number of effect evaluations are not distributed in an equal manner between the organisational sectors. The reason for this is that employees state that one sector have more operational (policy related) tasks than others. The de facto effect evaluations are tested on validity and creditability, based on the literature. This has resulted in the positioning of the de facto effect evaluations on the evaluation ladder. This shows that the de facto effect evaluations score low on the evaluation ladder and therefore are implemented in a fairly intuitive way. This means that there is no certain knowledge about the outcomes of the policy and that the municipality of Wierden is unable to draw conclusions about the success or failure of the policy.

The municipality of Wierden ignores the external factors in carrying out effect evaluations. In addition, the results indicate that not in all cases use is made of explicit goals, but that the relationship between the criteria and the instruments do not take place in an adequate manner as well.

Based on the de facto effect evaluations it seems that the municipality of Wierden is not able to learn in an optimal way of the de facto effect evaluations. Even more important, the learning ability of the municipality of Wierden with regard to the de facto effect evaluations is very limited. Interviews with employees as well as analyses form policy documents reveal that the municipality of Wierden only uses the results of the evaluations to adjust policy (single-loop learning). This means that it does not bring into question the assumptions of the policy (double-loop learning) and is not capable of learning to learn (triple-loop learning). In addition, the evaluations and their results only focus on cognisance to employees and to politics. In order to perform a successful learning process it requires, after cognisance, influencing standpoints and to create impacts of the findings. These last two aspects of the learning process is not explicitly done by the municipality of Wierden.

In conclusion, it can be stated that the municipality of Wierden does not perform effect evaluations in an adequate way which result in invalid findings. A major risk in this is that these uncertain results could be used for adjusting policies (single-loop learning). This means that policies are adjusted based on data that is not based on evidence and therefore could be erroneous. In order to create a more reflexive organisation, it is important for the municipality of Wierden to increase its learning ability with respect to effect evaluations. The validity of the evaluation is closely related to the learning ability. The municipality of Wierden can do this by performing effect evaluations in a more systematic way. In order to perform valid and adequate effect evaluations it is recommended to pay attention to the next action points:
- involving citizens and actors in the preparation, implementation and evaluation of policy;
- recording and enforcing responsibilities regarding effect evaluations;
- discuss the results of evaluations with those involved and use the results to double-, and when possible triple-loop learn;
- drafting of explicit (quantified) goals in advance;
- establishing the relation between the baseline measurement and the measurement after performing the policy (in the evaluation);
- establish the relation between criteria and instruments in respect of the goal.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Context

Policy implementation largely determines whether a policy is successful. But when can we speak of successful policy? The answer to this question can be given by performing policy evaluations. In the policy and government related literature, various definitions of the term 'policy evaluation' can be found. Although in parts they (may) vary, most definitions have in common the following: assessing policy on the perceived *content*, *processes*, or *effects* based on certain *criteria* (Blommestein, Bressers & Hoogerwerf, 1984; van de Graaf & Hoppe, 1996; Hoogerwerf &Herweijer, 2014; Bekkers, 2012). In this study, the focus is placed on the third aspect of the 'common' definition, namely: the effectiveness of policies. This definition also implies that an evaluation should always be carried out on behalf of certain criteria and the perception of the policy. Since the seventies of the last century, policy evaluation has become an important part of the policy process in Western societies (Hoogerwerf & Herweijer, 2014). Nowadays, policy evaluation is indispensable and forms an integral part of the policy process. What are the reasons for the (sudden) emphasis on policy evaluation? It should be clear that the demand for policy evaluation is accompanied by some developments within government. Hoogerwerf and Herweijer (2014) indicate that the fundamental reason for the emergence of policy evaluation is the result of the changing role of the government. Western societies underwent a transition from night-watchman state to social welfare state. This changing role has given the government more tasks which increases the need for citizens and private parties to study its efficiency and effectiveness (social justification). Additionally, questions around government intervention in the economic and social domain result in the demand for evidence of performance and accountability (Sanderson, 2002). Increased spending on public services reinforced this demand, as Sanderson (2002) presupposed. These arguments suggest that government responsibility and legitimacy play more important roles. Policy evaluation can thus provide evidence for governmental actions and interventions.

The 'external' call for performing policy evaluations (people from outside the government) is not the only incentive for conducting policy evaluations. Internally, that means the government itself, also sees the value and necessity to conduct policy evaluations. Korsten (2013) presented two arguments for this internal need for policy evaluation. First, many policies are not able to achieve the goals because the problem under consideration is simply too complex (uncertainty and pluriformity of problems). The second argument focuses on the intertwining of policies that require a cognitive basis of the policy maker (policy intertwining). These arguments substantiate the internal demand for policy evaluation and the need to come up with evidence in order to judge if policies actually are successful. Science, thus plays a crucial role in policy evaluation. Important in thinking about – and theorising of – policy evaluation is that conducting policies establishes the relationship between the 'governing' and the 'governed' (Lascoumes & Le Gales, 2007). This accommodates the expectation that the 'governed' wants to know if the 'governing' solve the policy problem(s) in question. Although citizens do not always explicitly acknowledge that they want to know whether a policy 'worked', it is expected that the quest for this is at all times present. Nonetheless, policy evaluations that focus on effectiveness can provide important information to improve the quality of policy and may also be important for the choices of political parties as well as the knowledge about effectiveness may contribute to broader discussions with other organisations (van Elk, van de Meer, van de Steeg & Webbink, 2011). In an era where the government interfered with more and more aspects of society, in an era of modernisation and globalisation and in an era of uncertainty and complexity, it can be taken for granted that policy evaluation becomes an increasingly vital element in government actions. But how do governments deal with policy effect evaluations in practice? By means of a case study, this question is applied on the municipality of Wierden.

1.2 Problem definition

The changing government (from night-watchman state towards welfare-state) as well as the increasing pluriformity of social problems has put the government in a more 'directing' role rather than the traditional 'executive' role (transition towards governance) (Hague & Harrop, 2013). The decentralisation of government tasks (in 2015) – youth care, employment and income and prolonged care for sick and elderly – towards municipalities has reinforced this role. In this context, at the municipality of Wierden the demand for policy evaluation for accountability and evidence of performances is also growing.

Despite the introduction of policy evaluation at the municipality of Wierden, it is acknowledged that this part forms a weak element in the policy process. This recognition stems from the little policy evaluations that are carried out which are focused on effect-measurement. Within the municipality of Wierden blooms the feeling that officials are 'just doing something' (who is responsible for what?). The emphasis lays on 'doing things' instead of 'thinking things over' and to pay attention to 'if what they are doing is done effectively'. This inadequate way of policy evaluation asks for an analysis in order to clarify how the municipality of Wierden measure the effectiveness of policies.

The aim of the municipality of Wierden therefore lies in the quest for performance with regard to policy effect evaluations. The context of this study as well as the notion of weak implementation of policy effect evaluations by the municipality of Wierden show that policy evaluation is not an easy matter. This research will focus on analysing the municipality of Wierden regarding policy effect evaluations. In this analysis is going to be searched for ways in which the municipality of Wierden can increase its learning ability. In response to the above, the following research question is formulated:

> *How does the municipality of Wierden perform policy effect evaluations and how can the learning ability of the organisation be increased based on the effect-measuring performance?*

In order to answer this research question, different sub-questions are formulated. The sub-questions are:
1. *What are the de jure obligations for the municipality of Wierden regarding policy effect evaluations?*
2. *To what extent does the municipality of Wierden perform de facto policy effect evaluations?*
3. *How can the learning ability for the municipality of Wierden be improved regarding policy effect evaluations?*

1.3 Research goal, scope and relevance

This section focuses on the goal of the research which is closely related to the research question. Additionally, it provides a demarcation of the research. In this demarcation is indicated what is disregarded and where the research specifically focuses on; the scope. Finally, attention is paid to the social and scientific relevance of the research.

1.3.1 Goal of the research
The goal of this research is to provide insight into the way the municipality of Wierden perform policy effect evaluations. The information resulting from this research amounts to creating awareness in order make adjustments in the way the municipality of Wierden makes use of policy effect evaluations on behalf of new insights (based on the conclusion and recommendations of this research). The new insights and the way in which the municipality currently performs policy effect evaluations can be used in the reorganisation of the municipality to increase its learning ability.

1.3.2 Scope and demarcation

As the research question suggests, this research focuses on policy effect evaluation of the municipality of Wierden. This means that the focus is laid on the third aspect of the definition of Hoogerwerf and Herweijer (2014), namely: the effects of policy. This research focuses only on the policy effect evaluations (ex post) of the municipal organisation itself. This means that policy evaluations carried out by other bodies are not part of this research. The Court of Auditors and other organisations that perform independently, implicitly or explicitl, policy evaluations are not included in this research. The reason for this is that the goal focuses solely on how the municipality itself could use policy effect evaluations. However, this does not mean that other bodies do not play an important role in evaluating policies. In addition, within the municipal organisation the research is demarcated by only focusing on the direct (explicit) policy evaluations. Indirect (implicit) policy evaluations are therefore not part of this research. An example of an implicit policy evaluation is when policy is being reformed when is experienced that the policy is not working properly which decision is not based on a conducted evaluation. This research also requires a demarcation in time. It is decided to take 2010 as starting point for the research (especially in the search for documents). In consultation with the municipality of Wierden, this starting point in time is assumed to be suitable because it allows to investigate important elements (such as strategic visions and changes within the organisation). Additionally, it may be assumed that 5 years provide enough information and data to get an adequate picture of the how the municipality makes use of policy effect evaluations. The demarcation of 5 years also offers the possibility to investigate periodic evaluations.

1.3.3 Relevance

Relevance of the research can be seen as the contribution to society and to science. Below, both contributions are elucidated.

Social relevance

The social relevance aims to contribute to knowledge about policy effect evaluation at the municipality of Wierden and the application of policy effect evaluations in practice. It allows the municipality of Wierden to (better) evaluate policies resulting in the creation of learning ability and in the stimulation of conducting effect evaluations. The knowledge and awareness are significant for the development of successful policies, for societal and political support of policies, and for the legitimacy of governmental actions.

Scientific relevance

The scientific relevance lies hidden in the application of scientific theories about policy effect evaluation (the ideal situation) in comparison with the practice (de facto situation). This research investigates how the municipality of Wierden performs policy effect evaluations in order to increase the learning ability of its organisation. It supplies new insights about the way policy evaluation is theoretically described in the literature and the way in which is dealt with it in practice: the contrast between theory and practice. The insight offers responsible persons and bodies grip on results of their actions. The review, but also the merging of scientific theories makes a complete and detailed picture of the theoretical situation.

1.4 Reading guide

This chapter has given substance to the inducement of the problem and the research questions which are distilled from the context/literature. This has resulted in the scope of the research as well as the relevance. Chapter 2 presents the relevant theories regarding policy effect evaluation. This theoretical framework is central to the research and serves as a basis for the analyses. The way in which the research is conducted and the strategies that are used, are described in chapter 3. In this chapter, the methods are disclosed that are used in collecting and analysing the research data. Chapter 4 then shows the results of the analyses. This chapter contains the core of the research in which all information (theories, data and the analysis of it) come together. Ultimately, this results in a conclusion in which recommendations are given regarding policy evaluation at the municipality of Wierden, chapter 5.

2. THEORETICAL FRAMEWORK

This chapter provides the theoretical foundations to be able to make statements about the functioning of the municipality of Wierden concerning policy effect evaluations. This framework will be used to put the findings of the research in a context in order to draw conclusions. Attention is paid to different components in relation with policy effect evaluations. First, the complexity of society is highlighted in which it becomes clear that policies are accompanied by uncertainties. In the second part, attention is drawn to different types of effect evaluations in relation with the degree of certain knowledge it produces. Additionally, the learning ability of the government as organisation in the context of effect measurement is elucidated. Finally, a theoretical conclusion is drawn in which expectations are stated on behalf of the theories.

2.1 The complexity of governing

"Government is not the solution to our problems, government is the problem". This famous statement by US-President Ronald Reagan on January 20, 1981 accentuates a significant change in the mind-set in which people try to manage society. Previously the idea lived that society is 'makeable' and that all problems could be solved by governmental actions (Bekkers, 2012). Good government was the utopian answer to 'makeability'. However, in uncertain and complex societies of today, this is proven to be an illusion. The way society is currently beheld and the way one is trying to influence society can be placed in a historical context, which roots date back to ancient Europe.

2.1.1 Understanding complexity

Rationalisation of society
The sociological development of Europe may be referred to as 'rationalisation'. It concerns a change in the understanding and practice of reason, in other words, it relates to the pattern of change of the way in which daily life is understood and beheld (Ossewaarde, 2013). In ancient Europe communities were ordered by value-rationality. This rationality can be seen as a society in communities inspired by shared values and common beliefs. Ossewaarde (2013) talks about a society that is declared from mythological explanations in which spells and curses are central. This ancient philosophy can be characterised as the 'logic of appropriateness'. This logic implies that one acts according to its own norms and values (Hague & Harrop, 2013). This is supported by March and Olsen (2009). They emphasise that this way of thinking is only based on the perspective that human actions sees as driven by rules of appropriate acting because they are considered natural, rightful, expected, and legitimate (a valued goal). In shaping European societies, value-rationality was criticised. They distanced themselves from the mythological explanations of reality (Ossewaarde, 2013). The democratic order was opposed to the given ideas about reality and so it was desirable to renounce from values. This replacement of ideal types can be typified as the transition to 'goal-rationality', also called instrumental or technical rationality (Ossewaarde, 2013). Goal-rationality can be seen as the 'logic of consequences', as opposite pole of the 'logic of appropriateness' belonging to modernist Europe. The logic of consequences, according to Hague and Harrop (2013), points to instrumental behaviour that is shaped by the achievement of a specific goal. In this logic, legitimising governmental actions is based on instrumental expediency (output legitimacy) and constitutional legality (input legitimacy).

Goal-rationality and the market-based society during the emerging industrialisation in the 19th century led to an unequal society (the working class on the one hand and the bourgeoisie: managers, factory owners, bankers, etc. on the other hand) (Ossewaarde, 2013). This has led to the ingress of social laws and social policies in governments in Western Europe in order to secure equality. However, the implementation of these laws and the embodiment of policies has had a significant side-effect. It is in this

period (late 19th century) that government was confronted with unintended effects of policies. The application of social rights did not seem to go fair in practice (despite the attempt to do so). This phenomenon is known as the so-called 'Matthew Effect', an important phenomenon that accentuates the limited capacity of governments in effective policies. Ossewaarde (2013) describes this phenomenon as:

> While the social citizenship status fulfils the manifest functions of reducing social inequalities between classes, creating equal opportunities and establishing a fairer distribution of national prosperity accumulated in commercial society, in practice, social policies seem to result in better social positions for the already advantaged bourgeoisie. (p. 130)

This Mathew Effect (together with manifest functions and latent functions), is conceptualised by Robert Merton. He describes the manifest function as the conscious (recognisable) function as intended (envisaged) and describes the latent function as an unconscious (unrecognisable) function as it is not intended (not envisaged) (Merton, 1957). This phenomenon shows that there are some unexpected consequences they had not contemplated in advance: social laws guaranteeing equality which result into (only) benefits for the already advantaged social group, strengthen the inequality.

Although people were aware of the fact that in complex situations policy may be associated with unintended effects, policy effect evaluations was not introduced yet in order to clarify the effects of policies. The goal-rational approach of society led to the focus of policymakers on the input-side of the policy process instead of focussing on the output-side. So people acted only goal-rational. Also in the course of the ninetieth and twentieth century, Weber's bureaucracy took shape. According to Weber (in Ossewaarde, 2013) is bureaucracy a power structure in which the government – usually from behind their desks – governs citizens by laws, regulations and policies. This setup is fairly structured in nature (legal-rational) and does not take into account interests or interactions with citizens. Hoppe (2011) calls this type of governing the 'question-and-answer-game'. The citizens raise questions which are answered by the government, expecting government to solve the problems with its legal system. It is a "paternalist, statist canon, with emphasis on collective actions and faith in bureaucratic rationality and professional autonomy" (Hill & Hupe, 2002, p. 88). In short, the makeability of society was in this period deeply ingrained in government thinking and government actions.

All at sea: uncertainty in times of interventionism
It is now clear that the industrial revolution (including the Matthew Effect), the changing role of government and the rationalisation process has led to a society in which much uncertainty exists. One does not know for sure whether intended policy effects will be achieved thanks to the (possible) presence of latent functions. But how is it possible to (better) understand this complexity? What role does policy effect evaluation have in this uncertainty?

An important question that one may ask is: Is it possible to eliminate latent functions? The answer to this question depends on the degree of complexity of a problem. If the problem is complex, the 'law of unintended consequences' applies. This law implies that "…in complex situations, any action you take leads inevitably to unanticipated outcomes" (Green, 2014, p. 3). Additionally, Green (2014) makes a very important statement, namely that if one tries to grasp the complexity, one should not ignore it. "In the twentieth century, and especially since the Second World War, the work of the government in the Netherlands as well as in other Western societies drastically changed character" (Hoogerwerf &Herweijer, 2014, p. 152). Due to the changing role of government in public policies, the transition took place from night-watchman state to the welfare state. Several sociologists in the 1970s assumed that a post-industrial society is emerging (Ossewaarde, 2013). The change (or transition) which occurs here is that fewer hands are needed for the work to be done and that more information data is needed. This is defined by Ossewaarde (2013) as a society in which factories no longer mainly dominate, but where offices, computers and media plays a more important role. Hoogerwerf and Herweijer (2014) identified three reasons for the emergence of policy evaluation as reaction on 'grasping complexity':

1. Interventionism: the changing role of the government (transition from night-watchman state to welfare-state) and the intensification of government in active policy domains.
2. Financial elaboration of interventionism: spending cuts and a more effective use of collective resources calls for greater accountability of policy implementation and policy effects.

3. Methodological developments: More methods and techniques are accessible which makes it easier to report the creation, implementation and effects of policy.

Interventionism as well as the complexity of the society has made the government realise that policy is not always effective. Since the seventies of the last century, evaluation of public policies in Western societies came into emergence (Hoogerwerf & Herweijer, 2014). As a response to this 'development' it became necessary to change the government and its organisation. The demand for change was crystallised by Osborne and Gaebler (1992). In the arguments in favour of chancing government, they emphasise among others as an analogy that government must 'steer' the boat instead of 'rowing' it. They also endorse the importance that governments should focus on outcomes instead of inputs. As a reaction to this incentive, New Public Management (NPM) emerged during the 1980s. Hague and Harrop (2013) describe this phenomenon as an anti-bureaucratic development. NPM enables the government to possess a market-based focus with a proactive attitude. In order to do this, governments need to focus on quantifying and measuring performances on the basis of performance indicators (Bekkers, 2012). "The definition also emphasis that evaluation should aim at helping the accountability of policy makers for the use of public resources, the effects of the implemented policies, and the reasons for choosing a specific alternative" (De Marchi, Lucertini & Tsoukiás, 2014, pp. 11-12).

The government is confronted with insoluble problems or problems that can be solved only partially (Hoppe, 2011). The uncertainty and complexity facing governments calls in question the capacity of the government to react on problems. "Most proximate policy makers felt that they were not longer as capable of formulating, implementing and evaluating policy as they were during the heyday of government during the 1960s and 1970s" (Hoppe, 2011, p. 37). This criticism on the question-and-answer-game of Weber's bureaucracy resulted in incentives to reinvent government (Hoppe, 2011). Nowadays, it is thus an illusion to think that the government is able to solve every problem. This can be seen as the path-dependency and the resource-dependency of the government in a commercial setting. The awareness that government is not always able to tackle problems, and perhaps to distance themselves from goals, has led to the replacement of the goal-rational society by a society which is characterised by what Ossewaarde (2013) defines as 'reflexivity'. Ossewaarde (2013) describes reflexivity of society as a growing compulsion to reflect on the consequences of ideas, emotions and activities. This distinctive way of thinking can be seen as social movements that emphasise the role of power relationships in the production of knowledge (Ossewaarde, 2013).

Reflexivity is also based on new information that is mainly available through the ICT. Reflexivity can therefore actually be seen as "...a lay reflex reaction to the ignored and unplanned consequences of modern scientific and technological achievements, such as the explosion of nuclear plans, ecological disaster, terrorist attacks, fraud involving financial products, inefficient bureaucracies, or ineffective procedures" resulting in a highly disordered society (Ossewaarde, 2013, p. 171). The post-rational approach to society has deep-rooted influence on the manner in which policy evaluations are conducted. In the Netherlands this led to the so-called 'VBTB-operation' (Van Beleidsbegroting Tot Beleidsverantwoording) at the turn of the century. This operation gave greater weight to actual results of policy (Bekkers, 2012) and that the realisation of policy and policy goals are measurable in order to assess them based on performance records to see whether policy is successful (Voermans, 2002). The government accountability for policy results emphasise the need and desire for evidence in complex situations. Therefore, governments started to focus on 'evidence-based policy' (EBP). EBP and NPM need professional evaluations for verification and legitimation to identify if government intervention had the desired and predicted effects (Power, 1999; Wesselink, Colebatch & Pearce, 2015). People want more control over government actions and want reliable knowledge about 'what works' (Sanderson, 2002), especially in a welfare state in which the government is active in many domains of society. EBP is seen as reducing uncertainty by basing policy on scientific evidence that, in effect, can provide 'hard data' about the achievability of the desired effects (Bulmer et al. in Sharman & Holmes, 2010). EBP is thus about 'learning from the past' in which empirical evidence is delivered on which policy is based (Pawson, 2002). With this evaluation, people seek to reduce uncertainty in order to minimise the latent functions. Van de Graaf and Hoppe (1996) shows this intricacy of uncertainty of variables in policy effects, demonstrating the 'law of unintended consequences' (figure 1 on the next page).

Figure 1 – General model of variables in evaluation of policy effects
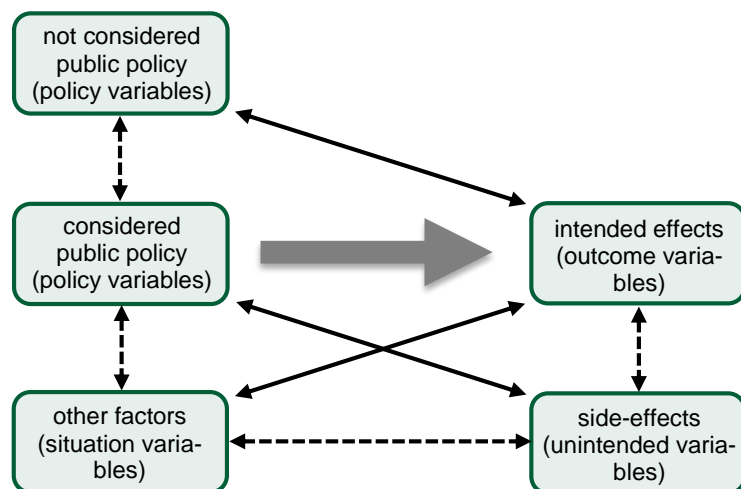
Derived from: van de Graaf & Hoppe (1996, p. 394)

The model shows that a distinction can be made between intended effects and side-effects (or manifest functions and latent functions). The various arrows indicate that there are multiple relationships. For example, the intended effects also depend on not considered public policy, but also on other factors in society (situation variables). This model provides an important replenishment to the distinction between manifest functions and latent functions, namely: there may be variables present that exercise, apart from the implemented policy, positive influence on the policy results (the problem). Although the manifest functions and latent functions are focused only on the intended and unintended effects, it is also possible that situation variables play a role in the implementation of policies and contribute positively to the effects of the policy. In such a situation, it is difficult to clarify the intended effects are due to the implemented policy. In short, van de Graaf and Hoppe (1996) clarify with this model that many factors play a role in a post-modern society and therefore providing evidence can be a fairly difficult task. This also reveals an important limit of evidence-based-policy. Evidence of prior policy implementations or researches does not assure factual and reliable data for future use by copying it. In order words, evidence-based policy may still be confronted with the 'law of unintended consequences'. Strassheim and Kettunen (2014) indicate that "public conflicts indicate that evidence-based policy rests on its own mechanisms of exclusion and selectivity…" (p. 260). EBP is no guarantee for certain knowledge about complex situations. Policy should therefore not be evidence-based, but should be evidence-inspired.

Imperfection of human beings: bounded by rationality
Unintended effects may be caused because too much information is needed to clarify a situation. The ICT has led to endless streams of information available forcing people to interpret their daily experiences (Ossewaarde, 2013). Bekkers (2012) calls this 'information overload' which creates a deficit of rationality. He interprets this as follows: "The cognitive abilities of people and organisations to understand this complexity, to fathom and to be able to undertake specific actions to achieve certain goals, fall fundamentally short" (p. 29). The quest for certainty (or the reduction of uncertainty) lies in making certain choices. Francot (2014) demonstrates this necessity: "Complexity…refers more in particular to social complexity: the type of complexity that emerges from the relationships between human beings and the myriad of options and possibilities that exist in our society" (p. 201). It also plays an important role in the chosen policy instrument. Policy instruments have their own 'force of action': "…as they are used, they tend to produce original and sometimes unexpected effects" (Lascoumes & Le Gales, 2007, p. 10). This endorses that in controlling uncertainty, the choice of policy instrument also needs to be considered and thus the relation between the goal (criterion) and the indicators.

Apparently, people are not able to conceive societal complexity. The reason for this, as it is acknowledged by Simon (1972) and Jones (1999), is the presence of 'bounded rationality'. They claim that

bounded rationality recognises that there is a lack of human cognitive. "Cognitive limits of human decision makers imposed limits on the ability of the organization to adjust to its environment" (Jones, 1999, p. 303). This means that people want to act rationally in order to change towards the desired situation, but in practice, as a result of emotional setting and cognitive capacity of human being, sometimes fail (rational deficit). According to Simon (1972) this introduces risks and uncertainties. The limitation of human cognitive ability also implies that people have incomplete information about alternatives, as Simon (1972) presupposed. Bounded rationality can thus be seen as the answer to the occurrence of the Matthew Effect. This is what van de Graaf and Hoppe (1996) imply in figure 1. The paradoxical phenomenon that takes place here is that people think that due to the policy the intended effects are reached, but in fact other factors (external factors) may have affected the intended effects and because people do not have sufficient cognitive capacity to enable them to provide insight into these influencing factors, it may be that the goals are achieved but that they are not or partly due to the policy itself. It may also occur that the policy has unforeseen (negative) effects: latent functions.

2.1.2 Structuration of problems
Till now the focus laid on the effects of policy implementation (problem solution). This is not illogical because people rather start working on solutions instead of working on problem finding (Klinkers (2002). A problem can be seen as the difference between 'is' (the current situation) and 'ought' (the desired situation). From the previous section can be learned that society is complex and that policy may be associated with unintended and/or unforeseen effects. In such a situation, it may be thought that the problem in question is not properly analysed. However, bounded rationality teaches us that we are simply not always capable of doing so. If society is that complex, how is it possible to solve/address (complex) problems? In other words, how is one able to analyse problems?

Approaching problems: the dimension of normative standards
The first dimension of approaching problems is the degree of consensus about normative standards. Here the 'logic of appropriateness' applies and relates mainly to subjectivity and opinions about norms. The 'logic of appropriateness' forms an important role in the way in which social relations are formed and organised. Ney and Verweij (2014) conceptualise this as the 'Cultural Theory': "Each way of life provides a specific set of 'perceptual lenses' that helps its adherents make sense of the world by defining focus and salience" (p. 625). This means that certain policy issues, in a given period of time, receive more attention while other policy issues remain at the background. This is also the reason why many policies fail to address issues in complex and uncertain problems or in intractable policy controversies (Ney & Verweij, 2014). In addition, one must also be aware of the fact that human acting synonymous is with rational acting (Bessant, 2008). Important conclusion Bessant draws from this is that people would be acting rationally, but this rationality is highly dependent on individual ideas, intentions, opinions and beliefs and therefore can vary greatly for each person. This means that "each individual 'has' an individual self, a mind and an independent set of wants or needs and is regarded as a static agent with clear boundaries that separates them off from other individuals" (Bessant, 2008, pp. 288-289).

Approaching problems: the dimension of knowledge
The second dimension of approaching problems is the degree of certainty about knowledge, which already got attention in the previous sections. Recall that bounded rationality in complex situations shows that people are unable to understand the complexity, resulting in uncertainty in which risks arise. If rationality fails, it means "…a mismatch between the decision-making environment and the choices of the decision maker" (Jones, 1999, p. 298). This does not only mean that one has a lack of information, on the contrary, one could also have too much information where the human brain is not able to deal with (information overload). This means that one cannot consider all aspects of a problem (Jones, 1999). In analysing (complex) problems, no complete search for information is enacted, in particular when factors are determined in order to characterise the problem. Providing insight into all possible alternatives – for addressing the problem and to come to a solution – is therefore not possible. This is precisely why policy effect evaluation is important, namely: to see whether the choices made are the right ones and whether it has led to the desired effects. In practice, this could lead to situations in which problems are not automatically followed by solutions, but that there is a set of solutions which can be applied to the

problems which arise (Jones, 1999). Iedema and Wiebinga (2010) call this model of problem approach the 'garbage-can model' in which problems and solutions will converge. This model is, according to them, mainly applies in situations when a policy problem has a large degree of uncertainty. The uncertainty about knowledge has in two ways a significant impact on conducting policy (in Bessant, 2008). Firstly, there may be uncertainty about the cause of the problem. Secondly, there may be uncertainty whether the policies actually address those aspects which causes the problem. This dimension refers to the 'logic of consequences' because it always is in search of cause and effect.

Typology of problems

The two dimensions indicate whether a problem can be regarded as 'complex' (unstructured problem) or as 'simple' (structured problem). Van de Graaf and Hoppe (1996) as well as Hoppe (2011) acknowledge these two dimensions and displays them in a matrix (see figure 2). The matrix identifies the two dimensions (certainty of knowledge and consensus about norms) resulting in four different problems, each with its own character.

Figure 2 – Types of policy problems

Certainty of knowledge

		large	little
Consensus about norms	large	*(1)* **structured problems**	*(2)* **moderately structured problems (goals)**
	little	*(3)* **moderately structured problems (means)**	*(4)* **unstructured problems**

Based on: van de Graaf & Hoppe (1996) and Hoppe (2011)

A problem can thus be considered as a complex problem, when:
   a. there is little certainty of knowledge about the problem, or;
   b. when there is little consensus about norms, or;
   c. when a and b take place jointly (defined as a wicked problem by Rittel and Weber (in Hoppe, 2011)).

From the above can be seen that problems not only the struggle is between the problem at hand and the knowledge available about the problem. Van de Graaf and Hoppe (1996) emphasise this: "…the image that is present at policy makers is in the first place the factual circumstances, events, and developments which, with respect to the policy in question are relevant; and in the second place norms and values, which they, with respect to the policy in question deem applicable or worth pursuing" (pp. 68-69). It is clear that policy makers always strive to transform problems into tamed problems (structured problems) to come to a problem-solving situation that is supported by everyone. Figure 1 shows that unintended consequences may occur. Figure 2 shows more thoroughly what variables could cause these effects. The variable 'consensus about norms' is very subjective in nature and as Bessant (2008) identifies, this varies for each individual. Controlling and striving for consensus is therefore not feasible in every case. A report of the 'Organisation for Economic Co-operation and Development' (OECD, 2000) made clear that lack of consultation with stakeholders may lead to non-compliance of proposed policies and regulations. The report endorses that adequate consultation will lead to an increasing commitment of stakeholders to the objectives of policies. Because of this, government needs to interact in order to solve problems (increase effectiveness). This is the reason that scholars argue that decision-making processes should be more based on deliberative grounds in complex problems: conflicts that arise in policy processes are often based on normative ethical issues and these issues cannot be solved only by rational analysis, but are part of a social construction of the policy problems (Kohoutek, Nekola & Novotny, 2013). The variable 'certainty of knowledge' indicates that evidence is needed to reduce uncertainty in order to minimise unintended (latent functions) and unforeseen effects.

2.2 Policy effect evaluation

Policy effect evaluation forms an important element in government actions, and more specifically in addressing complex problems such as section 2.1 shows. Figure 1 shows, however, that measuring these effects is not a simple task. This section will discuss different forms of policy effect evaluations in which different gradations can be made in the amount of evidence that a certain evaluation entails (provision of evidence).

2.2.1 Forms of policy effect evaluation

Evaluation ladder
In order to investigate whether policy has achieved the desired effects and whether that is due to the implemented policy, it may be clear that it must ensure facts. These facts can be disclosed by means of policy effect evaluations. In conducting these evaluations it is important that the results provide certain knowledge about the effects of the policy. In principle, this means that the effect evaluation provides insight into the variables of figure 1 and assesses the goals of the policy on basis of these variables. There are different forms of effect evaluation: from 'quick and dirty' to 'evidence-based'. Hoogerwerf and Herweijer (2014) refers to this distinction as: *intuitive* and *systematic* evaluations, respectively. According to them, intuitive evaluations are based on tacit (implicit) knowledge and common sense while systematic evaluations focus on scientific research methods. The dichotomy between intuitive and systematic can be viewed as a continuum (Hoogerwerf & Herweijer, 2014) where these different evaluations take a position on the line. The different evaluations designs can be divided into categories. The first category is to be regarded as (quasi-)experimental research designs. The second category is to be regarded as use-oriented and practical (between science and administrative practice) effect evaluations. In this study is referred to different evaluation designs which are described by van de Graaf and Hoppe (1996) and Korsten (2013). They show different forms where each form is associated with a certain degree of evidence it delivers (see figure 3 on the next page). These are both the most common evaluations and evaluations that deliver the most evidence. The gradation is based on the amount of evidence that a given evaluation design entails. This ranking is by Korsten (2013) referred to as the 'evaluation ladder' that consists of ten steps/stair-treads.

The evaluation ladder denotes that when one wants to have very precise knowledge about the effects of the policy, they should perform a 'real control group, pre- and post-test experiment'. This evaluation generates the highest degree of evidence regarding the effects of the policy. This is not only important for the accountability of the policy, but also plays an essential role in the learning ability of the government as section 2.2.2 will explain. However, this evaluation requires a lot of time, much expertise in the field of research methodologies and is very expensive. In practice, it is therefore not inconceivable that a government chooses a more straightforward research/evaluation. In addition, a government is also limited in conducting researches/experiments in which a control group is required. The Dutch legal system knows 'general principles of proper administration'. One of these principles is the 'equality principle'. This principle implies that similar cases should be treated similar and that governments should not make injustice distinctions (van den Berg, Meijer & Slager, 2010). This means that, for example, it is inappropriate for a municipality to submit a certain group of people to a specific policy while others are left out because they function as a control group. Though, the higher up the evaluation ladder, the more reliable the data from the evaluation and the more a government can learn from the evaluation. It also applies that: the higher up the ladder, the more complex, more expensive and time-consuming the evaluation.

Figure 3 – Evaluation ladder
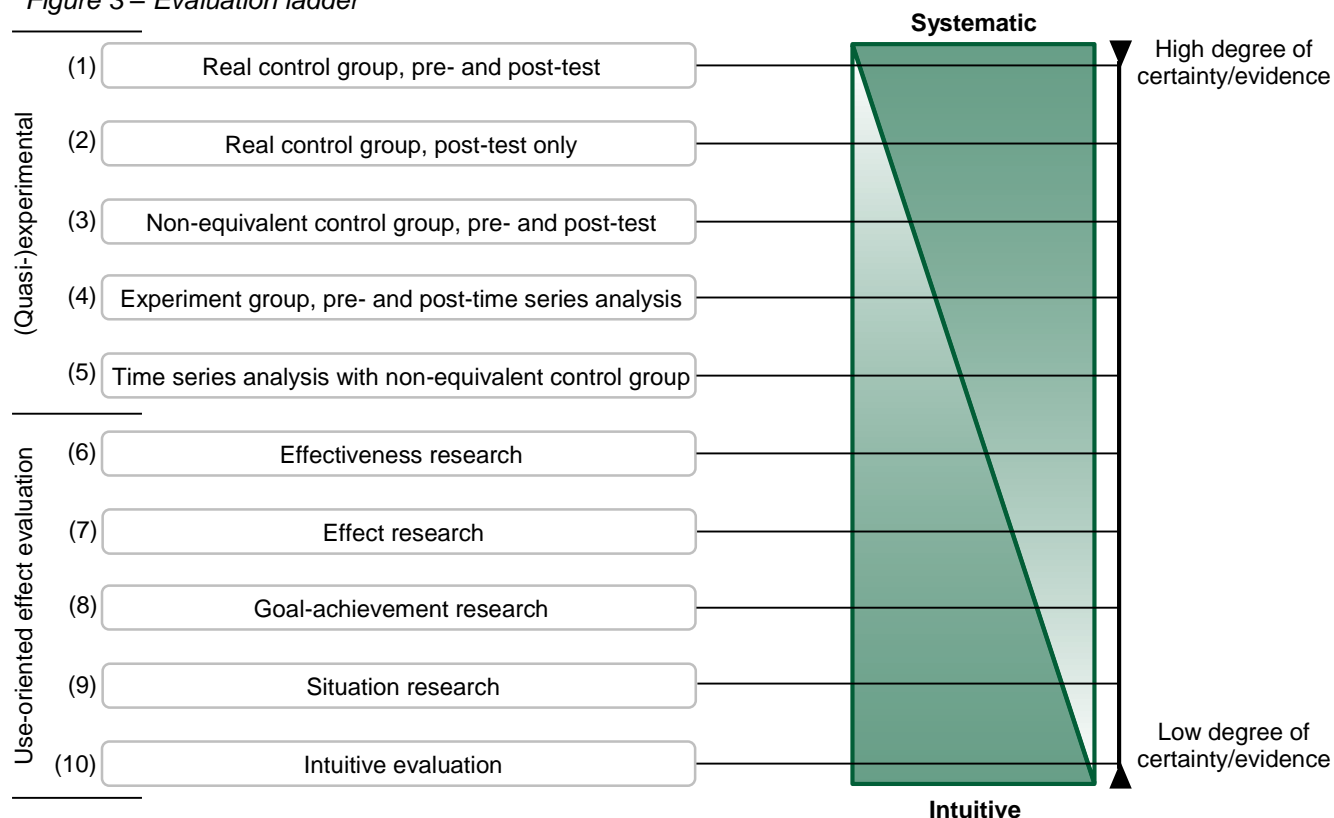
Policy evaluations can be distinguished in different moments of time. The distinction can be made between 'ex ante' (prior evaluation), 'ex durante' (evaluation during the policy implementation) and 'ex post' (evaluation afterwards) (van de Graaf & Hoppe, 1996). As already explained, this study focuses on ex post effect evaluations. Ex post evaluations require explicit goals. It evaluates whether intended goals and/or intended effects are realised and why the established goals are or are not achieved (Bekkers, 2012). This also means that prior knowledge about the intended goal is required in order to make statements ex post. In other words, ex post evaluations are often associated with an ex ante and/or ex durante evaluation. There lurks a major danger in this process, namely 'goal shift'. When a goal is modified, it means that the baseline measurement is no longer applicable, in other words, ex post results cannot be mirrored on the ex ante results (Iedema & Wiebinga, 2010). This means that in advance the goals must be clearly and thoughtfully written down. Table 1 provides an overview of the distinctive criteria of the evaluations that can be used as checklist for ranking effect evaluations.

It is worthwhile to mention that there are more types of evaluations; focusing on process, content or subjective understanding (see Korsten, 2013; Bekkers, 2012; Hoogerwerf & Herweijer, 2014), but these evaluations are not considered in this study as this research focuses on the effects of policy. Nevertheless, this study does not underestimate the importance of these evaluations. Hermeneutical evaluations are able to analyse social processes in contrast to effect evaluations (Blommestein, Bressers & Hoogerwerf, 1984). Korsten (2013) supports hermeneutical evaluations since governments nowadays rely on other governments and private parties and thus get involved in a network context. Though, because hermeneutical evaluations focus primarily on social processes, the power of effect evaluations is that they can provide evidence of intended goals. In practice, it may therefore occur that policy evaluations focus on both aspects: social processes and policy effects.

Table 1 – Criteria in relation with position on evaluation ladder

Type of evaluation	Features / criteria
Real control group, pre- and post-test	Experimental evaluation form in which a pre- and post-test is applied and which groups are randomly assembled. This set-up includes both an experiment group as a control group. R O X O R O O
Real control group, post-test only	Experimental evaluation form in which only a post-test is applied and whose groups are randomly assembled. This set-up includes both an experiment group as a control group. R X O R O
Non-equivalent control group pre- and post-test	Experiment evaluation form that is very similar to the 'real control group, pre- and post-test'. The only difference is that this evaluation form the groups are not randomised. O X O O O
Experiment group, pre- and post-time series analysis	Experimental evaluation form which includes a measurement at different time intervals. This involves several pre-tests and post-tests in order to make trends an autonomous developments visible. O O O X O O O
Time series analysis with non-equivalent control group	Experimental evaluation form which is similar to the 'experiment group, pre- and post-test time series analysis'. The only difference is that this form also includes a control group next to the experimental group. O O O X O O O O O O O O O
Effectiveness research	Effectiveness research that combines some strengths of effect- and goal-achievement research. As an attempt is made, as in the effect research, to find out the influence of the instrument empirically. In addition, as with goal-achievement research, explicitly the link is made between the goals of the policy.
Effect research	Disadvantage of both the situation research and the goal-achievement research is the uncertainty about the contribution of the instrument to the observed change. In effect research this is encountered by adapting the 'with-without' approach (comparing the situation 'with' the use of the instrument and the situation 'without' the use of the instrument).
Goal-achievement research	Establishing a relationship between what is pursued with an instrument (goal) and what has been achieved.
Situation research	Primary mapping changes made in the policy area in which no specific goals are expectations are established.
Intuitive research	Evaluation based on tacit (implicit) knowledge and common sense.

*Based on: van de Graaf and Hoppe (1996)*

2.2.2 Policy learning

Evaluation is learning! This means that evaluations are not only made to be accountable, but that they are drawn up utilisation-based. In short, policy evaluation is a learning process in itself. But what is then meant by policy learning? Bekkers (2012) defines it as "…a process in which actors endeavour to improve their actions on the basis of knowledge, experience and insight" (p. 304). In other words, it concerns a process of cognitive development and adaptation in order to improve the operation of an organisation in a systematic way. Policy learning can thus contribute in shifting the boundaries of 'bounded rationality'. Policy learning can therefore result in an increase of knowledge of the problem, increase of knowledge on related mechanisms, but perhaps more importantly it can relate to feedback loops of effectivity, perceptions of external events and perceptions of the consequences of alternative policies. In recent years, more attention is paid to perform policy evaluations utilisation-focused. This means that

people do not only use policy evaluations to see whether policies have or have not worked out. Utilisation-focused evaluations involve designing and the organisation of evaluation research in such a way that its results could play a role in the judgment and decision-making of a responsible policymaker or political body (van de Graaf & Hoppe, 1996). The aim of utilisation-focused evaluations is also to gain insight into the results of the actions of policymakers in order to intervene if necessary. The basic principle is that policy evaluations should be judged on their usefulness and actual use, as Patton (2008) stresses: "Therefore, the focus on utilization-focused evaluation is on intended use by intended users" (p. 425). Policy effect evaluations can play a vital role in the learning capacity of the government. Hoogerwerf and Herweijer (2014) describe this process of learning on behalf of three phases: cognisance of the findings, affecting standpoints and impact of the findings. The essence of this phasing is that a first important step is to communicate the findings of the evaluation. What has come about and what does it mean? This should not be limited to the directly involved persons, but must be shared in a broader context. This may influence on more fronts the standpoints of persons and gain its impact in the formulation of policies in the future.

Several scholars (Argysis, 2004; Papadopoulos & Warrin, 2007; Hoogerwerf & Herweijer, 2014; Sanderson, 2002) emphasise the importance of interactions and perceptions as a crucial element of policy evaluations in the learning process and in decision-making. The reflexivity of society is forcing organisations to be more reflexive about the impacts of their actions and policies (Argysis, 2004). "Effective learning is an important cause of effective action by individuals, groups, intergroups and organisation" (Argysis, 2004, p. 507). These actions should carefully get attention in drafting and communicating policy evaluations. Gerrichhauzen, Korsten and Fijen (2002) write that the management of an organisation obviously has the central task to stimulate learning in an organisation, but that the organisational culture offers little opportunities to do so. In this learning process are shared meanings, norms, values and ideologies important in how an organisation operates (Gerrichhauzen, Korsten & Fijen, 2002). Policy evaluation could change these shared meanings, norms and values by focusing on the three phases of Hoogerwerf and Herweijer (2014). In conclusion it can be said that the utilisation-focused policy effect evaluations can influence the actions of an organisation. There are several ways in which an organisation can learn. Argyris and Schön (1974) identify different 'learning loops' (figure 4).

Figure 4 – Learning loops
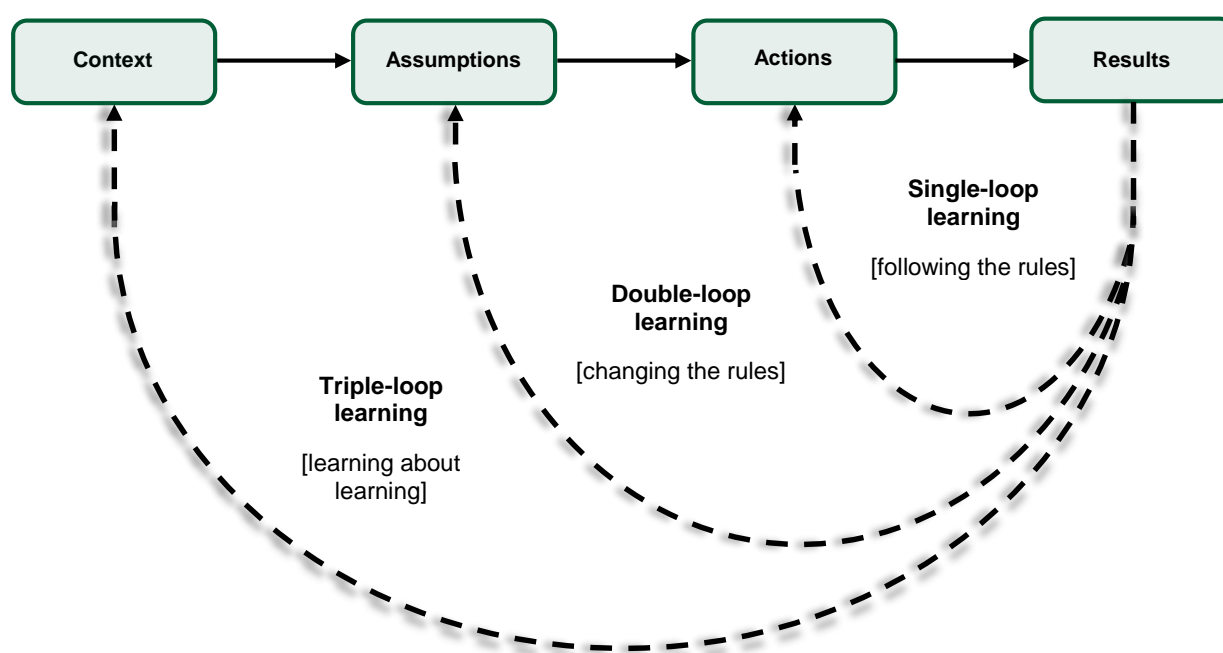
Single-loop learning
This learning loop mainly involves a comparison of the results of the policies and goals of the policy (Bekkers, 2012). This learning model is the most common model of action (Argyris, 1976). Figure 4 shows that the feedback loop focuses only on the actions and that these actions are adjusted during the learning process to better control the situation. It is therefore a learning process that focuses on the improvement of the interventions in order to achieve the goals which the organisation has in mind (Bekkers, 2012). In this mechanism occurs little to no learning because strategies, policies and procedures are established in which the policy makers spend time to identify and solve deviances (Argyris & Schön, 1974). This learning loop therefore focuses on 'following rules'. This process is by Lindblom (1979) defined as 'incrementalism'; a step-by-step better solving procedure in complex problems.

Double-loop learning
In double-loop learning the learning process goes further than following the rules. Here the assumptions (or rules) of the policy are challenged. Thus this feedback mechanism puts the focus on 'changing rules'. Argyris and Schön (1974) describe this process as reflecting on rules to draw conclusions about whether the 'rules' should be changed. With rules one is referring to the assumptions underlying the policies. In designing policies, it may be that policies are based on incorrect information or if new information leads to a different definition of the problem (Bekkers, 2012). In such as case, it is important to pull the assumptions in doubt in order to adjust if necessary. This requires 'out of the box' thinking.

Triple-loop learning
The triple-loop learning concerns 'learning to learn'. In this loop the context is taken into account. This learning process is not clearly stated in the published work of Argyris and Schön, but implicitly it is described by them as 'deutero-learning'. Although the fact that various descriptions exist of this learning process (Tosey, Visser & Saunders, 2011), Gerrichhauzen, Korsten and Fijen (2002, p. 4) refers to triple-loop learning "when participants of an organisation also learn about the context of learning" to be able to independently execute tasks. This learning process is of great importance when an organisation is in a rapidly changing environment.

Bekkers (2012) outlines that this system of feedback mechanisms can be described as 'cybernetic learning'. Such a learning process takes place within the governing body that monitors through feedback mechanisms what the contribution of executive organisations is on the intended results of a policy. Bekkers (2012) points out that this feedback mechanism indicates errors and problems and allows policy makers to undertake corrective actions to achieve the goals of the policy. Recognising the goals and the context in which they are realised is an important element in order to act adequately (Gerrichhauzen, Korsten, Fijen, 2002). Essential in the learning ability of a government is that the learning process does not only take place on the basis of single-loop learning in which the maintenance of processes is ensured through individual actions (solutions), but that one also takes into account the context and assumptions and thus on behalf of collective reflection approaches problems (Gerrichhauzen, Korsten & Fijen, 2002).

2.3 Theoretical conclusion

The theoretical framework sketches a complex picture of society in which the government, in solving problems, is encountered by uncertainty and risks. This complexity is the result of a historical development in which industrialisation, rationalisation and interventionism plays a central role. In a time of reflexivity, citizen participation and technological developments it is inevitable that government actions are influenced by external factors. This makes the establishment of certain performances not an easy task let alone the evaluation of it. Because society is so complex and because the government is faced with uncertainties, policy focused on effect measurement becomes an essential component of conducting policies. These policy evaluations are able to remove or reduce uncertainties and identify the reasons for policies being a success or failure. This is not only important for the legitimacy of the government and the spending of common money, but is also crucial for the learning ability of its organisation.

The municipality of Wierden indicates that they are doing too little effect evaluations. The theory teaches that in a thorough evaluation of policy effects many aspects of society can play a role. This requires highly specialised knowledge and expertise in research methods and in science. Because the municipality of Wierden is a relatively small rural municipality, it may therefore be expected that less attention is paid to the complexity and uncertainty of society in the implementation and evaluation of policies. This leads to the first expectation that this study draws.

Expectation 1: Municipality of Wierden expresses little understanding of the complexity and uncertainty of policy effects.

Because it is expected that little attention is paid to the complexity and uncertainty of policy effects, it can also be predicted that the policy evaluations, focusing on policy effects, do not take into account the influence of external factors and insight into these factors. This is a very important part in policy effect evaluations. When this is done thoroughly, it is possible to indicate to what extent policies have worked on the basis of definite evidence. However, in the case of the municipality of Wierden it is expected that more intuitive evaluations occur. This brings this study to the second expectation which is closely related to the first expectation.

Expectation 2: Municipality of Wierden scores low on the 'evaluation ladder'.

Besides the de facto policy implementations of the municipality of Wierden there are also some de jure obligations. These obligations focus on laws of other governments and focus on the provisions established by the council and/or college of the municipality of Wierden itself. It is expected that these provisions are being respected as they are legally required and because the municipality of Wierden considers them important. The third expectation focuses on those provisions.

Expectation 3: The de jure mandatory actions aimed at policy effect evaluations are respected by the municipality of Wierden.

As stated earlier, it is expected that the municipality of Wierden will score low on the evaluation ladder. If this is the case, it means that the evaluations have little certain knowledge to draw conclusions about the actual effects of policies. In this situation, it may be expected that the learning capacity is minimal. This brings the study to its final expectation.

Expectation 4: The (organisational) learning ability of the municipality of Wierden in policy effect evaluations will primarily take place on the basis of single-loop learning.

The expectations are in some instances in line with each other, but focus individually on a separately important aspect of policy effect evaluations. The expectations are tested against the research results in chapter 4 in order to conclude whether the expectations are true or if they could be rejected. Before this data and the analyses are presented, the focus is laid on the methodology of the research in order to indicate which methods are used in collecting and analysing data.

3. METHODOLOGY

This chapter describes in which way the research will be conducted. For this, a research design is set up which answers the question: What is the design? Furthermore, it provides a closer look at the adopted strategy together with the units of analysis answering the question: What is investigated in this design? Finally, attention is paid to data collection and data analysis methods. This eventually answers the question: How is it going to be investigated?

3.1 Research Design

3.1.1 Design model
The research question will be answered by conducting a qualitative research strategy. This approach is therewith open and flexible, moreover, it examines the backgrounds of the data collected (Verhoeven, 2010); which is the purpose of the research. The qualitative research can be characterised as a 'single case study'. In single case studies "usually one 'thing' is studied in detail" (Johnson & Reynolds, 2008, p. 149). In this research that will be the policy effect evaluation of the municipality of Wierden. In order to get a complete picture of the way in which the municipality of Wierden carries out policy effect evaluations and in order to understand the philosophy that they adopt, it is important to examine different departments, administrative units and managers. Thus, within this 'case' more elements (units of analysis) can be distinguished which are going to be scrutinised. These 'units of analysis' together contribute in understanding the situation regarding policy effect evaluation at the municipality of Wierden. Figure 5 shows the research model in which the research design is depicted together with the different stages.
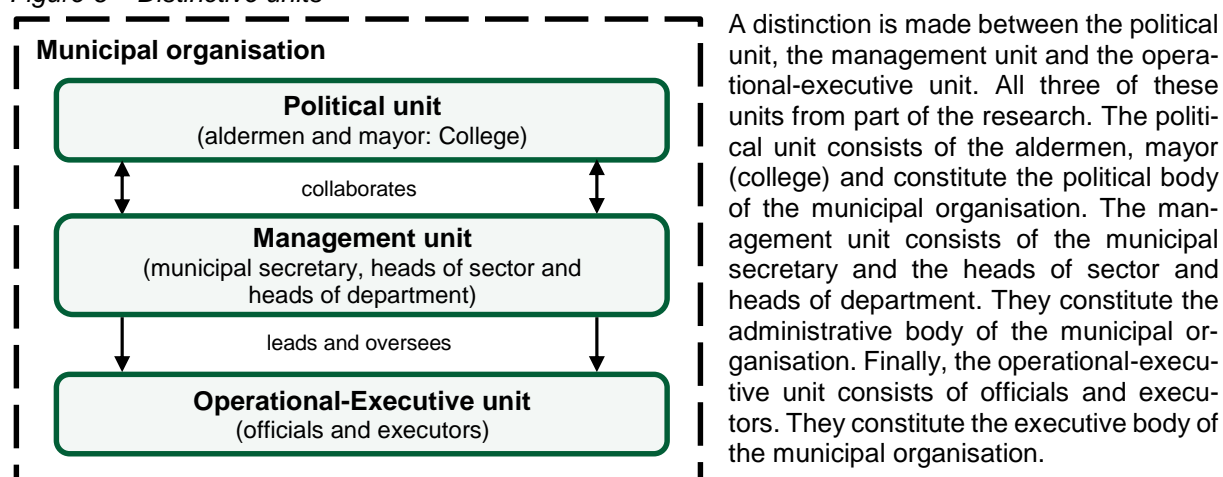
Figure 5 – Research model
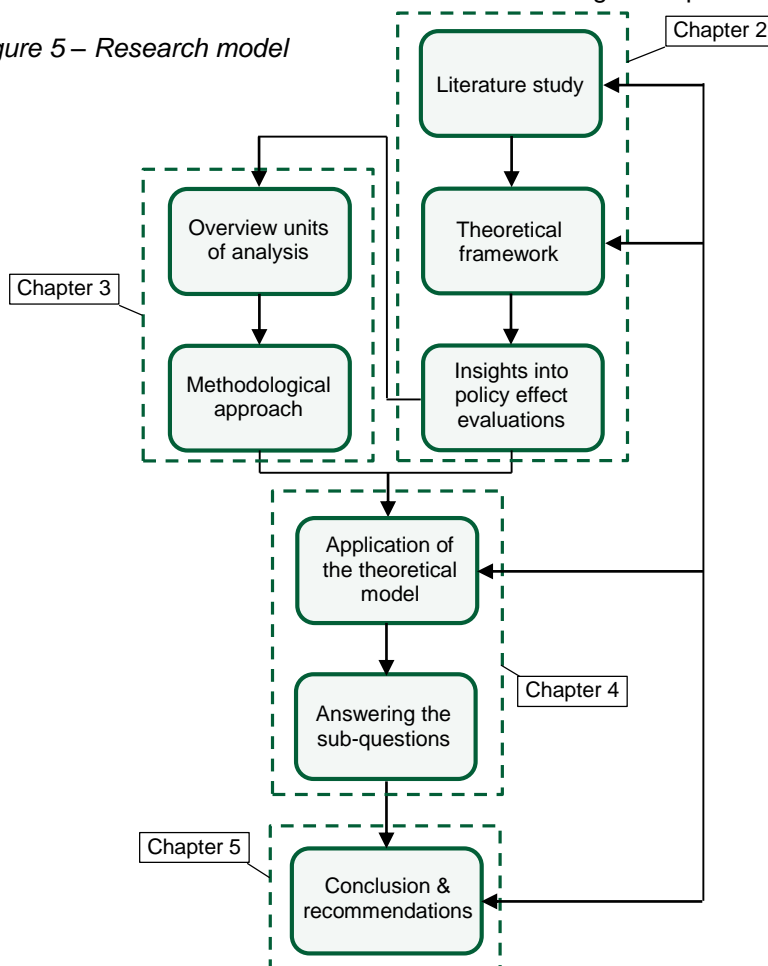
An important feature in which this research is different from other case studies is that the research is not aimed at testing or (re)developing theories (de Vaus, 2001). This research uses theories in order to understand the case under consideration: policy effect evaluation of the municipality of Wierden. The choice for this design is based on the fact that the goal of the research is to obtain insights about the policy effect evaluation at the municipality of Wierden and to solve the problem for the case (de Vaus, 2001). In this context, the research focuses on creating awareness and on providing new insights regarding the case. The design of the research may partly also be typified as 'descriptive' because it describes the case and puts it in the context of the theory.

On the basis of the literature study, a theoretical framework is designed. In this theoretical framework an attempt was made to create a clear picture of the different policy effect evaluations and the applications of the evaluations together with their backgrounds and intentions. This forms the first phase of the study and consists mainly of desk research. The second phase in the study contains the methodology indicating the units of analysis and the strategy in conducting the research. The third phase focuses on the case itself. In order to draw conclusions, several parts of the organisation need to be analysed. By mirroring the theoretical framework on the findings of the analysis, a conclusion can be formulated showing the attitude of the municipality of Wierden concerning policy effect evaluations. Based on this descriptive analysis, suggestions and recommendations can be made. Since the research focuses on understanding the case, the theory plays a central role in this research. This makes the process of the research an iterative process, as figure 5 indicates.

3.1.2 Research strategy

In the previous section the design of the research is explained as well as the phases that can be distinguished. This section elaborates on the parts to be examined, in other words, the units of analysis. As indicated above, there are several parts within the municipal organisation that is going to be investigated. This means that a focus is to be applied. As de Vaus (2001) points out: "To describe everything is impossible: there must be a focus" (p. 225). A distinction is made between three units within the organisation, see figure 6.

Figure 6 – Distinctive units

A distinction is made between the political unit, the management unit and the operational-executive unit. All three of these units from part of the research. The political unit consists of the aldermen, mayor (college) and constitute the political body of the municipal organisation. The management unit consists of the municipal secretary and the heads of sector and heads of department. They constitute the administrative body of the municipal organisation. Finally, the operational-executive unit consists of officials and executors. They constitute the executive body of the municipal organisation.

The differentiation in three units is made because it can be expected that different assumptions/motives in evaluating policies exist between the units. Appendix 1 provides the organogram of the municipality of Wierden in which these three distinctive units are indicated. The three units are analysed with combined methods as the following section will indicate.

3.2 Data collection

Qualitative research focuses on the subjective signification or the meaning given to the situation around policy effect evaluation at the municipality of Wierden (Verhoeven, 2010). In order to be able to find relevant data and to analyse it, is chosen for a combined methodology: document analysis and interviews. Qualitative research can be divided by several steps: (1) gaining entry; (2) category and observation; (3) data recording; and (4) analysis (Dooley, 2009). The first step is self-evident; internally there is access to documents and relevant persons for interviews. Step 2 is made up of the units of analysis that have been distinguished. The data record of step 3, is collected by means of the methods discussed in this section, after which they are analysed (step 4).

3.2.1 Document analysis

To be able to find relevant information regarding policy effect evaluation at the municipality of Wierden, several documents are examined. This method of data collection and data analysis is called 'document analysis' and is an empirical observation which makes use of written record. Johnson and Reynolds (2008) note that in order to use the written record, it is important to gain access to the material in order analyse them. The municipality of Wierden provides access to this information and helps in the search for relevant material. Because policy effect evaluation can be found in many documents, there are many different documents that form part of this document analysis. These documents are named '*running record*' because they are produced by an organisation, the municipality of Wierden, because they are easily accessible and because they are available for long periods of time (Johnson & Reynolds, 2008). The list of all consulted documents for analysing the situation regarding policy effect evaluations is provided in appendix 2. Also, the documents are included in the bibliography.

A focus is provided in the analysis of the documents. This focus is the coalition agreement of 2010-2014 together with all the related (policy) documents from 2010 till 2015. The document analysis results in an overview of the 'de facto' performed effect evaluations, the ambitions for evaluations, types of policy effect evaluations carried out and the nature and scope of the evaluations. It also offers insight into the ipso jure necessary evaluations. The data (information) which is gained form the documents are used in two ways. First, the information is used to describe what the municipality actually does concerning policy effect evaluation as well as what de jure is necessary. This provides an overview of the state of affairs of the empirical policy evaluation of the municipality of Wierden. This means that in providing an overview, use is made of quantitative data (e.g. number of evaluations, evaluation numbers by sector, etc.). Furthermore, the document analysis is used as a basis for the interviews. The information about certain documents, provide insight into the way the municipality makes use of policy effect evaluation. In this way backgrounds of certain choices, events or the lack of it in these documents will be elucidated by means of the performing interviews.

3.2.2 In-depth interviews

Interviewing is a method of data collection in which individuals are subjected to a series of questions which are then recorded, written or digital (Johnson & Reynolds, 2008). Only performing a document analysis is not sufficient, therefore, interviews also form part of this research. In this research, interviews are used for two reasons. Firstly it is used, as specified above, to place the findings from the document analysis in a context to better understand the situation regarding policy effect evaluation at the municipality of Wierden. This allows to find explanations for certain choices, events or the lack of it because the document analysis does not provide sufficient information about this. The second reason for using interviews is that it examines the idea or thought (philosophy) of the municipality regarding policy effect evaluations. Through specific questions can be studied in depth how the municipality of Wierden approaches policy. The interviews are face-to-face taken in an open-ended discussion. The reason for face-to-face interviews is that all relevant persons are located in the municipal organisation and are accessible. The foregoing calls for 'in-depth interviews'. This type of interview "gives the interviewer a chance to probe, to clarify, to search for deeper meanings, to explore unanticipated responses, and to access intangibles such as mood and opinion intensity" (Johnson & Reynolds, p. 338). Preparation is of

great importance in this type of interview. The document analysis takes care of this preparation because it serves many purposes (Johnson & Reynolds, 2008, p. 340-341):
1. It saves the interviewee's time by eliminating questions that can be answered elsewhere.
2. It gives the researcher a basis for deciding what questions to ask and in what order.
3. Advance preparation helps the researcher to interpret and understand the significance of what is being said, to recognise a remark that shed new light on a topic.
4. The researcher's serious interest in the topic impresses the interviewee.

In order to use the obtained information during the interviews in a reliable way, the choice is made to record the interviews (tape recording). This method of 'making minutes' takes away the risk that afterwards certain views of the interviewee are written incorrectly. In addition, an important advantage of this 'method of recording' is that it allows the researcher to think during the interview about what the interviewee says, to make notes and to formulate follow-up questions (Johnson & Reynolds, 2008). The recorded interviews are than transcribed. The interviews will be conducted within the various units of analysis. The different units of analysis each have their own character and may differ in the way they perceive policy evaluations. The distinction between the units of analysis and the interviewees within these units are provided in a list in appendix 2.

3.3 Data analysis

3.3.1 Analytical model
In qualitative research it is about the meaning given by people to certain situations and behaviours (Verhoeven, 2010). Analysing the collected data is a step-wise implementation. In this implementation, the two methods (document analysis and interviews) are combined. The following model provides an overview of the analysis. This data analysis contains of various steps, explained afterwards.

Figure 7 – Analytical model
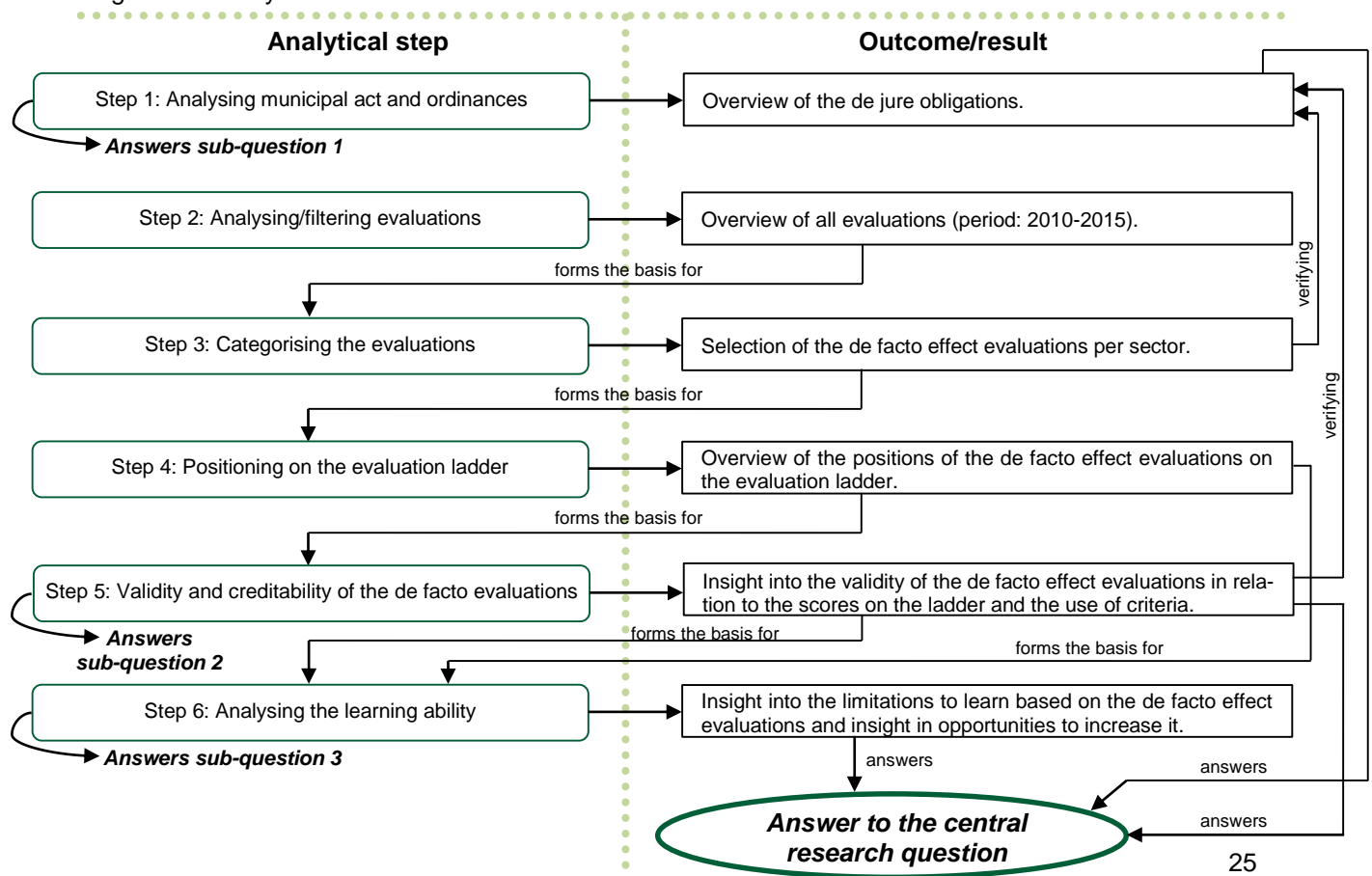
3.3.2 Elaboration of the analytical steps

The central research question in this research is: *How does the municipality of Wierden perform policy effect evaluations and how can the learning ability of the organisation be increased based on the effect-measuring performance?* In order to answer this question different sub-questions are established. Answers to the sub-questions are provided by following different analytical steps. This sections discusses the steps from the model in figure 7.The steps are divided into three analytical fields. These fields correspond to the paragraphing of the results and involve: de jure policy effect evaluations, de facto policy effect evaluation and learning ability.

Analytical field: de jure policy effect evaluation

This first analytical field answers the first sub-question: What are the de jure obligations for the municipality of Wierden regarding policy effect evaluations? Answer to this question creates an image of the obligations and codified ambitions regarding policy effect evaluation of the municipality of Wierden. This is important to understand their obligations concerning effect evaluation and to gain insight into the way the municipality of Wierden complies with these obligations.

Step 1: Analysing municipal act and ordinances

A municipality is not free in organising and implementing its organisation (Iedema & Wiebinga, 2010). There are several components that are set down in the municipal act and in ordinances. In this step, the municipal act and the different ordinances of the municipality of Wierden are analysed. This analysis focuses on explicit wordings about the effects of policy and policy evaluation. Based on the theory, the concept of 'effect evaluation' is operationalised resulting in keywords that are closely related to policy effect evaluation (the assessment of the effects perceived on the basis of criteria). The table below shows this operationalisation. The keywords are in Dutch to explicitly indicate which terms are used in the analysis.

Table 2 – Operationalisation of effect evaluation

Keywords	Item
(beleids)effect	The effect that the implementation of policy has brought about.
beoogd effect	The effect that is intended to be achieved by means of the implementation of the policy in question.
doel(stelling)	The desired future situation where the policy is meant to focus on.
doeltreffendheid	The extent to which the measures taken in the policy have resulted in the realisation of the intended effects.
effectiviteit	The extent to which the measures taken in the policy have resulted in the realisation of the intended effects.
evaluatie	Assessment of the perceived effects of policies.
externe variabelen/effecten of neveneffecten	Variables and effects that are not part of the intended effects (see figure 1), but influenced the results of the policy.
maatschappelijke effecten	Effects of policy on society.

The documents are read in which is sought for provision/obligations regarding effect evaluation based on the keywords in table 2. These obligations will be presented in order to analyse in a later step whether the municipality of Wierden meet these obligations. The obligations can be checked against the results which are objective measurable. The subjective obligations are checked based on the theory and the interviews.

Analytical field: de facto policy effect evaluation

This field of analysis answers the second sub-question: *To what extent does the municipality of Wierden perform de facto policy effect evaluations?* The answer to this question creates an important image of how the municipality of Wierden performs effect evaluations (in the period of 2010-2015). This is important to be able to make statements about the performance regarding effect evaluations and to verify the de jure obligations. This field of analysis contains several steps.

Step 2: Analysing / filtering evaluations

The second step is to collect all evaluations from the period 2010-2015. In this step use is made of the DECOS-system that the municipality of Wierden has. This system is a document management system in which the municipality is able to work business case-oriented. This system has many search functions. These search functions are used to collect all evaluations over de period of 2010-2015. The system provides to search specifically (on date, search terms, etc.). In the search for evaluations the terms 'evaluatie' and 'onderzoek' will be used. In addition, the evaluations that are found will be approved by employees of the municipality in order to include evaluations that are still missing.

Step 3: Categorising the evaluations

The list of evaluations concern all evaluations undertaken in de period of 2010-2015. In this step, the evaluations are categorised on 'effect evaluation' or 'other evaluation'. The implementation of this step is done by analysing the evaluations based on the keywords in table 2. Based on these keywords can be determined if the evaluation focuses on effectiveness or if it concerns another evaluation. This results in the de facto effect evaluations of the municipality of Wierden. The DECOS-system provides insight to which organisational sectors the evaluations belong. On basis of this information it is possible to classify the de fact evaluations per sector. Additionally, in this analysis is checked whether the objective de jure obligations are met.

Step 4: Positioning on the evaluation ladder

To draw conclusions about the validity of the de facto effect evaluations are these evaluations positioned on the evaluation ladder. Thereby the theory will be explicitly used. Table 1 presents the various forms of effect evaluations and the features/criteria. Based on these features/criteria the de facto effect evaluations are tested resulting in the granting of evaluations forms to the de fact effect evaluations. This leads to the position of the de fact effect evaluations on the evaluation ladder. This step will thoroughly be looked at the parts that are missing in the de facto effect evaluations and what consequences this has for the validity and reliability of the evaluations. This is done by making use of figure 1 and the different theories to draw conclusions about the validity and consequences of the de facto effect evaluations.

Step 5: Validity and creditability of the de facto evaluations

This step explains in more detail the validity of the evaluations on basis of the results from the fourth step. In this step, the de facto evaluations are tested on the interrelationship of the criteria. The definition of effect evaluations used in this research endorses that an effect evaluation is the assessment of the perceived effects on behalf of criteria. Therefore an effect evaluation requires criteria. The de facto effect evaluations will be studied in which is examined whether there is an explicit criterion on which the evaluation focuses. When a criterion is present is checked whether any underlying criteria/indicators have been associated with the criterion. In checking the evaluations the focus lies on explicit notifications of criteria and the relations between the criterion and criteria. The results are mirrored on the theory in order to draw conclusions about the validity of the de facto evaluations. In this analysis is also examined whether it is consistent with the objective de jure obligations.

Analytical field: learning ability

The last field of analysis answers the third sub-question: *How can the learning ability for the municipality of Wierden be improved regarding policy effect evaluations?* The answer to this question together with the other sub-questions complete the answer to the central research question (see figure 7). The validity of the de facto effect evaluations are an important basis for het learning ability of the municipal organisation. It is therefore important to examine to what extent the municipal organisation has the ability to learn in order to make statements about how they can improve this learning ability.

Step 6: Analysing the learning ability

In this step, the results from the previous steps will be analysed by means of the theory about the learning ability. The theories about the learning ability provides insight into how the results can be used based on the three learning loops. Through insight into the validity of de facto effect evaluations (in

previous steps), statements can be made about the learning ability of the municipal organisation by focusing on the three learning loops. Improvements can be proposed based on the findings of the analysis in order to increase the learning ability.

3.3.3 The use of the interviews in the analysis

The interviews are used to extent the finding from analyses by analysing opinions, experiences and insights from employees. In interviewing the employees takes place on behalf of the units of analysis. From each unit, different employees are interviewed with different functions (see 3.1.2). The interviews audiotaped and transcribed. The interviews are analysed based on a coding system. The codes are presented in table 3.

Table 3 – Codes analysing interviews

Code	Explanation
departments	Answers that address the differences in departments in relation to effect evaluations (numbers, reasons for differences, management steering, etc.).
explicit goals	Answers relating to the formulation and inclusion of goals in policy and policy evaluations (SMART formulation of objectives, explicit goals, etc.).
good/wrong	Answers that address in general terms the strengths and weaknesses of the municipality of Wierden regarding policy effect evaluations.
policy effects	Answers that focus on the effects of policies and the mapping of these effects. Complexity and uncertainty of policy is also part of this code.
responsibility	Answers that focus on the responsibilities of employees of the municipality of Wierden regarding policy effect evaluations.
satisfaction	Answers that focus on opinions, experiences and insights of goals that focus on satisfaction of citizens and the relationship between satisfaction and underlying policies.
sufficient	Answers that focus on the number of effect evaluations that is being performed and opinions about the amount of effect evaluations and effect measurements.

These codes are chosen on the basis of the theory and complemented with the findings of the analysis. The codes are a reflection of the parts that need more clarification in terms of the findings. The analyses and coding can therefore be seen as a complement and substantiation/explanation of the findings of the analytical steps and provide a picture of how employees of the municipality of Wierden look at certain aspects. After coding the interviews, the coded answers are categorised per unit of analysis and by code. This data is then used in the results.

3.4 Reliability

Reliability concerns the degree to which the obtained data, by means of data-collection methods, are reliable. The research data must be repeatable, that is a prerequisite for the reliability of a study (Baarda, 2012). This equates to the replicability of the research. In other words, during the study you rule out accidental mistakes. To increase the reliability of this research, use is made of combined methodology, namely: document analysis and in-depth interviews. The two methods are used in line with each other in order to minimise the occurrence of random errors (interpretation errors). In qualitative research with interviews, it is difficult to find consistent findings as interviews are based on the perceptions and experiences of the interviewees. For that reason also documents are analysed in order to not only offer perceptions and experiences, but also to be able to analyse procedures, results and facts (empirically). The interviews will be audio taped and transcribed. In this manner, the reliability of the research is ensured.

3.5 Validity

Reliability does not guarantee validity: if the reliability is good, it does not automatically mean that the validity is also good (Baarda, 2012). Reliability is however a precondition for validity. What, then, is validity? Validity means that you measure what you are intended to measure (legitimate and purity) as Verhoeven (2010) defines it. Validity is therefore a very important aspect for the generalisation of the research results. However, the generalisability of a case study is difficult because the case has a specific set of characteristics and will have a specific outcome (de Vaus, 2001). Nevertheless, this research has a certain value in wider generalisation, this is after all, the purpose of social science as de Vaus (2001) emphasises: "The point of a case study would be to see if a theory actually worked in a real life situation" (p. 239). In contribution of this research in 'if theories worked in real life' is that the practice is mirrored on the theories resulting in conclusions and recommendations. The validity of the research will be guaranteed by conducting a variety of interviews in which the questions are partly based on the factual findings of the document analysis. In addition, in preparation for the interviews, an introducing text will be sent to the interviewees in which the inducement, purpose, context and quality assurance is specified. In the latter a number points of attention are described to increase the quality of the interviews. Points of attention such as excluding politically-desired responses and to exclude conversations between people already interviewed and people yet to be interviewed. This ensures that already interviewed persons will not influence persons yet to be interviewed. The transcribed interviews will not be divulged and will only be submitted for the replicability of the research (reliability) in order to exclude politically-desired answers.

4. RESEARCH RESULTS

The objective of this chapter is to analyse to what extent the municipality of Wierden is making use of policy effect evaluations in relation with the de jure obligations and de ambitions they intend. This chapter answers the various sub-questions. First, attention is drawn to what the municipality of Wierden de jure ought to do regarding policy effect evaluation. The second section discusses the de facto effect evaluations. In this section attention is paid to the amount of evaluations and the validity of the effect evaluations. The final section discusses the learning ability of the municipality of Wierden regarding effect evaluations.

4.1 De jure policy effect evaluation

Municipalities are not totally free in setting and implementing their organisation. Some parts of the organisation of a municipality are laid down in the Constitution and the Municipal Act (Iedema & Wiebinga, 2010). The Municipal Act, chapter 14, provides the administration and control of municipalities. In Article 213a of this act, attention is given to the efficiency and effectiveness of the by the municipality performed management. About the following main provisions are shown with respect to policy evaluation as effect-measurement (Gemeentewet (n.d.), retrieved from http/wetten.overheid.nl/):
- College need to periodically examine the efficiency and effectiveness of its management;
- The municipal council shall by ordinance compose regulation about this.

By this act, municipalities are forced to pay attention to effectiveness. However, this act is not concrete and leaves much room for interpretation for both the municipal council and the officialdom. One can have different interpretations of what is meant by 'performed management' and on how effects should be measured. Despite the name 'effectiveness research', it is not possible to put (on forehand) the research on the evaluation ladder based on the presented information. This should become evident in the de facto evaluations in the next section (section 4.2). The ordinance, on the basis of Article 213a of the Municipal Act, was drawn up by the municipal council of Wierden (entry into force: 07-03-2004) and imposes further requirements on the execution of Article 213a of the Municipal Act. This 'Ordinance for periodic research by the college on the efficiency and effectiveness of the by the college performed management of the municipality of Wierden' contains the next important provision (Gemeenteraad Wierden, 2004, p. 1):
- The college assesses at least once a council period the programmes on effectiveness (it may also be paragraphs or other levels of budget or annual documents or parts thereof).

The ordinance defines 'effectiveness' as: "the extent to which the performance and the intended social effects of policies are actually being achieved" (Gemeenteraad Wierden, 2004, p. 1). This comprehensive elaboration of Article 213a reveals an important evaluation aspect: the frequency indicates how often the municipality must carry out such an evaluation.

In addition to the Municipal Act and ordinance regarding efficiency and effectiveness, the municipality of Wierden also has a document in which, to some extent, policy evaluation is legally framed. This legal document is the 'Organisatiebesluit 2009' (entry in force: 01-07-2007). This document contains, among other things, the method of process and the structure of the municipality as well as the tasks and responsibilities of the municipal secretary. The municipality of Wierden included in this document that the method of process is characterised by working with 'explicit goals' and by a 'systematic approach in order to optimise goal-orientation, efficiency, effectiveness and quality of process and product' (Gemeente Wierden, 2007). Departments are responsible for the periodic evaluation of activities (Gemeente Wierden, 2007, p. 2). This means that this 'Organisatiebesluit 2009' claims that policy that is located lower in the policy hierarchy should also be evaluated. Although no interpretation/definition is given to the term 'periodically', it seems that a department (or part of it) can interpret this themselves. Additionally, the 'Organisatiebesluit 2009' also describes the tasks of the municipal secretary. As head of the municipal employ, the municipal secretary is responsible for, among other things, the quality (efficiency

and effectiveness) of the functioning of the administrative organisation (Gemeente Wierden, 2007). This is an important implication. This means that the municipal secretary (indirectly) is responsible for the method of process of the organisation as mentioned above.

The de jure effect evaluations are a reflection of the mandatory evaluations and the codified ambitions of the municipality of Wierden regarding effect evaluation. These de jure effect evaluations give no univocal and clear picture of how these evaluations need to be carried out. It is also not specifically clear how the effectiveness of the organisation is ensured because there seems to be much room for interpretation. The interviews clearly show that there are many different opinions in all units of analysis regarding the implementation of policy evaluations and the way this is done properly along the different units of analysis. In order to see whether the municipality of Wierden complies with the de jure obligations, the following checklist is made.

☑ Municipality draws up regulation about efficiency and effectiveness via ordinance.

❑ At least once a council period the municipality checks on effectiveness.

❑ Working with explicit objectives.

❑ Optimisation of goal-orientation, efficiency, effectiveness and quality of process and product.

Based on the notification of de jure policy effect evaluations, only the first criteria can be checked. The other criteria should be checked in the following section in which the focus is laid on the de facto effect evaluations.

4.2 De facto policy effect evaluation

4.2.1 Actually exported evaluations
The uncertainty and complexity of society has led to a more reflexive society. Reflexivity recognises the importance of reflection of effects and the check of actions on effectivity. In conversations with employees of the municipality of Wierden it appears that the municipality of Wierden has a pragmatic culture characteristic that is often implementing and evaluating policies intuitively. All units endorse that too few policy evaluations are performed. Besides the pragmatic culture characteristic of Wierden it also appears that the 'issues of the day' result in overshadowing the need for evaluation. One experiences that policy evaluation forms a weak part of the policy process at the municipality of Wierden. By means of a document analysis, research has been done on the number of evaluations conducted in the period 2010-2015. Figure 8 is the result of this analysis.

Figure 8 – Amount of effect evaluations (period 2010-2015)
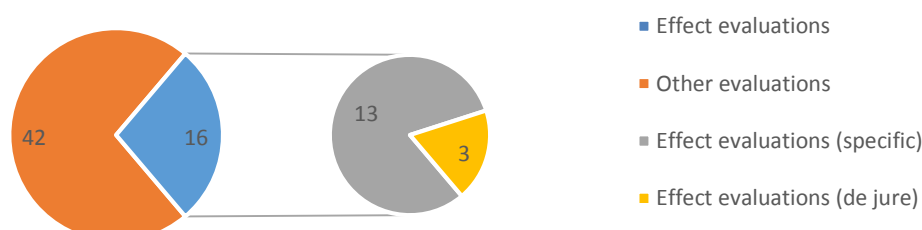
This chart shows that a total of 58 evaluations have been detected in the period 2010-2015. Of this number, there are 16 designated as effect evaluations (based on het formulated definition of policy effect evaluation). It is difficult to prove whether this number is low because no analysis has been conducted on the relationship between policies and policy evaluations. Despite the interesting information this analysis could deliver, it is not part of this study as it is a very comprehensive analysis which goes beyond the scope of this research. There are no objective data to draw conclusions about whether the number

of effect evaluations is little or much. One may assume that 16 effect evaluations in a period of five years is not much since it may be assumed that much more policy is implemented in that period. At the same time, 16 effect evaluations is relationally seen not much because there are conducted 58 evaluations in total. Figure 8 splits the effect evaluations in 'specific' and 'de jure'. The de jure effect evaluations correspond to the evaluations described in Article 213a of the Municipal Act. These evaluations cover more policy fields and are aimed at policy programmes, managing products or other annual reports. The specific effect evaluations are evaluations focused on a particular policy field and are not directly obligated by laws and/or regulations. Figure 9 shows the proportion of the number of effect evaluations in relation to the sectors. The organisational relationships and components (organogram) are shown in appendix 1.

Figure 9 – De facto effect evaluations by sector (period 2010-2015)
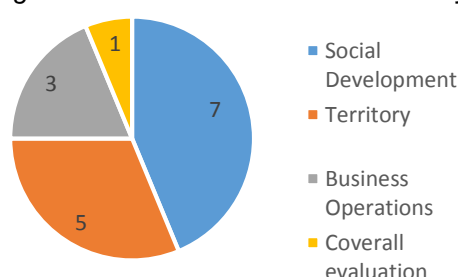
The differentiations are according to the different units attributed to the fact that the sector Social Development carries out the majority of policy tasks, and for that reason also performs more evaluations. Sector Territory and sector Business Operations have fewer operational tasks; for one it is more common sense than for the other. The 'coverall evaluation' is one of the de jure effect evaluations aimed at measuring the effectiveness of policy programmes. This is a municipality-wide evaluation and therefore it cannot be attributed to a specific sector. This evaluation (and partly the other de jure effect evaluations) ensures that the municipality of Wierden fulfils the second part of the de jure checklist because once a council period the municipality assesses the programmes on effectiveness (looking and the period 2010-2015). The screening of the amount of policy evaluations focusing on effect measurement and the relationship between the various sectors reflect the de facto policy evaluation carried out by the municipality of Wierden in de period 2010-2015. The low level of implementation shows that only a part of the policies are subjected to a policy effect evaluation. This means that the municipality of Wierden is not very reflexive in the actions they carry out. Despite the major consequences in failing to perform effect evaluations (as will become clear later), the amount of evaluations says nothing about the validity and the adequate implementation of the evaluations. The next section focuses on the validity and credibility of evaluations to better understand the way the 16 evaluations are carried out.

4.2.2 Validity and creditability of evaluations
It has become clear that according to the employees too few effect evaluations are carried out within the municipality of Wierden and that there are differences in the numbers by sectors. Conducting policy is a purposeful action (Hoogerwerf & Herweijer, 2014; Bekkers, 2012). The achievement of these goals can be associated with complexity and uncertainty as is shown in figure 2 by van de Graaf and Hoppe (1996) and Hoppe (2011) in four distinctive problem types. The fate of policy is that in many situations the 'law of unintended consequences' applies. All units recognise the presence of external factors. However, there are recognisable differences in views on how the municipality of Wierden takes into account these external factors. The respondents from the political unit indicate that the municipality hardly takes into account external factors that can influence policy while a few claim to do so. In the management unit, opinions are sharply divided about the way external factors are visualised. Some respondents indicate that the policy process has 'loose ends' and not sufficiently pays attention to the validity of the policy while others claim to make use of valid and objective data. Most notable are the reactions from the operational-executive unit. One of the assumptions in this unit is that external factors intuitively are visualised because they experience themselves when external factors have influence such as the strike

of the economic crisis or the emergence of a bad financial situation. Figure 1 shows that policy effects may contribute both positively and negatively to the policy and that these effects could be visible, but could also be invisible. The municipality of Wierden creates the image that it keeps up little account of the potential effects associated with the implementation of policies and that these effects receive little attention in policy evaluations.

To get an objective picture of the extent to which the municipality of Wierden visualises policy effects (by means of policy effect evaluations), a document analysis is performed. The 16 de facto effect evaluations are positioned on the evaluation ladder of figure 3. The positioning is carried out based on the criteria presented in table 1. The result is shown in figure 10. The document codes correspond with the effect evaluations listed in appendix 2.

Figure 10 – De facto evaluation score municipality of Wierden
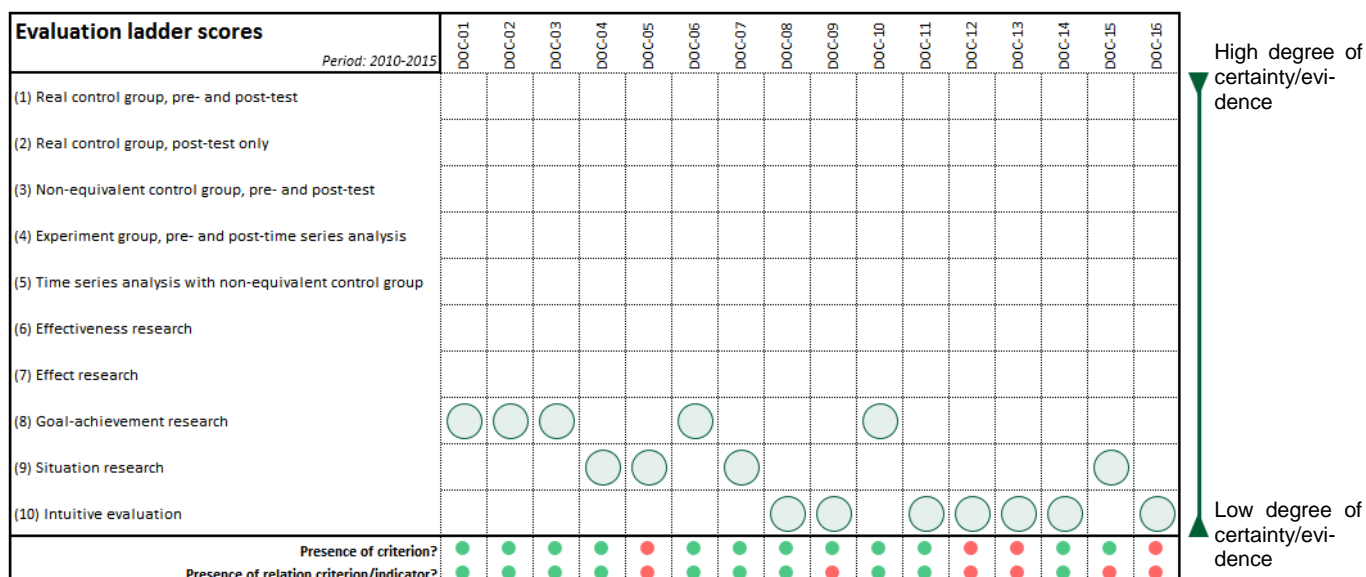

Figure 10 shows two important results leading to the conclusion in the manner in which effect evaluations are valid and credible. The first concerns the positioning of the effect evaluations on the evaluation ladder. The second concerns the way in which criteria are provided to measure the goals and the relationship between the goals and the criteria. Below these two important results are examined in detail.

Positioning on the evaluation ladder
In section 2.2.1, the evaluation ladder is presented (figure 3). This evaluation ladder is included in figure 9 and is put in relation with the 16 de facto effect evaluations. Figure 10 implies that the higher up the evaluation ladder, the more assurance/evidence it provides in respect to the policy effects. The figure makes clear that the de facto effect evaluations score low on the evaluation ladder (positions 8, 9 and 10). This means that the second expectation can be ratified: *'Municipality of Wierden scores low on the evaluation ladder.'* Figure 10 shows that the effect evaluations at the municipality of Wierden are carried out in a fairly intuitive manner. On the lowest stair-tread (step 10), the evaluations are positioned which are conducted intuitively. On stair-tread 9 are situated the evaluations which only indicate changes in the policy field and in which no specific goals or expectations are prepared. At the highest level (stair-tread 8) are the evaluations located in which the relationship is established between the goals and what has been achieved. These stair-treads (8, 9 and 10) deliver little certain knowledge/evidence for the validity and reliability of the results. The 'law of unintended consequences' applies in almost all situations. This implies that there is too little certain knowledge available about the problem and/or there is little consensus on the policy issue in question. These two variables can ensure that policy effects emerge that have not been foreseen and which negatively or positively influence the results of the policy.

When one scores low on the evaluation ladder it means that the policy effects are not identified in a reliable and valid way. The municipality of Wierden therefore is not valid in implementing policy effect evaluations.

The municipality of Wierden is unable to make credible statements about the success or failure of policies when there is little valid and reliable information about the effects of policies. It may be even more important for the learning capacity of the organisation and the ability to evaluate policies in order to learn. This aspect is given more attention in section 4.3. The theoretical framework shows that bounded rationality leads to deficiencies in human cognitive, but also that each individual has its own mental world with their own experiences, perspectives and insights. When the municipality of Wierden thus perform policy evaluations in a (very) intuitive manner (depending on the stair-tread), this can lead to a certain mind-set that may be contrary to reality, but could also be contrary to the perspectives of citizens and societal organisations. In all de facto effect evaluations of the municipality of Wierden, all external factors are ignored. There are five evaluations conducted on stair-tread 8 (goal-achievement evaluation). The advantage of this stair-tread is that it creates the relationship between the predetermined goals and the actual changes while the evaluations on the other two stair-treads do not contain specific goals. A very important remark one must place is that the municipality of Wierden never can say for sure if changes in the policy are due to the policy (instruments). It is therefore essential that the municipality of Wierden is aware of this fact. To clarify where the municipality of Wierden lacks regarding policy effect evaluations, use is made of figure 1. Figure 11 shows an edited version of figure 1 showing the focus of municipality of Wierden regarding effect evaluations. The focus is accentuated in order to indicate what is ignored.

Figure 11 – Shortcomings in effect evaluations at the municipality of Wierden
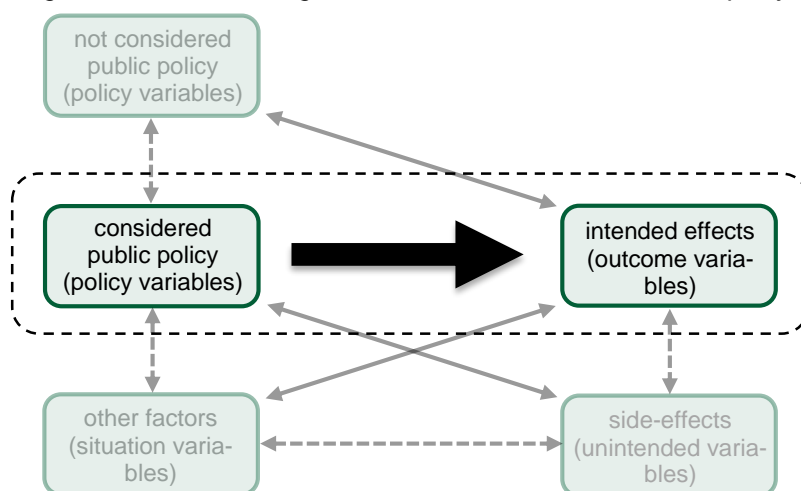
Derived from: van de Graaf & Hoppe (1996, p. 394) (modified)

Figure 11 shows that in effect evaluations of the municipality of Wierden, no attention is given to side-effects of policies, not considered public policy and other factors. When the municipality of Wierden has no insight into the actual effects of policies (both positive and negative) it is not fully able to adjust policies. Overall, this results in the following important implications for the municipality of Wierden:
  ➢ one does not have certain knowledge/results about the actual functioning of policies;
  ➢ one does not know which factors (not considered public policy, other factors and side-effects) and variables (consensus about norm and certainty of knowledge) assert influence on the policy;
  ➢ one cannot make statements about the contribution of different instruments in the success or failure of policies;
  ➢ there are no specific directions to adjust policy or to terminate policy or these directions are 'assumptions' (intuitive);

➢   the learning capacity of the governmental organisation is minimal and can even produce nega-
   tive effects (as will become clear later).

The consequences of carrying out effect evaluations on 'low stair-treads' are significant. The municipal-
ity of Wierden should therefore be cautious in the conclusions it draws and the way it uses the results
from evaluations when it performs evaluations low on the evaluation ladder. The claim: "we can conclude
that the measures have worked well" (Gemeente Wierden, 2014a, p. 2) is incorrect and unjust. In addi-
tion, it is worth noticing that it is not always possible to perform high on the evaluation ladder, as de-
scribed in the theoretical framework. This requires, namely, expertise on scientific level, requires much
time and effort and in addition the law (could) forbid(s) experimental methods. This study therefore aims
not to encourage the municipality to perform on the highest stair-tread possible. It aims (such as the
goal of the research describes) to clarify what the current state of affairs is regarding policy effect eval-
uations and to create awareness what this means for the learning capacity of the organisation. The
image that figure 10 and 11 creates is contrary to the way the municipality of Wierden itself assesses
this aspect. Since there is no consensus on how the municipality of Wierden takes into account external
factors and because the analysis shows that this is minimum, it can be concluded that after the second
expectation, the first expectation has come true: '*the municipality of Wierden expresses little under-
standing of the complexity and uncertainty of policy effects*'. There are many different opinions on the
accuracy of the municipality of Wierden in carrying out effect evaluations. Despite the recognition of
employees that the municipality of Wierden carries out too few evaluations, and that too little attention
is given to this aspect, there are employees who think that the municipality of Wierden is doing this
properly. However, this analysis shows differently. Additionally, there are employees who are aware that
effects are not taken into account and that more attention should be paid beforehand (what is intended
with the policy?). The political unit and the management unit put forward to integrate policy evaluations
more into the policy process of the organisation. The ability of employees to conduct policy effect eval-
uations in a proper way is very controversial in nature. These different perceptions indicate that little
clarity exist in the organisation concerning the adequately implementation of effect evaluations. This is
not in accordance with the results from the analysis since it shows that the municipality of Wierden does
not implement effect evaluations in an optimal way. The municipality of Wierden is therefore not aware
of the extent to which its organisation implements policy effect evaluations adequately.

Presence of criteria and the interrelationship of the criteria
Recall that the definition of policy effect evaluation is focusing on assessing policy on effects based on
certain criteria (Blommestein, Bressers & Hoogerwerf, 1984; van de Graaf & Hoppe, 1996; Hoogerwerf
& Herweijer, 2014). The description of the different evaluations forms, as they are presented in the
evaluation ladder, express that establishing a goal is essential for the validity of effect evaluations. The
evaluations forms (stair-tread 9 and 10) do not make use of specific goals or expectations and can
therefore not provide an accurate picture of the effects of the policy and are not able to establish a
specific relationship between the intended goal and the actual results. Partly for this reason, these eval-
uations are very intuitive. How is one able to make statements about the implementation of policies
when no measurement results are available on what one intends to realise? Using goals and measuring
the goals on the basis of certain criteria/instruments is essential. This means that it must be clear in
advance what one intends to do in order to make statements afterwards (pre-test and post-test), as the
political unit and management unit wishes. Section 4.1 (de jure policy effect evaluation) prescribes that
the organisation should work with explicit goals. The operational-executive unit claims that little use is
made of explicit goals while people in the management unit are controversial on answering this question.
Some argue that not much use is made of explicit goals while others claim that it is common business.
Beside this controversial idea of the situation do all units acknowledge that quantifying goals (SMART-
formulation) an important element forms of the implementation of policies and the evaluations of it. Ex-
plicit goals are important to assess policies on effectivity. The de facto effect evaluations are submitted
to an analysis to determine the extent it uses explicit criteria. In this analysis, an important distinction is
made. First is looked at whether the criterion is described (is present) and then they are assessed on
whether the indicators/instruments are measured in relation to the criterion. Figure 12 (on the next page)
displays the result of this analysis.

Of the 16 de facto effect evaluations, there are 12 that uses explicit goals. Of these 12 evaluations, there are 10 which establish the relation between the indicators/instruments and the criterion while there are two who do not. It should be understood that when no criterion is present, no relationship can be made with the indicators/instruments (see figure 10) and therefore no valid arguments can be made about the effectiveness of the policy.

Figure 12 – Explicit criterion of the de facto evaluations

■ Criterion

■ No criterion

■ Indicators/instruments in relation with criterion

■ Indicators/instruments not in relation with criterion
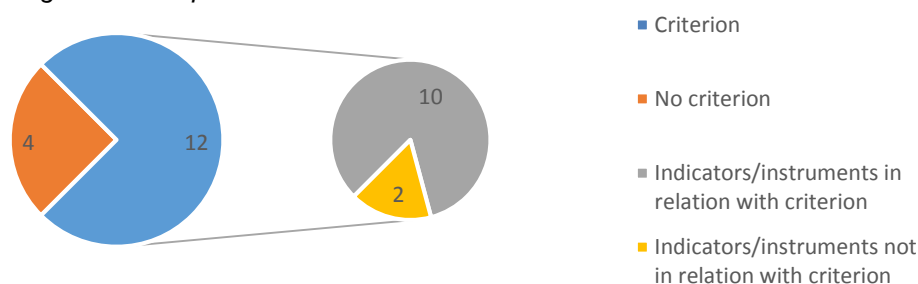
This analysis is important for the way in which the results are valid and reliable, but also for the useful-ness of the evaluation results. The analysis indicates that there are three different situations with respect to the use of criteria and indicators/instruments (see the bottom part of figure 10):
1. The evaluation makes explicitly use of a criterion and there is a relationship between the indi-cators/instruments and the criterion (two green bullets).
2. The evaluation makes explicitly use of a criterion, but there is no relationship between the indi-cators/instruments and the criterion (1 green bullet, 1 red bullet).
3. The evaluation does not make explicitly use of a criterion and for that reason it is not possible to make a relation between the indicators/instruments and the criterion (two red bullets).

Because not in all cases use is made of explicit goals it means that the municipality of Wierden is not always able to make statements about the implementation of policies and about the results obtained from the evaluations. In two cases, there is no link between the indicators/instruments and the criterion. In such policy evaluations, the indicators/instruments are measured on effectiveness. That means that these indicators/instruments and their outcomes are not put in relation with the goal of the policy in order to indicate the contribution of each indicator/instrument to the actual intended goal. These two evalua-tions can only provide results on effectivity of the indicators/instruments and is not able to make state-ments about the effectiveness of the policy in general (the originally intended goal of the policy). The textbox below reflects on behalf of a fictitious example what this means in practice. In this example, only the situation is created around criteria and the relationship between indicators/instruments and the cri-teria. This example does not represent the position on the evaluations ladder although this is closely linked to each other as indicated earlier.

-Textbox – criteria-

Situation 1 – Criterion and indicators/instruments in relation with criterion
Goal: Reduction of 10% in the criminality in the municipality of Wierden. In order to achieve this goal the policy aims to deploy 100 extra police officers (instrument 1) and to place 10 cameras (instrument 2). The criterion is crime (in numbers). The evaluation examines the extent to which the instruments contributed to the reduction in crime in order to make judgments about whether or not the goal is achieved.

Situation 2 – Criterion and indicators/instruments not in relation with criterion
Goal: Reduction of 10% in the criminality in the municipality of Wierden. In order to achieve this goal the policy aims to deploy 100 extra police officers (instrument 1) and to place 10 cameras (instrument 2). The criterion is crime (in numbers). The evaluation only examines whether the deployment of the instruments are accomplished (progress). No attention is paid to the extent the instruments contribute to the reduction of criminality.

Situation 3 – No criterion and therefore also no relationship between indicators/instruments and criterion
There is no goal and no explicit criteria. The evaluation does not make clear what it wants to measure and therefore it is not possible to draw conclusions about the success of the policy. The only thing the evaluation measures is the extent to which the instruments are deployed.

The examples in the textbox are simplifications. In practice, it is not always possible and easy to formulate quantitative goals. This is also highlighted by the various units. In addition, figure 10 shows that in many cases (evaluations on the 9th and 10th stair-tread) no use is made of specific goals (quantified). However, not making use of specific goals leads to difficulties in generating specific results. The criteria is distilled from the policy goal and therefore it is very important that a goal is explicitly formulated. The instruments that are used can be seen as the factors which must ensure the realisation of the goals. In figure 11, this is described as the 'policy variables'. As the theoretical framework indicates does each instrument contain its own 'law of unintended consequences'. Lascoumes and Le Gales (2007) explain that instruments have their own 'force of action' and that not with certainty can be stated that the deployment of instruments lead to the achievement of the goal(s) when no systematic evaluation is being executed.

It is essential to understand that the evaluations in the different stair-treads are positioned on the basis of the method for carrying out effective measurements. This method takes into account the use of specific goals, but does not take into account the validity of the measurement of indicators/instruments in relation with the criteria as explained above. However, there is another aspect resulting in invalid evaluation results. This invalidity is related to the 'coverall evaluation'. This evaluation (DOC-01) focuses on (broader) policy and management products. This means that multiple policy fields form part of the evaluation. Typically, this evaluation focuses on the satisfaction of citizens. The multiple year policy plans contain goals related to the satisfaction of citizens (70% satisfaction about the living environment, 70% satisfaction about employment, etc.). In its evaluation (DOC-01) is described literally: "All together, the implementation of the programmes lead to the satisfaction in the goals..." (Gemeente Wierden, 2010c, 2011d, 2012b, 2013b). This implies that the implementation of the programmes (the specific policies) aim to achieve the goals regarding satisfaction. One may doubt whether the satisfaction of citizens says something about the results of policies (lower in the organisation) because this is independent of whether the intended goals of the policies are achieved. It is only valid when one aims in all policies to achieve satisfaction among the citizens about the functioning of the municipality. This is not the case in the de facto effect evaluations and therefore the satisfaction does not say much about the effectiveness of policies because it ignores all external factors in measuring the actual effectiveness and it ignores the goals of the underlying policy documents. In the textbox below is illustrated this invalidity of measuring satisfaction by elaborating on the fictive example of criminality.

<div style="border:1px solid">

-Textbox – satisfaction-

In the multiple year policy plan are goals included based on the satisfaction of citizens. One of the goals is: 70% satisfaction among the citizens on 'safety of the environment'. In the evaluation of this plan, all goals are checked on the extent to which the goals are achieved. One of the underlying policy documents (part of the implementation of the programme) aims at reducing crime as is indicated in the programme of the multiple year policy plan. In this policy document the goal is to reduce the crime with 10%. In order to achieve this goal the policy aims to deploy 100 extra police officers (instrument 1) and to place 10 cameras (instrument 2). The criterion is crime (in numbers). [It is possible that for this policy a separate policy evaluation is being conducted on the extent to which the goal (10% reduction) is achieved.] This policy succeeded in achieving its goal (18% reduction of crime).

The goal for 'safety of the environment' is not met (62% of the citizens are satisfied on the 'safety of the environment'). The implementation of the policy plan for reducing criminality forms part of the execution of the programme. Because the multiple year policy plan describes that implementation of the programmes aim to achieve the goals regarding satisfaction, it will mean (in this particular case) that the implementation of the instruments leads to dissatisfaction among citizens even while the goal of the specific policy plan (reducing criminality) is met.

</div>

The municipality of Wierden presupposes in this evaluation (DOC-01) that the effectiveness of the implemented policy a reflection is of the satisfaction of citizens on certain aspects. However, the example shows that this does not have to be the case and that these should be regarded as two separate measurements/evaluations. This invalid phenomenon is not consistent with the definition of effectiveness given in the ordinance of the municipality: "the extent to which the performance and the intended social effects of policies are actually being achieved" (Gemeenteraad Wierden, 2004, p. 1). The intended social effect in the policy plan regarding criminality is to reduce criminality with 10%. This means that this

should be measured when one conducts an effect evaluation instead of citizen satisfaction. The municipality of Wierden therefore cannot claim that the policy contributes (positively or negatively) to the satisfaction of citizens despite that it may be that in some cases it does (this should be made clear by a separate research). Bounded rationality also teaches that one does not know all the alternatives. So despite the potential satisfaction of citizens, there may be alternatives that are better (more effective and/or more efficient) to achieve the intended goals. Both the political unit and in the management unit recognise that citizens satisfaction says little about the policy implemented by the municipality. In the operational-executive unit, the majority endorse that it does not measure the effectiveness of the implemented policies. Despite the sceptical image on the satisfaction measurement of policies, all units express that citizen satisfaction form an important part of governmental actions. The next section will pay more attention to this aspect.

Below the de jure checklist is presented and completed. Based on the analyses it can be concluded that the municipality of Wierden does not fully comply with the de jure obligations. Based on the results of this checklist, it is possible to reject the third expectation: *'the de jure mandatory actions aimed at policy effect evaluations are respected by the municipality of Wierden'.*

✅ Municipality draws up regulation about efficiency and effectiveness via ordinance.

✅ At least once a council period the municipality checks on effectiveness.

❌ Working with explicit objectives.

❌ Optimisation of goal-orientation, efficiency, effectiveness and quality of process and product.

In section 4.1 has been announced that the municipality has drawn, as prescribed by law, the ordinance for efficiency and effectiveness. Subsequently, this ordinance stipulates that once a council period the programmes, components thereof or other annual documents should be subjected to an effect evaluation. This is fulfilled by the municipality of Wierden because it conducted in 2013 a municipality-wide effectiveness evaluation (DOC-01) and because it has conducted two other evaluations (DOC-02 and DOC-03) on parts of the organisation/programmes (regardless of the validity of these evaluations). Figure 10 shows that not in all cases use is made of explicit goals/criteria. This means that the method of process is not characterised by working with explicit goals and for that reason the third aspect of the de jure checklist is not met. The analysis (score on the evaluations ladder, the different perceptions and invalidity of the relations between the goal/criterion and the indicators/instruments) shows that there is not a systematic approach to optimise goal-orientation, efficiency, effectiveness and quality of process and product.

4.3 Learning ability municipal organisation

The ordinance Article 213a describes that the goal of performing effect evaluations is to learn (Gemeente Wierden, 2004). The theoretical framework shows that learning a process is for improving actions based on knowledge, experiences and insights. This section examines the extent to which the municipality of Wierden is able to do so based on the de facto effect evaluations.

4.3.1 Learning ability of the de facto effect evaluations
The theoretical framework shows that the learning ability of effect evaluations can make a very important contribution in reducing uncertainty and shifting the boundaries off bounded rationality. Three learning loops are described in the theory. The first uses the results (from evaluations) in order to change the actions and to adjust policies. The second uses the results (from evaluations) to pull the assumptions of the policy in question. The last type of leaning loop concerns learning on an abstract level; learning to learn (context). The results and findings from section 4.2 are ingredients for how the municipal organisation is able to learn (its learning ability). Unreliable evaluations lead to unreliable results. This means that the learning process of these results are also unreliable. The learning ability of the municipality of

Wierden is therefore minimal. The results from section 4.2 show that the ability to learn is minimal since no account it taken of external factors. This means that the municipality of Wierden is unable to learn of factors that may influence the policy. The municipality of Wierden is unable to adequately improve the organisation (based on the de facto effect evaluations) on behalf of knowledge, experience and insights because: 1) it has no certain knowledge, 2) the experience is based on assumptions that are uncertain and 3) because it might provide false insights. Even more important, efforts to learn from the results of the de facto effect evaluations could be dangerous and risky. When the municipality of Wierden uses the results from the de facto effect evaluations to learn, it could mean that this learning process is based on erroneous data because the de facto effect evaluations are implemented in a fairly intuitive way. In short, the results of the de facto effect evaluations can hardly be used by the organisation to learn.

The employees widely support that the municipality of Wierden is very poor in using the results of the evaluations. In the evaluations is not indicated what the purpose of the evaluation is, and is often not described what is done with the results. The ordinance Article 213a explicitly prescribes that the results of the de jure evaluations are provided for cognisance to the council. The political unit endorses this and adds that the results are only used in order to adjust policies. The management unit emphasises that learning the intention is of policy evaluations, but that this is done deficient. The results are only used, according to the respondents in this unit, for cognisance/information provision and sometimes results in adjusting policy. Additionally, sometimes the results will be discussed with stakeholders. This is also recognised by the operational-executive unit. The foregoing illustrates some important aspects. The learning process consists only of providing knowledge/cognisance. This means that in the process description of Hoogerwerf and Herweijer (2014) in using the results, only attention is given to the first aspect (cognisance of findings) while no attention is paid to affecting standpoints and the impact of the findings (the learning process is not completed). In addition, the foregoing indicates clearly a single-loop process. The results are in some cases only used to adjust the policy (in order to adapt actions). According to the theoretical framework can this be characterised as single-loop learning. It indicates that the municipality of Wierden does not undertake double- or triple-loop learning. In short, the validity of the evaluations (such as figure 10 shows for the situation of the municipality of Wierden) is closely related with the learning capacity of the organisation. The learning ability of the municipality of Wierden, based on the de facto evaluations, results in the following features:

> ➢ There is no focus on the learning ability of effect evaluations because two elements of the learning process are missing (affecting standpoints and impacts of the findings).
> ➢ The learning ability of the effect evaluations is based solely on single-loop learning.
> ➢ It is not able to learn optimally since the implementation of the effect evaluations is not done in a systematic manner (see figure 10 and 11).

As already indicated, there is a significant risk in using the results of the de facto effect evaluations because the adjustment and/or renewal of policies is based on 'assumptions' as the de facto effect evaluations do not provide reliable facts. Based on the features of the learning ability of the municipality of Wierden it is possible to conclude that the fourth expectation is fulfilled: '*the (organisational) learning ability of the municipality of Wierden in policy effect evaluations will primarily take place on the basis of single-loop learning'.* In complex problems (see figure 2) the municipality of Wierden is not able to find out which variable (consensus about norms and certainty of knowledge) influences the policy process. However, the various units endorse that in many situations stakeholders are involved in evaluating policies. This is reflected in the involvement of stakeholders in the communication of the evaluation results. This then leads to adjustment of the policy in most cases, according to the respondents. This involvement is also described in the coalition agreement and in all the multi-year policy plans. This means that although the evaluation is carried out intuitively, the broad involvement of stakeholders may result in a wider supported learning process on the basis of different perceptions. This may indicate support for the results and can also mean that the results are more likely to be true. Despite this subjectivity based on perceptions, one cannot with certainty state that policy was implemented successfully of not. Therefore, involving stakeholders could be positive, but the actual learning process highly depends on the validity of the evaluations. Despite the indication in the ordinance (Article 213a) that the goal of evaluating is to learn from it, the learning ability of the municipality of Wierden leads not or barely to the reduction of uncertainty of complex problems (unstructured problems).

4.4 Answering the sub-questions

This chapter presents the results of the research. On the basis of these results, the sub-questions can be answered. The sub-questions together provide the answer to the main question (see chapter 5). This section answers the sub-questions.

Sub-question 1
What are the de jure obligations for the municipality of Wierden regarding policy effect evaluations?

An analysis was performed in which the municipal act and various ordinances are inspected on behalf of obligations and codified ambitions. This has resulted in a list of obligations for the municipality of Wierden in relation to policy effect evaluations which answers this sub-question. These are the following obligations:
- Municipality draws up regulation about efficiency and effectiveness via ordinance.
- At least once a council period the municipality checks on effectiveness.
- Working with explicit objectives.
- Optimisation of goal-orientation, efficiency, effectiveness and quality of process and product.

Providing insight into these obligations is considered important for assessing how the municipality of Wierden performs effect evaluations (part of the central research question). The assessment of the compliance with the obligations provides, in relation with the theory, interviews and document analyses, information on the reasons why these aspects are not met.

Sub-question 2
To what extent does the municipality of Wierden perform de facto policy effect evaluations?

The evaluation ladder together with figure 1 from the theory are used to link the de facto effect evaluations with forms of effect evaluations in order to position them on the evaluation ladder. The theories have been used to make statements about the way the municipality of Wierden carries out effect evaluations. Based on the theory the sub-question can be answered with the conclusion that the municipality of Wierden performs effect evaluations on a fairly intuitive way in which the results of these evaluations deliver little to no certain information concerning the functioning of the policy. Also, there are cases in which no use is made of explicit goals or no relation is being made between the criterion and the criteria/indicators of the policy. These results are also used to check whether the de jure obligations have been met.

Sub-question 3
How can the learning ability for the municipality of Wierden be improved regarding policy effect evaluations?

To make statements about the way the municipality of Wierden can increase its learning ability it is essential to first look at the extent to which the municipality of Wierden is able to learn based on the de facto effect evaluations. Therefore an analysis is performed based on the findings of the answer on sub-question 2. The validity of the de facto effect evaluations shows that the municipality of Wierden is unable to learn double- and/or triple-loop. Currently the municipality of Wierden only learns on the basis of single-loop learning. This results in a risk because this learning ability is possibly based on incorrect information. To increase the learning ability of the municipality of Wierden it is therefore required to increase the validity of the effect valuations and to start learning double- and triple-loop.

The answers to the sub-questions form the building blocks of answering the central research question. This is highlighted in the following chapter.

5. CONCLUSION

This chapter includes the conclusion of the research. This conclusion answers the central research question on behalf of the sub-questions. In addition, based on the conclusion recommendations are formulated in terms of policy effect evaluations. This is followed by a discussion in which the results are interpreted and limitation are identified.

5.1 Conclusion

The call for performing this research came from the feeling of the municipality of Wierden that it pays little attention to the implementation of policy effect evaluations. The municipal organisation is searching for bottlenecks and areas for improvement for implementing effect evaluations in and adequate way. The research focuses on the screening of effect evaluations at the municipality of Wierden for the period 2010-2015. To identify the bottlenecks and areas of improvements, a research question is formulated.

> *How does the municipality of Wierden perform policy effect evaluations and how can the learning ability of the organisation be increased based on the effect-measuring performance?*

It can be concluded that the municipality of Wierden does not perform policy effect evaluations in an adequate way. Also, the learning ability of the municipality of Wierden based on the de facto effect evaluations is minimal. In order to be more reflexive as a government, it is crucial to work with results that are valid and credible. This means that increasing the learning ability starts with increasing the validity of the evaluations. A higher validity of the findings of evaluations provide reliable data to adjust policy actions (single-loop learning), to check the assumptions of the policy on truth (double-loop learning) and to be able to learn how to learn (triple-loop learning). More systematic evaluations are needed to increase the learning ability of the municipal organisation. Section 5.2 (recommendations) provides nuanced recommendations in order to realise more valid effect evaluations. The answer to the research question is based on various conclusion.

Firstly, it can be concluded that the municipality of Wierden complies with the legally determined (by law) obligations in respect of effect evaluations, but that it fails to fully comply with its 'own' codified ambitions (by ordinances). Not complying with the codified ambitions concerning effect evaluations conducts influence on the validity of the effect evaluations currently being conducted (de facto effect evaluations 2010-2015). Not in all cases use is made of explicit goals and the quality is not guaranteed in an optimal manner in terms of goal-orientation, efficiency and effectivity of process and product.

Secondly, it can be concluded that the municipality of Wierden ignores the 'law of unintended consequences'. It does not take into account external factors in evaluating policies. This results in the positioning of the de facto effect evaluations low on the evaluation ladder. This means that the de facto effect evaluations are implemented on a fairly intuitive way. The results and findings of these de facto evaluations provide little to no certain knowledge about the performance of the policy in terms of effectivity. In addition can also be concluded that in some cases inconsistencies exist between the goals of the policy and the goals used in multi-year policy plans. The objectives in the multi-year policy plans are based on satisfaction, but this does not provide evidence for measuring the effectiveness of the underlying policy of the municipality of Wierden. Also, in some cases no criteria are linked to the evaluation resulting in effect measurements focusing on the effectiveness of the instrument instead of the goal(s) of the policy, or the effectiveness of the instruments are not put in relation with the goal(s). In conclusion, this means that he municipality of Wierden does not know with certainty what the social effects are of conducted policies.

Thirdly, it can be concluded that based on the de facto effect evaluations the municipality of Wierden is not able to learn in an adequate manner of the policies conducted. The municipality of Wierden is only able to undertake the single-loop learning process, but this is confronted with a major risk because the knowledge, experience and insights are based on intuitive results (uncertainty). This means that it is only able to make policy adjustments and is not able to test assumptions or to learn in a wider context of the policy. In addition, it can also be concluded that the learning process is not completed, but that the results of the evaluations will be used for cognisance and that it does not result in influencing stand-points and in impacts of the findings.

The theoretical framework forms the theoretical context form which the results are approached. This had resulted in four expectations regarding the performance of the municipality of Wierden on effect evaluations. The results of the research were able to make statements about whether or not the expectations are met. In conclusion, this had resulted in the following.

✔ Exp. 1: Municipality of Wierden expresses little understanding of the complexity and uncertainty of policy effects.

✔ Exp. 2: Municipality of Wierden scores low on the 'evaluation ladder'.

✘ Exp. 3: The de jure mandatory actions aimed at policy effect evaluations are respected by the municipality of Wierden.

✔ Exp. 4: The (organisational) learning ability of the municipality of Wierden in policy effect evaluations will primarily take place on the basis of single-loop learning.

The more complex the problem, the higher the demand for systematic evaluations. Systematic effect evaluations result in findings from which one can learn. This can better guarantee the certainty of knowledge in adjusting policies (single-loop learning) and in the assumptions of the policy (double-loop learning). This results in less complex problems in the future (structured problems).

5.2 Recommendations

Conducting policy effect evaluations in a valid way is important for the learning ability of the organisation and is also important to be reflexive. Complex problems demand a thorough approach and adequate measurement of results because of many uncertainties that goes along with the complexity. The Matthew Effect shows that well-intentioned policies can lead to adverse effects/developments. It also has become clear that in complex problems the external effects are not always foreseen (as the Matthew Effect illustrates), and that for this reason it is very important to verify afterwards what the policy brought about. To increase the learning ability of the municipality of Wierden based on the effect evaluations, some recommendations could be presented.

Conducting more valid evaluations requires a use-oriented approach. Utilisation-based evaluations are important to provide certain evidence. When one is formulating policies it is recommended to take into account what one wants to achieve at the back. This means that ex ante the establishment of explicit goals is essential. Also, the quantification of these goals is important to obtain an objective measurement. The evaluation ladder shows that systematic evaluations have an ex ante measurement, a so-called baseline measurement (zero-measurement). It is recommended to do this in order to compare the situation ex ante with the situation after the implementation of the policy (ex post). To pay more attention to the effect evaluations it is also important to clarify who is responsible for proper implementation of the evaluations. Therefore, it is also advisable to assign responsibilities and to better enforce it in order to comply with the codified ambitions of the municipality of Wierden. Better communication and better control on how evaluations should be carried out so that people are aware of what is expected from them. To ensure this in an optimal way and to establish some structure it is advised to create a guideline/directive in which employees are coached in how evaluations can be prepared in an adequate manner. This guideline/directive can also be used to steer in which way the criteria should be related to the instruments. It is recommended to clarify this thoroughly in order to be able to attribute the results of the evaluations to the policy (increasing the validity).

In addition to the methodological recommendations, there are also some recommendations on how citizens and actors should be involved. Involving citizens and actors in the preparation, implementation and evaluation of policies is recommended because this may result in consensus on the problem approach which might result in better compliance of the policy/regulations. Even if there is no consensus established, it is important to get in advance a clear understanding of the possible negligence of citizens and/or actors regarding the policy. In addition, it is strongly recommended to make clear what is done with the results of the effect evaluation. Discuss them with involved people and citizens rather than just distributing the results for cognisance/information provision. These discussions can lead to generating new knowledge, experiences and insights (learning ability based on perceptions) and is able to achieve double-loop learning.

In summary, the recommendations lead to the following action points. The action points are categorised based on 'process' and 'content'. These action points result in increasing the validity of the effect evaluations and ensure the development of the learning ability of the municipal organisation.

Action points (recommended)

Process
- involving citizens and actors in the preparation, implementation and evaluation of policy
- recording and enforcing responsibilities regarding effect evaluations
- discuss the results of evaluations with those involved and use the results to double-, and when possible to triple-loop learn

Content
- drafting of explicit (quantified) goals in advance
- establish the relation between the baseline measurement and the measurement after performing the policy (in the evaluation)
- establish the relation between criteria and instruments in respect of the goal

It is recommended to record the above action points in a guideline/directive so that it is easy accessible and one is aware of all aspects of conducting an adequate effect evaluation.

5.3 Discussion

The results of this study are valid and reliable. By applying a combined method, the findings are verified in two ways. The only remark in this methodology is that interviews are context-driven snapshots of the situation and are very subjective based on individual experiences, insights and knowledge (as bounded rationality teaches us). Additionally, the document analysis provides objective data on the actual performance of the municipality of Wierden in terms of effect evaluations. Important in this research is the way the results need to be interpreted and the limitations resulting therefrom. The results are objective data and is mirrored on the theory in which an optimal picture is scooped of the way in which the evaluations should be performed. However, a government organisation is always dependent on politics. One can strive to ensure adequate implementation of policy evaluations, but when politicians decide that little or no attention should be given to this phenomenon, this should be followed up by the organisation. This research encourages the municipality to perform effect evaluations adequately, but is aware that this is not always possible/desirable despite the fact that this is a prerequisite for increasing the learning ability. It can also be questioned whether a pragmatic and relatively small municipality is able to carry out evaluations in an adequate way. A lot of expertise, time and money is needed. This could therefore mean that the action points cannot all be optimally safeguarded in the implementation of all policies. Also, a government can consider the net benefit from the money that comes along with performing effect valuations and the learning process flowing from the evaluations results. It is often subjective and politically dependent on whether the benefits outweigh the costs.

In interpreting the results one must realise that the learning ability is based solely on the performance of the de facto effect evaluations. This means that all other possibilities to learn as an organisation have been omitted. There may be several ways in which an organisation optimise its ability to learn. Examples include guest presenters, publications of various governments, scientific publications etc. Also, one must realise that a proper implementation of policies and a proper execution of effect evaluations may result in the reduction of uncertainty, but that nevertheless external factors could occur which negatively influence the policy. As the theory describes is even evidence-based policy falsifiable and could thus be confronted with the 'law of unintended consequences'.

The complexity of society also demands a continuous development of the government to bridge the gap between the municipality together with its organisations and society. Herein politics play a crucial role and can provide guidance and awareness. Measuring the actions of governments on effectivity will therefore remain to play an essential role. Evaluating effects is a pluriform and dynamic activity, but is also a very important activity.

Suggestions for further research
This research provides an overview of how the municipality of Wierden carries out policy effect evaluations in relation to the de jure obligations and in relation to the learning ability of the organisation. Some suggestions can be made for further research to support the municipality of Wierden in conducting adequate effect evaluations and to present opportunities therein. This leads to the following suggestions:

 ➢ Research into the relationship between policy and policy evaluation
 In this research can be explored which policy should be subjected to evaluation and what can be regarded as policy. As the results of this research indicates, this is not part of this research because it is a thorough analysis which goes beyond the scope of this research. This research may be relevant to obtain an overview of the relationship between policy and policy evaluation and can provide information about the municipality of Wierden for evaluation prioritisation in policies.
 ➢ Comparison research municipalities of Twente
 It is not superfluous to conduct a comparison study between municipalities in Twente with respect to the performance of municipalities concerning policy effect evaluations. This provides an overview of the performances of different municipalities on this aspect. From this, municipalities are able to use examples to better and more adequate implement effect evaluations.
 ➢ Cooperation municipalities with respect to policy evaluation
 As the research indicates, an adequate effect evaluation requires lots of time, money and expertise. Research may reveal if it is possible to create a partnership between municipalities. Also, such a research may reveal whether municipalities acknowledge the importance of the phenomenon 'policy effect evaluation'.

BIBLIOGRAPHY

Below a list is provided with all consulted sources in order to carry out this research. The sources are displayed according to the APA-guidelines.

Argyris, C. (1976). Single-Loop and Double-Loop Models in Research on Decision Making. *Administrative Science Quarterly, 21*(3), 363-375.

Argyris, C. (2004). Reflection and Beyond in Research on Organizational Learning. *Management Learning, 35*(4), 507-509. doi:10.1177/1350507604048276

Argyris, C., & Schön, D. (1974). *Theoyr in Practice: Increasing Professional Effectiveness.* San Francisco: CA: Jossey-Bass.

Baarda, B., Bakker, E., van der Hulst, M., Fisscher, T., Julsing, M., van Vianen, R., & de Goede, M. (2012). *Basisboek Methoden en Technieken. Kwantitatief praktijkgericht onderzoek op wetenschappelijke basis.* Groningen/Houten: Noordhoff Uitgevers.

Bekkers, V. (2012). *Beleid in beweging. Achtergronden, benaderingen, fasen en aspecten van beleid in de publieke sector.* Den Haag: Boom Lemma.

Bessant, J. (2008). The sociology of policy-making in the modern state. *Journal of Sociology, 44*(3), 283-300. doi:10.1177/1440783308092885

Blommestein, H., Bressers, J., & Hoogerwerf, A. (1984). *Handboek beleidsevalautie.* Samson: Alphen aan den Rijn.

De Marchi, G., Lucertini, G., & Tsoukiás, A. (2014). *From Evidence-Based Policy Making to Policy Analytics.* Paris: Université Paris Dauphine.

de Vaus, D. (2001). *Research Design in Social Research.* London/Thousand Oaks/New Delhi: Sage Publications.

Dooley, D. (2009). *Social Research Methods.* (H. Vos, Ed.) Essex: Pearson Custom Publishing.

Francot, L. (2014). Dealing with complexity, facing uncertainty morality and ethics in a complex society. *Archiv fur Rechts- und Sozialphilosophie, 100*(2), 201-218.

Gemeente Wierden. (2004). *Verordening voor periodiek onderzoek door het college naar de doelmatigheid en doeltreffendheid van het door het college gevoerde bestuur van de gemeente Wierden.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2005). *Inspraakverordening gemeente Wierden.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2007). *Organisatiebesluit 2009.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2009). *Nota Meerjarenbeleid 2010-2012.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2010a). *Coalitieakkoord CDA-PPW 2010-2014.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2010b). *Evaluatie pilot dienstverlening Werkplein Noord Twente.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2010c). *Nota Meerjarenbeleid 2011-2013.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2011a). *Doeltreffendheidsonderzoek 2011. Sociaal Cultureel Werk, Algemene jongeren.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2011b). *Doeltreffendheidsonderzoek 2011. Werk, Ingkomen en Zorg.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2011c). *Evalautie afschaffen inkomenstoets Wmo regiotaxipas.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2011d). *Evaluatie cameratoezicht WIEZO.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2011d). *Nota Meerjarenbeleid 2012-2014.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2012a). *Evaluatie SOWECO.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2012b). *Nota Meerjarenbeleid 2013-2015.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2013a). *Doeltreffendheidsonderzoek 2013. Gemeente-breed.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2013b). *Nota Meerjarenbeleid 2014-2016.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2014a). *Evaluatie activiteitenplan weren roeken.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2014b). *Evaluatie groene diensten.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2014c). *Evaluatie klimaatprogramma.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2015a). *Evalautie beleidsnota duurzame energie.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2015b). *Evaluatie bezuinigingen op het openbaar groen.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2015c). *Evaluatie cultuureel arrangement Wierden.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2015d). *Evaluatie intentieverklaringen gemeente-verenigingen voor plaatselijk en gemeenschappelijk belang.* Wierden: Gemeente Wierden.

Gemeente Wierden. (2015e). *Evaluatie nota Wierden weet van kunst.* Wierden: Gemeente Wierden.

Gerrichhauzen, J., Korsten, A., & Fijen, H. (2002). De lerende organisatie in een ambtelijke cultuur. *Bestuurswetenschappen, 3.*

Green, D. (2014). *Of Ants and Men. The Unexpected Side Effects of Complexity in Society.* Clayton: Springer.

Hague, R., & Harrop, M. (2013). *Compararitive Government and Politics: An introduction.* Basingstoke: Palgrave Macmillan.

Hill, M., & Hupe, P. (2002). The Rise and Decline of the Policy-Implementation Paradigm. In M. Hill, & P. Hupe, *Implementing Public Policy: Governace in Theory and in Practice* (pp. 85-115). London/Thousand Oaks/New Delhi: Sage Publications.

Hoogerwerf, A., & Herweijer, M. (2014). *Overheidsbeleid. Een inleiding in de beleidswetenschap.* Alphen aan den Rijn: Kluwer.

Hoppe, R. (2011). *The Governance of Problems. Puzzling, powering and participation.* Bristol: The Policy Press.

Iedema, R., & Wiebinga, P. (2010). *Profiel van de Nederlandse Overheid. Organisatie, beleid en besluitvorming.* Bussum: Coutinho.

Jones, B. (1999). Bounded Rationality. *Annual Reviews Political Science, 2*, 297-321.

Klinkers, L. (2002). *Vakvereisten voor politiek en beleid.* Meise: Klinkers Public Policy Consultants BVBA.

Kohoutek, J., Nekola, M., & Novotný, V. (2013). Conceptualizing Policy Work as Activity and Field of Research. *Central European Journal of Public Policy., 7*(1), 28-59.

Korsten, A. (2013). *Wat ervan terecht komt. Zicht op beleidsevaluaties.* Retrieved from http://www.arnokorsten.nl/downloads.html

Korsten, A., & Hoppe, R. (2006). *Van beleidswetenschap naar kennissamenleving. Voortgang, vooruitgang en achteruitgang in de beleidswetenschap.* Arno Korsten. Retrieved from http://www.arnokorsten.nl/downloads.html

Lascoumes, P., & Le Gales, P. (2007). Introduction: Understanding Public Policy through Its Instruments. From the Nature of Instruments to the Sociology of Public Policy Instrumentation. *Governance: An International Journal of Policy, Administration, and Institutions., 20*(1), 1-18.

Lindblom, C. (1979). Still muddling, not yet through. *Public Administration Review*, 517-526.

March, J., & Olsen, J. (2009). *The logic of appropriateness.* Oslo: University of Oslo.

Merton, R. (1957). *Social Theory and Social Structure.* Glencoe: Free Press.

Ney, S., & Verweij, M. (2014). Exploring the Constributions of Cultural Theory for Improving Public Deliberation about Complex Policy Problems. *The policy Studies Journal, 42*(4), 620-638.

Organisation for Economic Co-operation and Development (OECD). (2000). *Reducing the Risk of Policy Failure: Challenges for Regulatory Compliance.* Paris: OECD.

Osborne, D., & Gaebler, T. (1992). *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector.* London: Penguin.

Ossewaarde, M. (2013). *Theorizing European Societies.* London: Palgrave Macmillan.

Papadopoulos, Y., & Warin, P. (2007). Are innovative, participatory and deliberative procedures in policy making democratic and effective? *European Journal of Political Research, 46*, 445-472.

Patton, M. (2008). *Utilization-Focused Evaluation.* London/Thousand Oaks/New Dehli: Sage Publications.

Pawson, R. (2002). Evidence-based Policy: In Search of a Method. *Evaluation, 8*(2), 157-181.

Power, M. (2000). The Audit Society - Second Thoughts. *International Journal of Auditing, 4*, 111-119.

Sanderson, I. (2002). Evaluation, Policy Learning and Evidence-Based Policy Making. *Public Administration, 80*(1), 1-22.

Sharman, A., & Holmes, J. (2010). Evidence-Based Policy or Policy-Based Evidence Gathering? Biofuels, the EU and the 10% Target. *Environmental Policy and Governance, 20*, 309-321. doi:10.1002/eet.543

Simon, H. (1972). Theories of Bounded Rationality. In C. Mc Guire, & R. Radner, *Decision and Organization.* (pp. 161-176). Amsterdam: North-Holland Publishing Company.

Strassheim, H., & Kettunen, P. (2014). When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. *Evidence & Policy, 10*(2), 259-277. doi:10.1332/174426514X13990433991320

Tosey, P., Visser, M., & Saunders, M. (2011). The origins and conceptualizations of 'triple-loop' learning: A critical view. *Management Learning, 43*(3), 291-307. doi:10.1177/1350507611426239

van de Graaf, H., & Hoppe, R. (1996). *Beleid en politiek. Een inleiding tot de beleidswetenschap en de beleidkunde.* Bussum: Coutinho.

van den Berg, M., Meijer, C., & Slager, F. (2010). *Recht. Een inleiding.* Baarn/Utrecht/Zutphen: ThiemeMeulenhoff.

van Elk, R., van der Meer, F.-B., van der Steeg, M., & Webbink, D. (2011). *Zicht op Effectiviteit van Beleid. Studie naar evaluatieontwerpen voor onderwijs- en wetenschapsmaatregelen.* Retrieved from http://www.beleidsevaluatie.info/wp-content/uploads/2011/06/cpb-achtergron-ddoument-zicht-op-effectiviteit-van-beleid.pdf

Verhoeven, N. (2010). *Wat is onderzoek? Praktijkboek methoden en technieken voor het hoger onderwijs.* Den Haag: Boom onderwijs.

Voermans, W. (n.d.). *Van beleidsbegroting tot beleidsverantworoding.* Retrieved June 25, 2015, from Leiden University Repository: https://openaccess.leidenuniv.nl/bitstream/handle/1887/3916/360_318.pdf?sequence=1
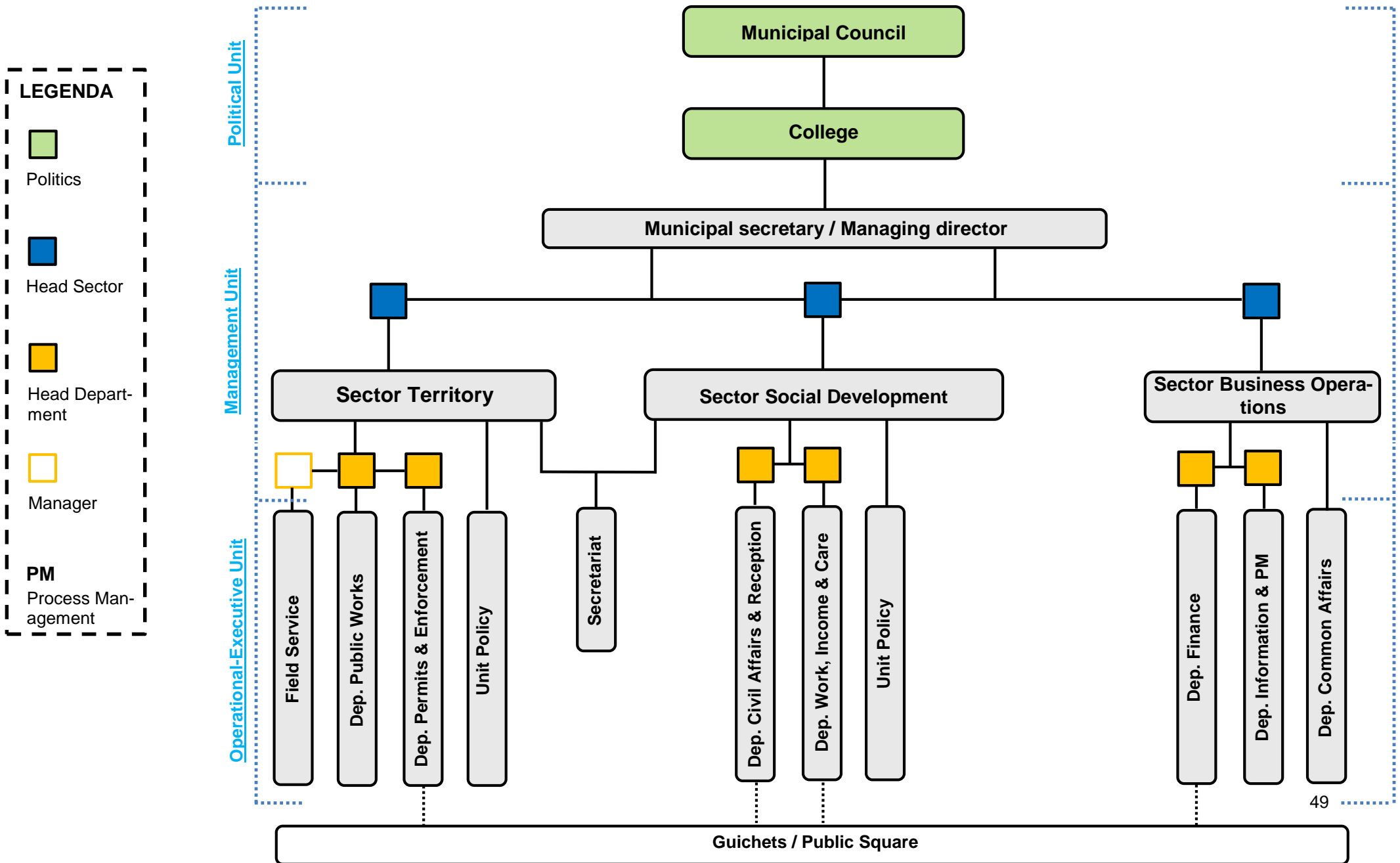
ATTACHTMENTS

Appendix 1: Organogram municipality of Wierden

Appendix 2: List of policy documents and interviewees

Appendix 1: Organogram municipality of Wierden

49

Appendix 2: List of policy documents and interviewees

Below two tables are displayed. Table 1 provides an overview of all the consulted internal documents of the municaplity of Wierden. Table 2 provides an overview of all interviewees including relevant data. All the interviewees are employees of the municipality of Wierden from different departments and with different functions.

Table 1 – List of intern consulted documents

Document name	Type of document	Evaluation ladder code
Coalitieakkoord 2010-2014	Political document	n.a.
Doeltreffendheidsonderzoek 2011 (Sociaal Cultureel Werk Algemene Jongeren	Research/evaluation	DOC-02
Doeltreffendheidsonderzoek 2011 (Werk Inkomen en Zorg)	Research/evaluation	DOC-03
Doeltreffendheidsonderzoek 2013 (Gemeentebreed)	Research/evaluation	DOC-01
Evaluatie activiteitenplan weren roeken	Evaluation	DOC-04
Evaluatie afschaffen inkomenstoets Wmo regiotaxipas	Evaluation	DOC-05
Evaluatie beleidsnota duurzame energie	Evaluation	DOC-06
Evaluatie bezuinigingen op het openbaar groen	Evaluation	DOC-07
Evaluatie cameratoezicht WIEZO	Evaluation	DOC-08
Evaluatie cultureel arrangement Wierden	Evaluation	DOC-09
Evaluatie groene diensten	Evaluation	DOC-10
Evaluatie intentieverklaringen gemeente-verenigingen voor plaatselijk en gemeenschappelijk belang	Evaluation	DOC-11
Evaluatie klimaatprogramma	Evaluation	DOC-12
Evaluatie nota Wierden weet van kunst	Evaluation	DOC-13
Evaluatie pilot dienstverlening Werkplein Noord Twente	Evaluation	DOC-14
Evaluatie SOWECO	Evaluation	DOC-15
Evaluatie stimuleringsregeling woningbouw en bedrijventerreinen	Evaluation	DOC-16
Inspraakverordening gemeente Wierden	Ordinance	n.a.
Nota Meerjarenbeleid 2010-2012	Policy document	n.a.
Nota Meerjarenbeleid 2011-2013	Policy document	n.a.
Nota Meerjarenbeleid 2012-2014	Policy document	n.a.
Nota Meerjarenbeleid 2013-2015	Policy document	n.a.
Nota Meerjarenbeleid 2014-2016	Policy document	n.a.
Organisatiebesluit 2009	Ordinance	n.a.
Verordening Artikel 213a gemeentewet	Ordinance	n.a.

Table 2 – List of interviewees

Interviewee	Interviewer	Function	Date
Political unit			
J.H.M. (Henk) Robben	C.C.G. (Chiel) Mensink	Mayer of the municipality of Wierden	14-October-2015
T.P. (Theo) de Putter	C.C.G. (Chiel) Mensink	Alderman of the municipality of Wierden	30-September-2015
R. (Rick) Schasfoort	C.C.G. (Chiel) Mensink	Alderman of the municipality of Wierden	30-September-2015
D.F. (Dianne) Span-Kiekebeld	C.C.G. (Chiel) Mensink	Alderman of the municipality of Wierden	09-September-2015
Management unit			
R.B. (Bjørn) van den Brink	C.C.G. (Chiel) Mensink	Municipal secretary / managing director	02-September-2015
M.A.E. (Marie Jeanne) Boltong	C.C.G. (Chiel) Mensink	Head sector Social Development	24-November-2015
W.W.G. (Wouter) Brinkman	C.C.G. (Chiel) Mensink	Head department Common Affairs	14-October-2015
H.F. (Henk) Hesselink	C.C.G. (Chiel) Mensink	Head department Work, Income & Care	19-November 2015
H.M. (Hans) Hubers	C.C.G. (Chiel) Mensink	Head department Public Works and deputy head sector Territory	12-November-2015
B. (Ben) Maas	C.C.G. (Chiel) Mensink	Head department Civil Affairs & Reception	13-Novemer-2015
M.G.J. (Mark) Seyger	C.C.G. (Chiel) Mensink	Head department Finance	12-November-2015
Operational-executive unit			
A.J. (Andrea) Holterman	C.C.G. (Chiel) Mensink	Policy maker Unit Policy	26-November-2015
J.M. (Jan Marten) de Jong	C.C.G. (Chiel) Mensink	Policy maker communication / consultant	26-November-2015
J.H.W. (Hans) Langen	C.C.G. (Chiel) Mensink	Strategic policy maker	12-November-2015
J. (Joop) Pasman	C.C.G. (Chiel) Mensink	Collaborator Public Works	19-November-2015