Radboud University Nijmegen

Master's thesis

From pilots to policy change

Researching the pilot paradox within internally and externally successful drought adaptation

pilots in Dutch dry rural areas

Author: Menno (M.N.J.) van Dongen

Master's Thesis for the Spatial Planning programme

Specialization: Cities, Water and Climate Change

Nijmegen School of Management Radboud University Nijmegen

Date: 08-2021

From pilots to policy change

Researching the pilot paradox within internally and externally successful drought adaptation pilots in Dutch dry rural areas

Colophon

Menno van Dongen Student number: s1065428

04-08-2021

Master's thesis Spatial Planning: Cities, Water and Climate Change Faculty of Management Sciences Radboud University Nijmegen

In combination with a graduate internship at Deltares

Supervisor Radboud University: Prof. Dr. S.V. Meijerink Supervisor Deltares: Drs. G.J. Ellen

KLIMAP



Radboud Universiteit



I. Preface

Before you lies the final report of my master thesis research into the conditions for internally and externally successful pilot projects aimed at drought adaptation in Dutch dry rural areas. With this research, I hope to complete the Spatial Planning master with the specialization Cities, Water and Climate Change at the Radboud University Nijmegen.

I have experienced the process of writing this thesis as quite intense, but on the other hand also really enjoyable, because of the nice meetings and discussions with my university professors, my colleagues at Deltares and other specialists in the field of water and soil management. In the first place, I would like to thank Sander Meijerink for the intensive support and the encouraging, sharp feedback on my research. I also would like to thank Huub Ploegmakers for the support with regard to the research design and QCA in particular. The QCA methodology was new for me, which made the whole process of writing this thesis extra challenging. The regular meetings with Sander and Huub were really helpful and therefore much appreciated. Additionally, I would like to thank Gerald Jan Ellen for the support during my internship at Deltares. I am thankful that I got the chance to learn more about Deltares, and enjoyed working on the various projects that I were involved with. During my internship I also got the chance to work on the subject of pilots and upscaling in a project for the municipality of Rotterdam, for which I would like to thank Heleen Vreugdenhil and Jitske van Popering-Verkerk. Finally, a thank you to Onno Giller and my fellow interns Merlijn Anderiesse and Jort de Vries, for the nice meetings and the feedback moments. I have found it very helpful to exchange literature and thoughts about the research process.

I learned a lot from conducting this research and hope to be able to apply my knowledge in the working field of spatial planning and water management in the near future.

I hope you enjoy reading this report.

Menno van Dongen

Rijsbergen, 04-08-2021

II. Summary

Temporal droughts are an increasingly problematic phenomenon in Dutch dry rural areas, resulting in economical, ecological, and societal damage on the short- and long term. At the same time, governments on all geographical levels are developing adaptation policies and programmes to adapt to the drought issue. This requires developing new innovative adaptive solutions. To realize this, pilot projects are used as innovation instruments. These pilot projects are strategically used to develop innovations within an environment that limits the impact of the innovative change (Turner & Muller, 2008). When the innovation that was desired is reached, the aim of many pilots is to scale the results up in order to realize a broader impact (Van den Broek, Van Elzakker, Maas & Deuten, 2020). However, scientific literature about pilots describes that there is a paradoxical relation between conditions that contribute to on the one hand the ability of the pilot to reach the previously stated innovative outcomes (the internal success), and on the other hand the ability of the pilot to scale up the innovation (the external success). This phenomenon is called the 'pilot paradox' (Van Buuren, Vreugdenhil, Van Popering-Verkerk, Ellen, Van Leeuwen & Breman, 2018). The pilot paradox distinguishes five dimensions that all have a paradoxical character with regard to the internal and the external success: position of the pilot, project design, resource distribution, participants and process design. The pilot paradox was initially developed as a qualitative phenomenon. This research focused on trying to empirically test the pilot paradox on a larger scale, by means of Qualitative Comparative Analysis. This research methodology offers the opportunity to seek for quantitative patterns, while at the same time also considering the context of the units of study in a more qualitative way (Schneider & Wagemann, 2012). Three QCA analyses were carried out individually to examine which conditions contribute to 1) the internal success, 2) the scaling up within the carrying organisations and 3) the scaling out to external organisations and new contexts. By taking on a comparative view, this research additionally tried to identify patterns that pointed towards the pilot paradox, as well as potential pathways to get around the pilot paradox.

The results of this research show that patterns can be identified that indicate a presence of the pilot paradox. When pilots are conducted in a 'safe haven' environment, on a relatively small scale and with the presence of internally connective boundary spanners, the pilot is more likely to reach internal success. However, with regard to the scaling up potential, pilots should to a certain extent stay connected to the carrying organisations. The project design should not be small-scale, but in fact aimed at generating results that are representative for the problem complexity. Furthermore, in light of the scaling out potential towards external organisations and contexts, the results should also be seen as representative and the identification of the broader embedding potential of the innovation should be a central part of the pilot's process design.

All in all, the analyses prove that three dimensions of the pilot paradox have an explicit paradoxical character. However, the comparison of the various QCA analyses also suggest potential pathways that can be followed in order to try to get around the pilot paradox. First, the process design dimension shows that an externally oriented process design does not

necessarily inhibits the chances of the pilot to reach the internally desired innovation. Also, a resource distribution that is composed out of regular budgets and other regular resources, does not inhibit the pilot's ability to become internally successful. Furthermore, analysis of primary and secondary data shows a potential strategy to use a phase-based approach, where various phases of the pilot process have a varying character in light of the pilot paradox. A new technical innovation can for example be developed in a small-scale 'safe haven' environment, but the testing and monitoring can be done in a new phase, on a participatory basis and with the involvement of potential future users. Also, the results of this research suggest a potential role for internal boundary spanners in terms of the scaling up potential. These boundary spanners can use their external network where they can serve as ambassadors of the pilot's results, to aim for a broader impact of the developed innovation. All in all, dealing with the pilot paradox requires a comprehensive balancing act by the initiator of the pilots. When the aim is to scale up the results of the pilot, the composition of a strategic scaling up plan can help to realize this desired broader impact. There is no universal formula for success, but there are influenceable elements that can help to get around the pilot paradox and realize a pilot that is both internally and externally successful.

Table of contents

1	. Introduction to the research	8
	1.1 Context and problem statement	8
	1.2 Research aim and -questions	8
	1.3 Scientific relevance	. 10
	1.4 Societal relevance	. 10
	1.5 Thesis outline	. 11
2	. Theoretical overview	. 12
	2.1 Innovation in climate adaptation planning	. 12
	2.2 The nature of pilots as tools for experimenting and innovation	. 13
	2.3 Scaling up pilots: how and what?	. 15
	2.4 The pilot paradox: conditions for internally- and externally successful pilots	. 17
	2.5 Conceptual framework	. 21
3	. Drought context in The Netherlands: from national policy to local pilots	. 24
	3.1 Water management in The Netherlands	. 24
	3.2 Regional drought in the Dutch policy context	. 24
4. Methodology		. 28
	4. I Main research strategy	. 28
	4.1 Main research strategy4.2 Ontological, epistemological, and methodological foundations of QCA	. 28 . 28
	 4.1 Main research strategy	. 28 . 28 . 30
	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 	. 28 . 28 . 30 . 32
	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 	. 28 . 28 . 30 . 32 . 32
	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 	. 28 . 28 . 30 . 32 . 32 . 33
	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 	. 28 . 28 . 30 . 32 . 32 . 33 . 33
	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 4.8 Research ethics 	. 28 . 28 . 30 . 32 . 32 . 33 . 33 . 35
5	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 4.8 Research ethics Operationalization 	. 28 . 28 . 30 . 32 . 32 . 33 . 33 . 33 . 35 . 36
5	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 4.8 Research ethics Operationalization 5.1 Outcome variable: internal success. 	. 28 . 28 . 30 . 32 . 32 . 33 . 33 . 33 . 35 . 36 . 36
5	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 4.8 Research ethics Operationalization 5.1 Outcome variable: internal success 5.2 Outcome variable: external success 	. 28 . 28 . 30 . 32 . 32 . 33 . 33 . 35 . 36 . 36
5	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 4.8 Research ethics Operationalization 5.1 Outcome variable: internal success 5.2 Outcome variable: external success 5.3 Condition variables: internal success 	. 28 . 28 . 30 . 32 . 32 . 33 . 33 . 35 . 36 . 36 . 36 . 37
5	 4.1 Main research strategy 4.2 Ontological, epistemological, and methodological foundations of QCA 4.3 Case selection 4.4 Condition selection 4.5 Data collection 4.6 Data analysis 4.7 Validity and reliability of the research 4.8 Research ethics Operationalization 5.1 Outcome variable: internal success 5.2 Outcome variable: external success 5.3 Condition variables: internal success 5.4 Condition variables: external success 	. 28 . 28 . 30 . 32 . 32 . 33 . 33 . 35 . 36 . 36 . 36 . 37 . 38

6.1 Single and multiple response test			
6.2 Data preparation: internal success			
6.3 Data preparation: external success 40			
6.4 Data preparation: pilot paradox conditions41			
7. Results and analysis			
7.1 Case characteristics			
7.2 Analysis of internal and external success			
7.3 Analysis of conditions for internal success			
7.4 Analysis of conditions for external success: scaling up			
7.5 Analysis of conditions for external success: scaling out			
8. Conclusions and discussion			
8.1 Answering the research questions 69			
8.2 Discussion: searching for pathways and putting the results in perspective			
8.3 Recommendations and further research73			
Bibliography75			
Appendix A) Outline of the survey			
Appendix B) Single and multiple response test			
Appendix C) Reliability test results			
Appendix D) Overview of the minimal formulas that were not included in the results and analysis chapter			

1. Introduction to the research

1.1 Context and problem statement

Over the last decades, water management in the Netherlands has been characterized by a focus on flood risk management and the protection of urban populated areas against flood damages (Ozerol et al., 2016). After the fatal flood of 1953, the Dutch government expanded its water management approach considerably, and aimed it at implementing structural measures for flood protection. This resulted in a relatively stable regime focused on flood risk management (Lintsen, 2002; Van Buuren, Ellen & Warner, 2016). Dutch water management field nowadays continues to face new and more complex challenges, with climate change leading to sea level rise and higher river discharges (Van Buuren et al., 2016). Contrary to the increasing national threat of flooding lies another problem related to climate change that is expected to be increasingly problematic in the future: temporal droughts (Philip, Van der Wiel, Wanders & Van Oldenborgh, 2020). During the summers of 2018, 2019 and 2020, high temperatures and low precipitation values caused periods of drought which had negative consequences for multiple sectors. A major impact can be found in the agricultural sector, where crop yields and grasslands were impacted due to the precipitation shortfall. But impacts were also found in a broader sense, which resulted in an estimated economic damage of 450 to 2080 million euros (Philip et al., 2020). Jeuken et al. (2012) show that next to short-term economic losses, droughts can lead to a significant long-term economic impact. Therefore, government authorities on multiple geographical levels increasingly consider droughts and freshwater shortages as first order problems, where it was formerly seen as more of a second order problem (Vidaurre et al., 2016). In the process of dealing with this problem, governments are seeking for public-private partnerships with local or regional farmers, entrepreneurs, and nature organisations to find a way to better manage freshwater supply and demand in periods of drought. Over the last years, a number of joint droughtadaptation pilots were designed and set up in the areas that are considered most vulnerable for drought. These pilots are mostly small scale and have an innovative character aimed at gathering new types of knowledge (STOWA, n.d.). However, in order to have a larger impact and to realise a sustainable soil- and water system in the future, the pilots need to be scaled up. This proves to be a challenge, considering that pilots and living labs sometimes develop under certain supportive conditions that are not present when the pilot is scaled up. Van Buuren et al. (2018) describe this phenomenon as the 'pilot paradox'. They argue that the conditions that determine the internal success of the pilot, prove to be barriers for upscaling of these projects (Van Buuren et al., 2018). This research therefore focuses on the presence of specific conditions for internal and external success in drought adaptation pilots.

1.2 Research aim and -questions

The aim of this research is to gain knowledge about specific conditions or combinations of conditions that could lead to a drought adaptation pilot that is internally and/or externally successful. Scientific theories about the barriers and conditions that play a role in the degree to which a pilot is successful are tested by means of a QCA analysis of an intermediate to large range of drought adaptation pilots. It is important to mention that specific contextual

factors can always have an influence on the degree to which a certain pilot is successful in the broader uptake. To take this context into account and to gain as much in-depth knowledge as possible, this research uses a mixed qualitative-quantitative approach. The research is organised around the following main question and sub-questions:

Main question

What conditions influence the internal and external success of drought-adaptation pilots in dry rural areas in The Netherlands?

The main question can be split up into smaller, more empirically focused sub-questions:

1. Which pilots are regarded as internally and/or externally successful and what is the nature of these drought-adaptation pilots in the Dutch context?

This sub-question focuses on determining the outcome variable or success factor. To answer this question, indicators for successful pilots and successful scaling up need to be determined. These indicators result from literature and will be tested by several questions within the survey that was distributed among involved actors of various pilots, together with additional secondary information sources (evaluation reports or desk research results), in order to determine which pilots are internally and/or externally successful. Also, the context of the drought adaptation pilots is discussed. In a holistic manner, patterns and highlights in the descriptive data are sought, in order to present an encompassing overview of the nature of the various pilots.

2. Which conditions contribute to the internal success of the drought adaptation pilots? This question is aimed at researching which conditions influence the extent to which the pilots were able to reach their previously stated main goals and ambitions, which are often aimed at realising innovation. Several conditions result from literature, that will be questioned by a survey among involved actors from the various pilots. The influence of the conditions will afterwards be tested by means of a QCA analysis.

3. Which conditions contribute to the scaling up of the drought adaptation pilots? This question focuses on researching which conditions influence the extent to which the pilot(s results) lead to broader effects or are being scaled up within the area of the initially involved organisation(s). Several conditions result from literature, that will be questioned by the survey among involved actors from the various pilots. Afterwards, a QCA analysis will be performed to research which (combination of) conditions lead to successful scaling up.

4. Which conditions contribute to the scaling out of the drought adaptation pilots? This question focuses on researching which conditions influence the extent to which the pilot(s results) lead to broader effects or are being scaled up outside the area of the initially involved organisation(s). Several conditions result from literature, that will be questioned by the survey among involved actors from the various pilots, after which a QCA analysis will be performed to research which (combination of) conditions can lead to successful scaling out.

5. To which extent do the conditions for internal success collide with the conditions for external success?

This last sub-question is aimed at comparing the outcomes of the two QCA analyses, in order to research to what extent the pilot paradox can be recognized in the case of drought adaptation pilots in Dutch dry rural areas. By comparing the results of the two QCA analyses, the recognized patterns are discussed and translated into potential strategies for developing pilots that can be internally and externally successful at the same time.

1.3 Scientific relevance

This research focuses on empirically testing conditions for successful upscaling of pilots and the presence of the 'pilot paradox' in drought adaptation pilots in the Netherlands. Where Van Buuren et al. (2018) identified the phenomenon and qualitatively tested the presence of the phenomenon in six cases, this research tries to take a next step in empirically testing the presence of the paradox on a larger scale. By using Qualitative Comparative Analysis (QCA), the research focuses on searching for (combinations) of conditions that contribute to the internal and external success of pilots and tries to identify the possible presence of the pilot paradox on a larger scale. Especially in the field of spatial planning, there is a lack of empirical studies focused at testing the success or effectiveness of interventions (Brody & Highfield, 2005). Moreover, there is a lack of empirical studies into the scaling up of pilots and the way pilots can have influence on the broader policy regime (Vreugdenhil, Taljaard & Slinger, 2012).

QCA offers the opportunity to conduct empirical research without neglecting the specific context of the cases. This holistic approach suits the current demands of sustainable development planning, which requires holistic and integrated policy, action, and evaluation of outcomes (Nadin, 2007). Next to the conditions identified by Van Buuren et al. (2018), the research will also try to integrate conditions for successful upscaling that are identified by other scholars in the literature review. The qualitative dimension of QCA offers the opportunity to also take a closer analytic look into the context of the pilots and the way in which the conditions potentially influence the success by leaving room for suggestions and further in-depth explanations from the research participants.

1.4 Societal relevance

The Dutch water management sector faces a dual challenge in dry rural areas. First, periods of drought are expected to occur more often in the future. Also, peak rainfall is likely to increase as a result of climate change. The current soil- and water system in Dutch dry rural areas is not capable of dealing with these phenomena (STOWA, n.d.). The Dutch national government has therefore provided investment packages aimed at improving fresh water supply and division in the dry rural areas (Rijksoverheid, 2020). Dutch water authorities and local and regional governments are looking to gain knowledge about spatial, societal and physical consequences of climate change for their region and ways to realise adaptation measures in order to reduce the negative consequences. The KLIMAP programme focuses on this topic by developing climate development pathways to reach a more sustainable soil-and water system in Dutch dry rural areas (STOWA, n.d.). Within the KLIMAP programme

many different parties are involved, including Wageningen Environmental Research, Deltares, regional water authorities, provinces, Radboud University, Wageningen University, and others (STOWA, n.d.). Innovative drought adaptation measures are tested in pilots and living labs. To reap as many benefits as possible from the innovation in these pilots, the upscaling challenge needs to be considered. This research focuses on pilots that are finished, in order to learn about conditions for internal and external success of the projects. This knowledge can eventually be implemented in the design of new pilots in the KLIMAP programme.

1.5 Thesis outline

This report consists of eight chapters, of which this introduction chapter is the first. Chapter two provides an overview of relevant scientific concept and literature about, among others, the role of pilots in innovation in climate adaptation planning and the conditions that can contribute to the success of these pilots. After this theoretical discussion, chapter three discusses the drought adaptation policy context in The Netherlands on different geographical scales. Afterwards, chapter four discusses the chosen research strategy and methodology. Also, reliability-, validity- and ethical related considerations are discussed in this chapter. Additionally, chapter five step-by-step describes the operationalization process of the various theoretical dimensions into researchable elements that can be questioned by means of survey questions. This is, therefore, the tipping point in this research report. From chapter six onwards, the survey results are gathered and analysed. This process starts in chapter six with the description of data preparation for the QCA analyses. In this chapter, the cases are attributed values based on the survey answers and secondary data analysis. Afterwards, chapter seven step-by-step describes the three various QCA analyses that were performed. Finally, chapter eight presents the main conclusions of this research and puts these conclusions in their context in the discussion section. This final chapter also makes recommendations for further research.

2. Theoretical overview

This chapter provides an overview of the relevant scientific literature with regard to the role and challenges of pilots in climate adaptation planning. First, the role of innovation in climate adaptation planning is discussed. From this broad innovation basis, in paragraph 2.2 the nature of pilots as a tool for innovation is described. The third section analyses literature about scaling up and diffusion of pilot results. Subsequently, the fourth part of this chapter reflects on the main theoretical framework of this research: the pilot paradox and the conditions for internal and external success. This chapter results in the conceptual framework that is provided in section 2.5 and will be used for the structuring of the research.

2.1 Innovation in climate adaptation planning

The climate adaptation planning field is characterized by high dynamics. Local and regional governments all over the world are increasingly facing effects of climate change and are taking steps to adapt to hazard risks. To adapt to climate change issues, they must identify and assess risks, develop adaptation plans, implement these plans, and monitor and evaluate them (Shi, Chu & Debats, 2015). Here, governments are often restricted by pathdependencies and traditional planning practices. The ability of cities or regions to adapt to climate change issues is called their 'adaptive capacity' (Engle, 2011). This capacity depends on many underlying factors, such as institutions, governance, economic structures, and management (Engle, 2011). The adaptive capacity is also linked to the room for innovation. Aase, Chapagain & Tiwari (2013) refer to this as the 'innovative capacity', which focuses on the capability to master and implement new design features, "irrespective of whether they are new to their competitors, their country or the world" (Aase et al., 2013). When something can be regarded as 'innovative' is highly debated in scientific literature. Something can be called an innovation based on, for example, its degree of "newness". This degree of "newness" is hard to identify, because a certain technique can be new to certain (groups of) people, while other (groups of) people already knew about the technique (Varis & Littunen, 2010). Moreover, innovation can mean different things for different people. Where academics count new knowledge or new frameworks as innovative, for entrepreneurs it needs to make money before it can be regarded as innovative (Massa & Testa, 2008; Varis & Littunen, 2010). Broadly speaking, innovation literature distinguishes two types of innovation. The first is focused on the object of change, which can be for example product, process, or market innovation (Schumpeter, 1934; Varis & Littunen, 2010). Within climate adaptation planning, innovations can have multiple characters. Technical innovations are the most common type of innovation. Especially with regard to drought adaptation planning in dry rural areas, the development of new techniques and smart technologies have proven to be of large importance to realise innovation. The most obvious example is the development of new irrigation techniques to cope with regional droughts and to irrigate more efficiently (Smithers & Blay-Palmer, 2001). However, the object of change in climate adaptation innovation can also be social. Especially the relation between climate innovation and social inequalities and taking locally based actions to change current practices sees the object of change as more social than technical (Rodima-Taylor, 2012). Another object of change can be institutions, when institutional arrangements change internally or when new institutional arrangements

are set up (Rodima-Taylor, 2012). Bauer & Steurer (2014) additionally see partnerships and policies as the object of change in climate adaptation planning. The establishment of new partnerships are expected to result in collaborative advantages for the partners, while at the same time-sharing risks and resources results in broader policy innovation (Bauer & Steurer, 2014). The second differentiation in innovation focuses on the earlier described degree of "newness" or "radicalness", where a certain innovation can be classified as something that is entirely new for everyone, or just an improvement of already existing technologies (Varis & Littunen, 2010).

Innovation inherently goes with challenges and risks. Hunter, Cassidy & Ligon (2012) state that in planning for innovation, the majority of ideas will be subject to failure. Because this failure results in financial losses for the leading organization, a sound plan for monitoring and learning from failures should be in place (Hunter et al., 2012). Additionally, this is where pilots often come into the picture. Pilots are seen as tools that can help to realize innovation, oriented towards policy testing and the exploration of new solutions for problems in the social-physical environment. Because of the often-small scale, temporal character and clear boundaries of pilots, the potential risks that a certain innovation fails can be managed better (Van Buuren et al., 2018). The nature of different types and roles of pilots in innovation processes will be elaborated on in the next section.

2.2 The nature of pilots as tools for experimenting and innovation

Overall, pilots can help to limit the impact of change, thereby reducing risks and providing a prototype for new methods and practices. They can therefore be strategically used to " unfreeze the status quo and prepare the ground for a shift towards open innovation" (Turner & Muller, 2003; Boscherini, Chiaroni, Chiesa & Frattini, 2010). In scientific literature, different types of pilots, or experiments, are distinguished. These different types of pilots also have different theoretical backgrounds in literature. Sengers, Wieczorek & Raven (2016) provide a comprehensive overview of the differences between types of pilots in the sustainability transitions literature. They distinguish five types of pilots: niche experiments, bounded sociotechnical experiments, transition experiments, grassroots experiments, and sustainability experiments (Sengers et al., 2016). Niche experiments are founded in Strategic Niche Management (SNM) and are being developed as market niches in order to eventually realize broader regime shifts. It is being argued that these experiments indirectly influence transitions but do often not lead to big successes that highly influence the regime (Hoogma, Kemp, Schot & Truffer, 2002; Sengers et al., 2016). Bounded socio-technical experiments have a more social character that developed as criticism of the technical focus of SNM experiments. It puts emphasis on the social experiment and the role of the civil network (Sengers et al., 2016). Transition experiments are more radical in their ambitions, and proactively aim for the exploration of radical new solutions for societal problems (Van den Bosch & Rotmans, 2008). These experiments inherently focus on deepening and learning about conditions, broadening the scope, and learning from other pilots, and scaling up as main parts of the pilot design and also have a strong normative dimension (Sengers et al., 2016). Fourth, grassroots experiments have a bottom-up character and respond to a certain local situation of a community. These experiments are often more informal and also have a normative character

(Sengers et al., 2016). Finally, sustainability experiments are planned and are aimed towards substantial sustainability gains (Berkhout, Angel & Wieczorek, 2009). These types of experiments can also be seen as an overarching category, that can include elements of various other types of experiments (Wieczorek, Raven & Berkhout, 2015).

Next to this categorization of experiments based on primarily scientific movement theories, Vreugdenhil et al. (2012) make a distinction in three broad pilot types that is more based on what the pilot aims to achieve. This distinction consists of: research pilot projects, management pilot projects and political-entrepreneurial pilot projects. These categories all have various uses. Research pilot projects are mostly guided by exploration or evaluation uses, for example when a certain new innovation needs to be tested or evaluated on a small scale. Management pilots often focus on communication or triggering dialogue, problem mitigation, policy implementation or overcoming insurance issues by testing on a small scale (Vreugdenhil et al., 2012). Finally, political-entrepreneurial pilots are used as an incentive or advocacy tool, but can also be used in a strategic political game with commercial interests or hidden intentions (Vreugdenhil et al., 2012).

These different uses also influence which phases are passed through in the process. As shown in figure 2.1, the typical project view influences the scope of the pilot process. A pilot that is used for a political game often focuses on designing intervention and implementation. However, a pilot for a societal innovation process comprises all of the six phases, from initiation to diffusion (Vreugdenhil et al., 2012).



Figure 2.1: Pilot project activities and typical project views on inclusion of activities for different pilot uses

Note. Reprinted from "Pilot projects and their diffusion: a case study of integrated coastal management in South Africa", by Vreudgenhil, H., Taljaard, S. and Slinger J.H., 2012, *International Journal for Sustainable Development*, *15*(1/2), 155.

This research uses an inclusive project view for the society-wide transition to a climate adaptive society, therefore also embracing diffusion challenges as a part of the pilot project. More concretely, this means that pilots aimed at climate adaptation mostly aim for diffusion of their innovations to non-partners of the pilot or society as a whole (Bauer & Steurer, 2014; Van den Broek, Van Elzakker, Maas & Deuten, 2020). This diffusion, or scaleup, is seen as required to realize actual adaptive change, but also is associated with difficulties. The different ways of diffusion and the associated challenges will be discussed in the next sections.

2.3 Scaling up pilots: how and what?

Hartmann & Linn (2008) define the term 'scaling up' as "expanding, replicating, adapting and sustaining successful policies, programs or projects in geographic space and over time to reach a greater number of people" (Nair & Howlett, 2015). Different authors have different categorizations of how scaling up can take place. For example, Van Winden & Van den Buuse (2017) distinguish three categories: roll-out, expansion and replication. Roll-out happens when an actor decides to "use the pilot's test results to scale up [...] or apply the lessons of the experiment within their own organization" (Van Winden & Van den Buuse, 2017). Expansion happens when the pilot is expanded by involving new partners or users or extending the geographic range. Finally, replication occurs when it is implemented in another context (e.g. a different geographical area or organization) (Van Winden & Van den Buuse, 2017). The report "voorbij lokaal enthousiasme" uses a four-fold distinction of: growth, replication, circulation, and institutionalisation (Van den Broek et al., 2020). Growth happens when the network of the pilot expands, and the pilot becomes more embedded at different locations. Replication is defined comparably to van Winden & van den Buuse (2017). Circulation means that elements of pilots circulate to other parties. Finally, institutionalisation occurs when a pilot's outcome leads to a fully embedded innovation. This means that the innovation finds its way into regime policy and is also partly a social-political process (Van den Broek et al., 2020). Vreugdenhil et al. (2012) also see institutionalisation as the most comprehensive, and most challenging, form of upscaling. Furthermore, they define expansion as: scaling up the scale dimensions of time, space, and problem scope. This results in the inclusion of more administrative bodies than in the prior pilot. Next to these two forms of upscaling, they see the replication of the pilot project to other pilots or to small scale management projects as 'diffusion', rather than actual scaling up (Vreugdenhil et al., 2012). This is however just the tip of the iceberg with regard to the variety of different definitions that are mentioned in scientific literature. There are also authors who try to narrow this variety of different definition down into several main categories of scaling up. For example, Van Doren, Driessen, Runhaar & Giezen (2018) make a distinction in horizontal and vertical scaling up. Horizontal scaling up refers to the spatial growth of (parts of an) innovation, which results in a larger impact of the innovation. This horizontal growth contains different previously mentioned categories: replication, growth, circulation, roll-out and expansion. Vertical scaling up, however, applies when a pilot results in more structural knowledge development or structural policy related scaling up (i.e., actual policy change). This refers to the actual structural institutionalisation of innovations in regular policies. This vertical scaling up often also has a social-political character (Van den Broek et al., 2020). Gillespie (2004) therefore defines this form of scaling up as 'political scaling'.

Hughes, Yordi and Besco (2020) provide another view on the upscaling of pilots. They acknowledge that the process of scaling can take place by different means (e.g. technical, financial, institutional or political) but in addition make a distinction in the 'scaling up' and 'scaling out' of pilots. Scaling up is being defined as "widening of the scale of operation" (Hughes et al., 2020; Douthwaite, Kuby, Van de Fliert & Schulz, 2003). More specifically, this means moving from a contained pilot to an area-wide program. Scaling out is seen as the expansion of the pilot to new geographical areas or other governmental levels. The outcome that is pursued in scaling out is therefore a broader shift in the regime of sustainability or climate resilience, which is more uncertain and associated with more difficulties (Hughes et al., 2020). Figure 2.2 shows a theoretical model, as developed by Hughes et al. (2020). Their focus is cities, but this can be applied to other geographical areas as well.

Figure 2.2. A theoretical framework on the function of pilots in policy innovation and processes of scaling up and scaling out of these pilots



Note. Reprinted from "The role of pilot projects in urban climate change policy innovation", by Hughes, S. and Yordi, S. and Besco L., 2020, *Policy Studies Journal, 48*(2), 278.

Additionally to the scientific debate on when and how certain results of pilots are scaled up, the question 'what is being scaled up?' is also being discussed in literature. The object of diffusion occurs in different types and relates to the extent to which the pilot can bring about change. For example, pilots can diffuse new practices that alter the physical environment or the social environment (Vreugdenhil, 2010). Within the field of water management, changes in the physical environment nor to example when freshwater reservoirs are constructed. Changes in the social environment occur when new partnerships or cooperation networks are set up. Another object of diffusion is knowledge. The diffusion of knowledge happens through so-called 'learning effects'. McFadgen & Huitema (2016) make a distinction in participant learning and political learning. Participant learning has a more small-scale character and focuses on the pilot group itself. This also means that participants in the group learn from each other and have the possibility to learn about the opinions of others. Political learning, however, focuses on the learning processes that the pilot causes at policy makers (McFadgen & Huitema, 2016).

Research shows that the process of scaling up or out often does not run smoothly. Many factors can restrict the upscaling potential of pilot projects (Van Winden & Van den Buuse,

2017). Van Buuren et al. (2018) describe the difficulty of scaling up pilot projects as 'the pilot paradox'. This term indicates that " *the conditions that are necessary to give a pilot room to experiment and to learn [...] also seem to constitute the main barriers to the broader uptake and translation of its results into changes in policy goals, content or instruments*" (Van Buuren et al., 2018). What these conditions and/or barriers are according to scientific literature, will be examined in the next section.

2.4 The pilot paradox: conditions for internally- and externally successful pilots

Van Buuren et al. (2018) distinguish between the internal and external success of pilots. Internal success refers to the pilot's ability to realize its main ambitions, which can be technical innovations or new cooperation mechanisms. The external success is about the extent to which the pilot's outcomes are being implemented in other contexts, a larger area or even lead to a broader transition in the policy regime (Van Buuren et al., 2018; Van Buuren et al., 2016).

There are multiple factors that have an influence on the internal success of pilots. First, pilots are regarded as 'safe havens' distant from everyday practices, and therefore leave room for experimentation (Van Buuren et al., 2018; Hargadon & Douglas, 2001). Moreover, there are often many resources available for piloting, which together with the limited scale results in a limited number of risks. During the pilot, additional resources can become available to stimulate innovation. The limited scale of pilots also means that they have boundaries that can be clearly defined, and that risks or danger of negative external effects are reduced (Sanderson, 2002; Van Buuren et al., 2018). Additionally, for pilot projects to become successful, the presence of boundary spanners is essential, according to Van Buuren et al. (2018). Boundary spanners can be defined as people who are capable to bridge the gap between different problem conceptualizations, who are able to change their cognitive vision as a result of potential new insights and "who are able to link the organization they represent with its environment" (Van Meerkerk & Edelenbos, 2014).). This second part of the definition also hints on a possible role for boundary spanners with regard to the external success of pilots, as ambassadors of the innovation towards potential future users or application areas. Van Buuren, Vreugdenhil, Verkerk & Ellen (2016) researched multiple pilots within the field of climate adaptation and concluded that these boundary spanners are often present in these projects. The presence of boundary spanners subsequently relates to the role of leadership in (policy) innovations. There is a broad scientific debate about the role and dimensions of leadership in innovation. Overall, leadership is not defined solely by singular, managerial leaders, but can instead best be regarded through the interplay of different leadership functions. This means that leaders should be part of the network in projects, and their degree of support and commitment can potentially have a high influence on the success of innovation (Nam & Pardo, 2011). It also means that multiple actors in a (pilot) project network can perform different leadership functions. One actor can for example 'lead' by connecting multiple other actors to each other, while another actor can lead by ensuring that there is sufficient space for innovation or by disseminating the innovation back to regime policy (Meijerink & Stiller, 2013). This also shows that leadership can be influential for the connection between the pilot and the regime and thereby also for the external success of the pilot. This is also mentioned by Hartmann & Linn (2008), who describe that a charismatic leader or champion with a strong

vision and connective abilities can influence the external success of pilots. They have the connections to major stakeholders and government institutions to realize broader uptake of the innovation in the policy environment (Hartmann & Linn, 2008; Samoff, Sabatane & Dembélé, 2013).

Next to these boundary spanners, the pilot should contain competent participants (Van Buuren et al., 2018). Friedrich, Karlsson & Federley (2013) add to this that the participants need to be active partners during the whole process. Finally, the project design of pilots needs to include an open dialogue, and thereby enable processes as social learning and principled engagement (Van Buuren et al., 2018). Social learning processes are also essential to frame the common problem understanding and agreement on collective actions that have to be taken to overcome this problem. This common problem and goal understanding needs to be shared among all participants (Friedrich et al., 2013). The learning processes also relates to ex ante, ex durante and ex post knowledge creation by means of monitoring. The extent to which and the way in which monitoring takes place during the pilot process can determine the chance of reaching the end goals. Especially in innovative pilot projects, monitoring intensively can help to timely make adjustments that steer the pilot in the right direction (Martin & Sanderson, 1999; Vreugdenhil et al., 2012). Cuéllar-Gálvez, Aranda-Camacho, & Mosquera-Vásquez (2018) state that this also applies from a technical perspective, where continuous evaluation helps to improve the chance to reach technical goals by introducing timely modifications. This all takes place in an open and communicative environment (Muro & Jeffrey, 2008; Vreugdenhil, Slinger, Thissen & Rault, 2010). The open and communicative environment also means that power structures need to be transparent. All roles and procedures for decision-making need to be clarified (Friedrich et al., 2013).

The external success is also influenced by multiple criteria. The first condition is 'normative congruence', which is about the position of the pilot towards the values and beliefs of the regime as a whole (Hargadon & Douglas, 2001; Van Buuren et al., 2018). This needs to be bridgeable to successfully upscale the pilot (Van Winden, 2016). Van den Broek et al. (2020) refer to this as the degree of 'social-cultural embeddedness', which also considers the broader values, routines and views. Especially controversial innovations have to consider this dimension of upscaling (Van den Broek et al., 2020). The compatibility of the pilot's outcomes with the standard operating procedures (SOP) in public administrations and the extent to which the pilot can be used as instrument for policy change also matters (Hargadon & Douglas, 2001; Van Buuren et al., 2018). This is related to the embeddedness in laws and regulations, standards, and protocols. In order to implement the pilot's outcomes, these documents often need to be adjusted (Van den Broek et al., 2020). Hughes et al. (2020) state that scaling up is a process that needs to consider political and institutional dimensions. Additionally, lack of political will and institutional mismatches are seen as two of the main bottlenecks for scaling up (Davis, 2004). There is also a technical dimension to the compatibility of the pilot's outcomes with standards. Innovative outcomes need to be connected to existing technological infrastructure, which can demand adjustments to these existing systems. In other words: the pilot needs to be technologically embedded in order to be able to scale up (Van den Broek et al., 2020). This technological embeddedness is related to the perceived risks and the perceived technical maturity. This requires monitoring and

assessment by experts that can determine whether upscaling of the innovation is possible from a technical perspective (Cuéllar-Gálvez et al, 2018; Rae, Kerr & Maroto-Valer, 2020). Many innovations are project- and context specific. This means that when a certain innovation 'works' on one location (i.e. in the pilot), the desired effect on another location or on a larger scale cannot simply be assumed to work (Rae et al., 2020). Also, the innovation can produce negative consequences on other regions or can be less effective in other contexts. This means that in such a technical feasibility assessment, potential risks and externalities should also be included (Taylor, Harman & Inman, 2013). In the case of drought adaptation pilots, this means that the innovation that is being tested or developed in a pilot should also lead to significant freshwater savings and should have no negative externalities in terms of unwanted new flooding risks or unwanted freshwater surpluses. This requires a physical and technical assessment of the innovation and the potential consequences of implementation on a larger scale by a technical expert.

Naber et al. (2017) discuss the importance of learning processes in the pilot's design. To facilitate diffusion and upscaling, learning processes are required that focus on the alignment between 'the technical' and 'the social'. These learning processes should be reflexive and consider underlying assumptions and values (Naber et al, 2017). For example, to support the legal embeddedness of the pilot's outcomes, the involvement of legal experts can help to gain a common understanding among the participants about how to deal with possible legal issues when upscaling will take place (Van den Broek et al., 2020).

It is seen as necessary to organise learning and reflection during the execution of the pilot to enable timely adjustment of the pilot's aims, to prevent goal displacement. This occurs when the pilot is successful in realising an innovative outcome but fails to keep the wider environment connected (Van Buuren et al., 2018). A potential strategy to prevent goal displacement is to already involve future users of the innovation in the experimental phase in order to ensure trust and credibility. This is referred to as the intertwinement of upscaling and social inclusion. A diverse group of local actors contributes to the ability of the experiment to eventually scale up (Dijk et al., 2018; Ryghaug et al., 2019). Related to this, Ingram et al. (2015) argue that a large degree of interaction and a strong connection between the pilot and the regime is necessary to ensure that the regime assesses the pilot results as legitimate (Van Buuren et al., 2016; Van Buuren et al., 2018). The involvement of future users in the pilots can also help to learn about the market potential for the pilot's outcome.

Additionally, Davis (2004) considers insufficient funding plans as one of the major bottlenecks for scaling up. Pilots need to become economically embedded by establishing a viable business model and connecting to the demands of the market (Van den Broek et al., 2020). In this process, a resource distribution that is composed out of regular budgets and means can contribute to the scaling up potential of pilots. When there are too much temporary innovation funds, the resource distribution can be regarded as not representative (Van Buuren et al., 2018). Drought adaptation pilots have much diversity with regard to their economic feasibility and the creation of a viable business case. Pilots that focus on high-quality crops are mostly considered economically feasible, but realizing new types of irrigation for standard crops is often not economically viable. In this case, seeking for investors or public subsidies could be an alternative strategy (Ministerie van Infrastructuur en Waterstaat, 2019). This

shows that the business case does not necessarily need to be an economic one, but can also be a societal business case. When pilots are mainly financed by temporary innovation funds, it is important to actively seek for additional financial means that can be claimed when the innovation is scaled up (Franzel, Cooper & Denning, 2001).

Moreover, the articulation of expectations should be clear, robust, and grounded (Ryghaug et al., 2019). Ambitions and expectations must be clearly articulated and shared by the members. Especially when the pilot is aimed at highly innovative practices or new technologies, expectations should be substantiated by sub-experiments and scientific research (Naber et al, 2017; Ryghaug et al, 2019). Finally, the parties that have to adopt the results, need to see them as relevant, reliable, representative, and useful (De Moor et al, 2010; Van Buuren et al., 2018).

In line with the formerly discussed pilot paradox, the conditions for internal success and external success have a contradictory character. Van Winden and Van den Busse (2017) argue that some pilots fail to scale up because they are protected from real world legislation, market forces and norms too much and therefore not represent the real world enough. As a result, the pilot can suffer from regime alienation, where the distance to the mother organisations becomes greater over time, as the pilot focuses mainly on the internal learning processes (Van Buuren et al., 2018). Also, the large number of additional resources present during the pilot phase is often not available when the project is being scaled up (Sanderson, 2002; Hasluck, 2000). Van Buuren et al. (2018) argue that an important strategy to overcome the pilot paradox is to make sure that the broader uptake of the pilot's results is already one of the main targets in the pilot's design. The complete pilot paradox can be found in figure 2.3. This figure shows the five dimensions of the pilot paradox that both influence internal and external success in a different, contradictory way.

Dimensions	Internal success condition	External success condition
Position of the pilot	Safe haven at a distance	Congruence with carrying organisations
Project design and scope	Limited scale and reduced risks	Representativeness and generalizability of outcomes
Resource distribution	Sufficient number of resources for innovation	Results fit in existing resource distribution
<u>Participants</u>	Coalition of leading boundary spanners	External representativeness of stakeholders
<u>Process design</u>	Open design and mutual learning environment	Broader embedding potential is identified

Figure 2.3: Dimensions of the pilot paradox: internal vs. external

Note. Adapted from "The Pilot Paradox: Exploring Tensions between Internal and External Success Factors in Dutch Climate Adaptation Projects", by Buuren, A. van., Vreugdenhil, H., Van Popering-verkerk, J., Ellen, G.J., Van Leeuwen, C. & Breman, B., 2018, *International Journal for Sustainable Development*, *15*(1/2), 155.

2.5 Conceptual framework

Figure 2.5 provides a schematic representation of the conditions that, according to the literature, influence internal success and external success. Within the external success, the division between diffusion within the area or organization, and scaling out or institutionalization is made (Hughes et al., 2020; Vreugdenhil et al., 2012). This framework will be used to structure the research. The dimensions of the pilot paradox serve as a basis for this framework. Several conditions from other authors that were discussed in the literature review were merged into these dimensions when they showed a link to these dimensions.

The grey boxes represent the different dimensions of the pilot paradox, as was previously shown in figure 2.3. The white boxes summarize the internal and external sides of the paradox. These short statements show the main point that is mentioned in literature. This will be further operationalized into different elements in chapter five. The framework in figure 2.4 shows the conditions that potentially lead to an internally or externally successful pilot, indicated by the arrows in the figures. These complex causalities are studied in the context of the pilots, which is possible due to the holistic character of QCA research (Berg-Schlosser et al., 2009). The loop on the left and the right side of the model emphasizes the complex causality and the possibility of interaction and interdependence of the different variables. Eventually, the relation between the dimensions itself will also be researched, because the patterns that are sought often involve subsets in which different dimensions can operate together to realize a successful outcome.

Figure 2.4: Conceptual framework



Figure 2.4 shows the dimensions of the pilot paradox and how they potentially influence the internal and external success of pilots. With regard to the internal success, Van Buuren et al. (2016; 2018) define this as: "*the extent to which the pilot successfully realizes its main ambition*" (Van Buuren et al., 2018). This means that the results of the pilot comply with the goals that were set in the pilot design. Additionally, pilots can be regarded as internally successful when they lead to new cooperation mechanisms between the parties involved (Van Buuren et al., 2016). Another dimension that Van Buuren et al. (2016) discuss is the learning element, which can lead to successful learning outcomes. This also means that a pilot that is not successful in realizing its previously set substantive goals (e.g. testing or developing a new innovative technique), can be regarded successful in terms of gathering 'new knowledge about the fact that a certain technique or innovation does not work in

practice, and needs to be adjusted'. Following this line of reasoning, this would mean that pilots are never internally unsuccessful, because new information is always gathered, also when the innovation does not appear to work. Therefore, this research relates the internal success to the goals that were primarily set in the pilot's design and the different pilot's uses, as described in paragraph 2.2 (Vreugdenhil et al., 2012). The operationalization of this assessment of internal success will be further elaborated on in section 5.1.

As seen in paragraph 2.2, many possible definitions and categorizations of the process of scaling up pilots are given in scientific literature. Van Buuren et al. (2016) define the external success of pilots as "*the extent to which the knowledge or networks developed in the pilot project are being used for new initiatives, both on project and policy level.*" (Van Buuren et al., 2016; Van Buuren et al., 2018). When linking this definition to the broad spectrum of categorizations, as provided in paragraph 2.2, both replication or circulation (on project level) and broader institutionalization (on policy level) are considered in the external success. Hughes et al. (2020) do make this distinction between scaling up in the area of the organization itself (e.g. municipality, water board area) and scaling out into other policy contexts (e.g. other municipalities, or national policy). The categorization that will be followed in this research builds mostly on the definition of Hughes et al. (2020) but also includes elements of other definitions. This research mainly distinguishes the following categories of external success:

- Scaling up within the organization(al area): the pilot leads to some kind of scaling up effect within the carrying organisations or organizational area of the parties that were involved in the initial pilot.
- Scaling out to other organizations: the pilot leads to scaling out effects to other contexts or organizations beyond the initially involved actors, or the pilot's results lead to scaling out effects on a higher scale (e.g. from local to regional, or from regional to national).

Within this categorization of external success, the emphasis is not on the way in which scaling up processes work (for example, whether it is replication or circulation). Instead, the focus lies on the extent to which the results of the pilot were able to impact, change or modify broader regime practices. One could argue that a pilot that is scaled up to other organizations is more 'institutionalized' than a pilot that remained within the original organization. However, in practice, institutionalization, replication, or circulation can all take place in both categories. The difference between the previously mentioned categories lies in the geographical or organizational level of scaling up, not in the character of the scaling up process. This does not mean that this research fully ignores this character of scaling up. The qualitative dimension of the research will try to sketch an image of the pilot's policy context, and will also try to identify the way scaling up has taken place. However, for the QCA analysis, the abovementioned categorization will be followed.

3. Drought context in The Netherlands: from national policy to local pilots

As was described in the introduction, regional droughts have become increasingly problematic in the Dutch dry rural areas over the last years. This chapter sketches an image of the existing policy context with regard to drought issues. It will describe policy documents and policy steps that were taken on different geographical levels. The chapter provides more information about the context in which pilots can take place. First, the context of the Dutch water sector will be sketched, in order to understand the division of roles and (legal) frameworks on which water policy is based in the Netherlands. Afterwards, the focus is on the regional drought issue and the inclusion of this topic in national, regional, and local policies. Also, the role of pilots in innovation programmes in the Dutch context will be discussed.

3.1 Water management in The Netherlands

The Dutch water management sector traditionally focused on the battle against water, and had a technical and sectoral character. Realizing water security was the main priority and the approach to reach water security was guided by hydraulic engineers who developed large-scale hydraulic measures (Van den Broek et al., 2020). In the subsequent decades, more and more negative consequences of this engineering approach surfaced. The large-scale measures negatively affected traditional landscapes and natural ecosystems. The policy document 'Omgaan met water (Dealing with water)' characterised the start of a new approach to water management, that was more integrative because it also included water quality and both groundwater and surface water (Van den Broek et al., 2020). This integrative water management approach was expanded over the following decades, by including more interdisciplinary collaboration and knowledge development. More and more scientific disciplines became part of the water management sector, such as ecology, spatial planning, and law (Van den Broek et al., 2020; Rooy & Sterrenberg, 2000).

The Dutch water system nowadays contains a participatory and collaborative approach to water management (Özerol et al., 2016). The four main actors are the central government, the provinces, municipalities, and water authorities. In the process of drinking water purification, drinking water companies also play a large role. The 'Water Act' (2009) provides the national legal framework for water management in The Netherlands. While the main focus of Dutch water management has always been flood protection, the goals nowadays can be described as diverse (Özerol et al., 2016). This can also be seen in the Delta Programme. The first Delta Programme was established in 2011 and aims at a diverse set of goals, including improving water quality, providing a sufficient fresh water supply, and ensuring a sufficient discharge capacity of the main rivers during peak discharge periods (Rijksoverheid, 2010).

3.2 Regional drought in the Dutch policy context

In the most recent revision of the Delta Programme, an increasing focus on droughts and fresh water supply becomes clear. The Dutch government has enabled an additional investment package of 800 million euros for improving fresh water supply and division in periods of drought. Measures that need to be taken, according to the Dutch government, are

aimed at more efficient water usage and freshwater conservation in winter periods (Rijksoverheid, 2020). The Delta Programme of the Dutch national government is regionally implemented in regional Delta Plans or climate adaptation strategies. One example is the 'Deltaplan Ruimtelijke Adaptatie (Deltaplan Spatial Adaptation)', in which municipalities, waterboard authorities, provinces and the national government constituted seven ambitions to deal with climate change by adapting to its various consequences. This Deltaplan is focused on themes such as flooding and flood related effects, heat stress and drought. Another example is the 'Deltaplan Hoge Zandgronden (Deltaplan High Sandy Soils)', several provinces (Drenthe, Overijssel, Gelderland, Utrecht, Noord-Brabant, and Limburg) work together with regional water authorities, agricultural collectives and nature conservation organisations on measures for sustainable use of freshwater. The first phase runs from 2016 to 2021. Currently the second phase, which will run from 2022 to 2027 is under construction (Provincie Noord-Brabant, n.d.). However, the provinces and waterboard authorities already have constituted several additional policy and implementation programmes in the recent past with regard to the regional drought problem. First, the province of Noord-Brabant constructed a long-term vision and a complementary implementation programme with regard to climate adaptation. The province of Limburg and Brabant also created a shared strategic policy document called 'Strategy climate adaptation Southern Netherlands'. In this document, they present a shared course for the climate-proof design of urban and rural spaces. The province of Overijssel also followed this regional adaptation strategy, and presented their adaptation plan at the end of 2017 (Kennisportaal Klimaatadaptatie, n.d.). In Gelderland, the provincial climate adaptation programme is still under construction. Currently, an action plan was established to make a first step towards more climate adaptation measures on a provincial level (Provincie Gelderland, 2019). On a regional level, coalitions of municipalities and waterboard authorities were precursors on this trend. They already established regional adaptation strategies together with several regional stakeholders (Kennisportaal Klimaatadaptatie, n.d.).

Next to the provincial policy documents, the regional water authorities are the actors who lie at the core of all water related issues in the Netherlands. Most of these government authorities are geographically smaller than provinces and larger than municipalities (except for Limburg, Fryslan, Scheldestromen and Zuiderzeeland). Figure 3.1 shows the territories of the regional water authorities in the Netherlands. These water authorities have adopted drought innovation policies and implementation programmes in recent years. For example, in 2010, the water authority Limburg implemented a policy programme that aimed to raise the groundwater level in nature areas and stream valleys. The plan, which was called 'Nieuw Limburgs Peil', incorporated several measures to realize more water retention and more sustainable ways of irrigation in the agricultural sector (Waterschap Peel en Maasvallei, 2010). Another more recent example of an integrated policy development with regard to the drought problem, is the initiation of the Advisory Committee Drought in the province of Noord-Brabant. This joint decision was made by the water authorities 'Aa en Maas', 'Brabantse Delta' and 'De Dommel', and comprises of a new collaborative approach to deal with drought related issues (Waterschap Aa en Maas, 2021). All in all, water authorities are including drought issues into their policies more and more. Partly by including drought related implications into existing policies and procedures, but also by setting up separate policy programmes.

Figure 3.1: Waterboard authorities and their territories in The Netherlands (Unie van Waterschappen, 2018).



The development of knowledge and the testing of new techniques or partnerships lies at the core of the drought adaptation challenge. In the Dutch context, the Lumbricus programme was set up to search for innovative measures and strategies to deal with drought issues and to realise a more robust soil- and water system. Lumbricus researched the effectiveness and coherence of smart soil- and water management measures along four central themes. These measures were researched and tested in pilots or living labs (Lumbricus, 2021). Figure 3.2 shows the focus area of the Lumbricus programme: the high sandy soils in the dry rural areas in the Netherlands.

The Lumbricus programme ended in March 2021. According to the initiator, the programme resulted in clear answers, but also raised new questions. Therefore, the KLIMAP programme builds on the Lumbricus programme by researching not only into the effectiveness of certain climate adaptive measures, but also looking at economic sustainability and societal use of the measures (STOWA, 2020). A part of the KLIMAP objective is also aimed at developing instruments and models that regions can use themselves to reach their adaptation goals, and developing strategies for the broader societal uptake of successful adaptation innovations (Waterforum, 2020).

Figure 3.2: High sandy soils in the Netherlands, as focus area of the Lumbricus Programme (WUR, n.d.).



4. Methodology

This fourth chapter elaborates on the methodology that was used to conduct this research. First, the choice for the main research strategy will be discussed. Then, the ontological, epistemological, and methodological foundations of this research will be elaborated on. This paragraph also discusses the different uses of QCA and the nature of potential single set and subset relations. The third and the fourth paragraph respectively elaborate on the conditions for case selection and condition selection that were used in this research. Paragraph five discusses the data collection process, after which paragraph six will briefly dwell on the data analysis procedure. Finally, paragraphs seven and eight focus on the reliability, validity and the ethical considerations that played a role in the process of conducting this research.

4.1 Main research strategy

The strategy that was followed does not fit entirely in one of the four main research strategies as described by Van Thiel (2014, p.58) and can best be regarded as a hybrid form. The basis of the research followed a case study-based strategy. Multiple pilots were researched in their contexts. However, where case study research often focuses on a limited number of cases, this research is aimed at more systematically researching a medium to high number of case studies (Van Thiel, 2014, p.58). In the research process, a combination of primary and secondary data was examined. First, desk-research of relevant policy and evaluation documents and websites of the various pilots laid the foundation of the database. By making use of a survey, this database was supplemented with additional cases and information about success(conditions) in the cases. The research was to some extent also evaluative because it focused on pilots that were finished or in the last implementation phase.

The hybrid research strategy was chosen out of several alternatives. Another possibility was to study a limited number of cases qualitatively and more intensively by interviewing multiple involved actors per case study. Where this would probably have led to a more in-depth insight in these few examples, the influence of the local contexts would also be higher. Moreover, the variety of different types of pilots in the drought adaptation field is large. When only five cases were selected, it would be impossible to make statements about drought adaptation pilots as a whole. With regard to the societal challenge of drought adaptation, as well as the scientific debate with regard to the upscaling of pilots, the chosen strategy had a larger potential to eventually be able to make recommendations to address both debates in a more founded way.

4.2 Ontological, epistemological, and methodological foundations of QCA

The central element of the research methodology was qualitative comparative analysis (QCA). This research method was first developed by Ragin (1987). Back then, the main ambition of QCA was to *"integrate the best features of the case-oriented approach with the best features of the variable-oriented approach"* (Ragin, 1987, p.84). Therefore, QCA is seen as a combination between qualitative and quantitative research (Berg-Schlosser et al., 2009, p.5). This two-fold character of QCA leads to the fact that QCA research can include elements from both positivist and constructivist research approaches. When looking at QCA from a historical point of view, the comparative procedures that were carried out in the 18th and 19th century by Hume and John Stuart Mill provide the logical foundations of this way of researching (Berg-

Schlosser et al., 2009). The 'method of difference' and the 'method of agreement' both compared various circumstances that could result in a certain phenomenon, and searched for similarities and differences that could explain this phenomenon. These 'canons' use rather positivist assumptions with regard to causes and effects, without actually implying hard causal relationships (Berg-Schlosser et al., 2009). An important distinction that they seek to make is between factors that could have an influence on a certain outcome and factors that have no influence, which they eliminate. This connects to the 'falsification' norm of Karl Popper (Popper, 1963).

The main principles of QCA are 1) the combination of detailed case analysis and 2) the assumption of complex causality (Ragin, 1987). With regard to the meaning and the occurrence of causality, several differences between QCA research and statistical research can be identified. In essence, QCA follows a conception of causality that leaves room for complexity. This is referred to as 'multiple conjunctural causation' (Berg-Schlosser et al., 2012). This means that different causal pathways, using different combinations of conditions, can lead to the same outcome (De Meur, Rihoux & Yamasaki, 2009). This goes against the principle of additivity in statistical analysis. In QCA research, various combinations of conditions can be equally important to reach a certain desired outcome. This phenomenon is called 'equifinality', and is an important part of the epistemological foundation of QCA (Schneider & Wagemann, 2012). QCA additionally sees causality as being context specific and thereby rejects the assumption of permanent causality. Finally, QCA recognizes the possibility of causal asymmetry, which means that the presence of the outcome variable can require different explanations (Berg-Schlosser et al., 2009).

QCA uses a holistic approach, which researches case studies as a whole and also takes their contexts into account. This qualitative dimension can enable the researcher to provide more clarity on the gathered research data (Berg-Schlosser et al., 2009). Although QCA can be used in an inductive way of research, it often makes use of scientific theories to determine the relevant variables that need to be considered (Berg-Schlosser et al., 2009). This deductive way of research is used in this research. Scientific literature is used in order to formulate hypotheses regarding the conditions for success of the pilots.

There are two general types of QCA methodology: fuzzy set and crisp set. The main difference between the two types can be found in the number of values that the variables can take on. Crisp set QCA uses the principle of dichotomization, where a certain condition is either present (1) or not present (0). Fuzzy set QCA differs from crisp set QCA mainly because it allows partial membership. For example, cases can have a membership score of for example 0.33 or 0.40 (Schneider & Wagemann, 2012). This research will use crisp set QCA as a basis. Dichotomization is used in the research process. For example, whether the broader uptake of the pilot is mentioned as a main target in the pilot's design can be answered by either 'yes' (1) or 'no' (0). Sometimes this dichotomization is more difficult and requires determining a certain threshold value, which is the border between present (1) or not present (0). This also relates to the main criticism of crisp-set QCA, which is about the possible loss of empirical information. However, Berg-Schlosser et al. (2009) argue that, while loss of information can happen, the dichotomization of data can best be seen as a form of simplification of data that can be regarded legitimate, when it is done on the basis of theoretical considerations,

familiarity with cases or "*dialogue between ideas and evidence*" (Berg-Schlosser et al., 2009). These methodological considerations will eventually be discussed in detail in the step-by-step description of the survey generation process.

In QCA research, the relationships between the independent variables (conditions) and the dependent variable (outcome) can best be analysed by means of set-theoretic models. These models help to identify supersets and subsets that are sufficient or necessary for a certain outcome (Basurto & Speer, 2012; Berg-Schlosser et al., 2009). In order to analyse the complex causal relationships between the different variables, three concepts can be distinguished: 1) sufficiency, 2) necessity and 3) INUS. Figure 4.1 shows a conceptual representation of these three concepts, presented in a Venn diagram.





Note. Adapted from "An introduction to applied data analysis with qualitative comparative analysis", by Legewie, N., 2013, *Forum Qualitative Sozialforschung, 14*(3), 1-45.

Condition X is <u>necessary</u> for Y to happen, when Y cannot occur without the presence of X. But the presence of X alone is in this case not enough to produce Y. This means that the condition is present in all cases where Y occurs, but cannot produce Y on its own. This is shown in figure 4.1 by the large circle X around the smaller circle Y. In other words: there are cases in X that are not in Y, but all cases in Y are also in X (Legewie, 2013; Schneider & Wagemann, 2012). Condition X is <u>sufficient</u> for the outcome Y, when Y always occurs when X is present. However, other conditions next to X can also produce Y. As can be seen in figure 4.1, the circle X is smaller than the circle Y, so there are parts of Y that can be explained by other causes (Legewie, 2013; Schneider & Wagemann, 2012). Finally, INUS conditions are "*insufficient but necessary parts of a condition which is itself unnecessary but sufficient for the result*" (Schneider & Wagemann, 2012, p.90). This means that they are always combined with other conditions, which together form a sufficient combination for the outcome. In figure 4.1, this is shown by the small circle X that is partly out of the larger circle Y. However, when combined with other conditions, the circle can transfer into the larger circle Y, in order to produce a sufficient subset.

4.3 Case selection

The selection of cases was done by means of non-probability sampling, specifically a combination between purposive sampling and snowball sampling (Van Thiel, 2014, p. 46). In essence, the cases were selected on theoretical grounds and common characteristics. The

pilots that were initially selected for this research all have in common that they are located in Dutch dry rural areas and focus on drought adaptation. As discussed in paragraph 2.2, the variety of the nature of pilots can be quite large and pilots can serve multiple goals. The criteria by which the cases were selected are presented in figure 4.2. Additional to the criteria in this figure, it is important to mention that the subject of innovation can vary across the pilots (i.e. technical, organisational, societal, etc.). The research does not make a distinction from this perspective.

Figure 4.2: Case selection criteria

- 1. Cases need to be located in Dutch dry rural areas (sandy soils)
- 2. Cases need to be aimed at adaptation towards droughts or freshwater shortages*
- 3. Cases need to be completed or in the evaluation phase
- 4. Cases need to be designated as projects with a 'piloting' character by the initiator or other involved parties
- Cases are often executed by several actors aimed at testing an innovation in a specific area, and monitored in order to disseminate the results when judged as successful**

* This should be one of the goals, but does not have to be the only or main goal. Climate adaptation pilots mostly serve multiple purposes.

** This fifth criterium is used when the organisations do not use the wording 'pilot' themselves, but the project does have these pilot characteristics (based on the literature review in paragraph 2.2.)

Secondly, sampling of the cases in this research was not fixed from the start, but can best be regarded in terms of an iterative process. This reflects the QCA handbook by Berg-Schlosser et al. (2009), who argue that new cases can be added (or old cases can be dropped) in the process of research when, for example, new hypotheses arise. This reflects the snowball sampling type. Here, via the units of study, new cases can be added when they are found suitable and of added value for the research. In this research, new cases were added as a result of new information that was gathered in conversations with involved actors from other cases. On multiple occasions, these respondents pointed towards other pilots that they were involved in, that were subsequently added to the research. No new cases were added as a result of new hypotheses that arose.

Berg-Schlosser et al. (2009) distinguish several 'good practices' with regard to case selection. First, cases should share background characteristics. In this research, all cases were droughtadaptation pilots that are implemented in Dutch dry rural areas, on high sandy soils. Moreover, when selecting the cases, a clear definition of the outcome is necessary. In this case the desired outcome is to explain which conditions lead to internal or external successful pilots. Both terms of success were defined before the cases were selected. Finally, it is important that the researcher gains sufficient familiarity with (the types) of cases (Berg-Schlosser et al., 2009). This was pursued by executing an extensive a priori desk-research, several meetings with experts in the field and the inclusion of qualitative sections in the questionnaire.

4.4 Condition selection

The selection of conditions should be guided by theoretical criteria (Berg-Schlosser et al., 2009). As mentioned in chapter 2, this research used scientific literature about conditions for successful pilots as baseline. The conditions that were mentioned in scientific literature, were empirically tested in the field of drought adaptation pilots. Secondly, the number of conditions should be kept relatively low (Berg-Schlosser et al., 2012). Some of the conditions that were mentioned in literature have a degree of common ground. If this was the case, the criteria were merged into one overarching condition for the quantitative analysis. The open-ended questions were additionally used to acquire more in-depth knowledge about the actual relation between the condition with the dependent variable.

4.5 Data collection

First, an analysis of secondary data from policy documents, evaluation documents and internet sources was performed. A database was created to store the relevant information. This database provided an overview of the cases that were found and sometimes already links to information about whether or not specific conditions were present in the cases, and why. Secondly, a survey was used to gain primary data from involved actors in the pilots. This survey consisted of fixed-response questions, for example statements with five-point Likert scales ranging from 'strongly disagree' to 'strongly agree'. Complementary to these fixed-response questions that are best used for statistical analysis, the survey purposefully also contained open-ended questions or sections where the participants could explain their answers, in order to require as much in-depth information as possible for the eventual QCA analysis to properly function (Basurto & Speer, 2012; Berg-Schlosser et al., 2009). The survey also contained binary questions, which contain the lowest level of detail. However, as this research was aimed at identifying the presence or non-presence of certain outcomes, binary questions could be very useful. To ensure validity, these binary questions were complemented by sections where respondents could further explain their answers.

The survey was divided into three parts. The first part focused on the characteristics of the pilot. The second part focused on researching the outcome variables (i.e. the extent to which the pilots can be regarded as internally and externally successful). First, the internal success was measured by several statements. Afterwards, the external success was measured by several statements. Here, the distinction between scaling up and scaling out was made in determining the success score. In the third part, the conditions for success were measured. This part was categorized along the dimensions of the pilot paradox (as shown in figure 2.3). This means that the first conditions (which are part of the pilot paradox) were to some extent mirrored. However, they were treated as separate conditions to treat both sides of the paradox separately, and questioned by a five-point Likert-scale ranging from strongly disagree to strongly agree. Each condition was independently questioned by several items (Van Thiel, 2014, p.79). An alternative approach would be to question the dimensions directly by designing questions based on semantic differentials, which uses two extremes as opposite poles on a scale (Friborg, Matinussen & Rosenvinge, 2006). In this case, the two conditions (internal and external) that belong to one dimension could be placed on the opposite poles of

the scale. However, for this to properly function, two explicit opposites need to be placed at the end of each scale. As this research wanted to test the presence of the pilot paradox on a more neutral basis and really wanted to use the elements that resulted from the literature review as basis for the empirical process, there was chosen to focus the questioning on both sides of the paradox separately. In this way, the various elements could be treated individually to get more information about the presence of the specific elements of the dimensions. Moreover, questioning both sides of the paradox separately could potentially result in gaining more contextual information on the pilots. After the data collection, the medium Likert scores were calculated and served as a basis for assigning the values to the cases. This process of data preparation is further elaborated on in chapter six of this research report.

The survey was organised in an online environment (Google Forms) and distributed by email among main actors in the different case studies. The triangulation of different data sources (secondary and primary) was used to verify or strengthen the data of the survey when necessary. Possible bias of the information provided by respondents could in this way be identified and taken into account in the data preparation for the QCA analyses.

4.6 Data analysis

As mentioned earlier, QCA makes use of truth tables, which allow researchers to visualize and analyse central features of causal complexity (Schneider & Wagemann, 2012). The first step of data analysis therefore consisted of a manual examination of these tables. In this way, major patterns can possibly be noted already. As a next step, QCA software will be used to perform a number of different systematic calculations on the dataset. The QCA program presented necessary and sufficient (combinations) of conditions for success, out of which the main findings resulted. After this, the gathered qualitative data was used to further manually research whether or not the pilot paradox could be identified. The data analysis itself will be elaborated on further in chapter seven.

4.7 Validity and reliability of the research

4.7.1 Reliability

The central point in this research were the scores that were assigned to the independent variables and the dependent variable. In order for the research to be reliable, every step taken should be clearly elaborated on. This started with determining the factor of success by asking the question: when can the drought pilot be called successful? The next step was to determine whether or not the conditions were present in the pilots. In order to ensure reliability of the research a detailed description of the steps taken, and the motivation why certain steps are taken as they are, is given. By giving these elaborate and sound explanations of the measurement of all variables, repeatability of the research is pursued. When the research would be repeated in exactly the same way and under the same circumstances, this should lead to similar results (Van Thiel, 2014, p. 48). With regard to the items and questions in the survey, a reliability test (Cronbachs Alpha) was executed to ensure sufficient internal consistency. A variety of previously executed QCA studies use this test as a means to strengthen the internal reliability and consistency (Navarro, Llinares & Garzon, 2016; Miranda, Tavares & Queiro, 2018).

4.7.2 Validity

This research combined several research approaches, which all have validity-related implications that need to be dealt with. First, a distinction between internal and external validity needs to be made. Internal validity refers to the adequacy of the operationalization of theoretical constructs and the quality of the causal relationship between the dependent and independent variable (Van Thiel, 2014, p.49). The internal validity of QCA is debated because the researcher can influence the outcome of the research through the values that he or she assigns to the conditions (Basurto & Speer, 2012). Scientific literature poses several tools to address internal validity concerns. In the first place, acquiring as much qualitative data about the cases can help to minimise this measurement error (Schneider and Wagemann, 2012). The triangulation of data sources (qualitative and quantitative) strengthens the internal validity. With regard to the conditions, complementary strategies are available to limit measurement error. This research used a robustness test, by systematically trying different consistency values or thresholds (Thomann & Maggetti, 2020). Secondly, thorough viewing of the truth table and the results can also limit internal validity concerns. If, for example, the minimal formula is too complex or the truth table includes contradictions, the choice of conditions can best be reconsidered (De Meur, Rihoux & Yamasaki, 2009). Finally, there is no unanimity in scientific research with regard to the question how causal relationships between the independent and dependent variables are considered to really exist. Some argue that a valid connection between the independent and dependent variables is "plausible and free from logical contradictions" (Thomann & Maggetti, 2020). To strengthen this, a broad range of theoretical and empirical knowledge can help to distinguish implausible results. This improves the internal validity of the research. This was also pursued in this research.

The external validity, or generalizability, refers to the extent to which the outcomes of the research can be generalized (Van Thiel, 2014, p.49). In case study research, a problem with regard to external validity is the small number of units (Van Thiel, 2014, p. 92). The fact that this research focused on an intermediate to large number of cases strengthens the external validity. However, a large N alone does not ensure generalizability. Moreover, the fact that QCA is case-sensitive creates challenges to the external validity, especially when the cases do not form a representative sample, as in this research (Thomann & Maggetti, 2020). An important element that needed to be considered in the analysis and conclusion is to be careful with making general statements or draw general conclusions. The scope of the research should be clearly present when the outcomes are presented, in order to delimit the context in which the possible relations between variables apply. This empirical scope needs to be in line with the case selection rationale (Thomann & Maggetti, 2020). The discussion section of this research was also focused on putting the results into perspective, in order to make sure that the conclusions are not interpreted as all-encompassing statements, but rather need to be seen in their contexts.

4.8 Research ethics

Because research in the social science field often concerns collecting data from respondents, ethical considerations about the handling of data by the researcher should always be considered (Oliver, 2010). In order to ensure ethical quality, the five standards of ethical research as stated by the Economic and Social Research Council (ESRC) are applied and discussed below. Some of these standards mostly apply to qualitative research. However, because the foundation of QCA lies in between qualitative and quantitative research, these standards are also discussed.

First, research subjects need to be informed about the purpose, methods and uses of the research (Silverman, 2013, p.162; Wester, 2011). In this research process, the e-mails that were sent to (potential) participants to inventory pilots for the database and to distribute the survey, both elaborated on the purpose of the research and the way in which the data would be analysed. Secondly, the confidentiality of information and the anonymity of respondents must be respected (Silverman, 2013, p.62). Because this research mostly focused on policy issues and not on personal or individual problems, the confidentiality of information was not expected to be problematic. However, the participants of the survey had the option to skip a question, when they judged is as confidential. The anonymity of respondents was pursued by not including personal questions in the survey and not including names in the secondary data analysis. However, the fact that pilots often had one or two contact persons and the community of the pilots was relatively small could potentially lower the actual anonymity. Third, participants should be engaged on a voluntary basis (Silverman, 2013, p.62). In this research, participants were approached by e-mail and were completely free to decide whether or not to participate. Fourth, harm to participants needs to be avoided (Silverman, 2013, p.62). This research focused on the systematic analysis of policy related subjects, that were not considered harmful to the participants because of the standardisation of the gathered data into the QCA analysis and the pursued anonymity of the respondents. Finally, the independence of the researcher must be clear and partiality must be explicit (Silverman, 2013, p.63). This research was conducted as a master thesis research project, in which the researcher had no personal interest in the outcomes. The research was conducted within the framework of the KLIMAP project and combined with an internship at Deltares. However, both Deltares and the Radboud University are independent research organizations and were therefore not favoured by any particular outcome.

5. Operationalization

This chapter uses the conceptual frameworks, that were described in section 2.6, as a starting point and describes the next steps in the process of operationalization. During the operationalization process, choices were made to ensure sufficient researchability of the variables. These choices are discussed in this chapter.

5.1 Outcome variable: internal success

As described in section 2.6, the extent to which a pilot can be regarded as internally successful is embedded in the question: has the pilot successfully realized it's previously stated main ambition?" (Van Buuren et al., 2016). Internal success can also relate to the creation of new knowledge about a certain problem, or the development of new networks or collaboration mechanisms between parties that did not cooperate before the pilot (Van Buuren et al., 2016; Van Buuren et al., 2018). This can in practice result in the situation where pilots are never regarded as unsuccessful because learning always takes place, even when the innovation itself fails. To avoid the situation that all pilots are seen as successful by the respondents, measuring the internal success will be operationalized by first asking the respondents about the main goal of the pilot (question 13). The subsequent question (14) will focus on the extent to which the pilot has reached the three options of internal success (reaching goals, knowledge development and new collaboration mechanisms). If the respondent answers the first question by ticking 'knowledge development' as the main goal of the pilot, the creation of new knowledge will be seen as sufficient to regard the pilot as internally successful. When the respondent ticks another option (technical innovation, showcase innovation, etc.) at the first question, only the creation of new knowledge will not be seen as sufficient for an internal successful pilot. In this way, the success of the pilot can be determined as reliably as possible, without the influence of possible bias of the respondents.

5.2 Outcome variable: external success

In the QCA analysis, the external success of pilot projects is categorized in two categories:

- Scaling up within the organization(al area): the pilot leads to some kind of scaling up effect within the carrying organisations or organizational area of the parties that were involved in the initial pilot.
- Scaling out to other organizations: the pilot leads to scaling out effects to other contexts or organizations beyond the initially involved actors, or the pilot's results lead to scaling out effects on a higher scale (e.g. from local to regional, or from regional to national).

As described in section 2.6, this categorization is guiding in the QCA analysis. For the QCA analysis to function, each case needs to be assigned to either 'no scaling up' (0), 'scaling up within the organization(al area)' (0.5) or 'scaling out to other organizations' (1). The combination of questions 15 and 16 in the survey (see appendix A) is aimed at gathering the information to be able to assign the cases to a specific class. The first question asks the respondents whether the pilot's results are being diffused. If the respondent answers this
question by ticking 'no', this leads to a classification 'no scaling up (0)'. If the respondent chooses for one of the other options, the second question is focused on gathering information about the organizational scale of the scaling up process. Respondents can answer from which actor (e.g. regional waterboard authority, local municipality) to which other actor the diffusion took place. Based on the answer in this question, the cases were attributed either a value of '0.5' for scaling up, or a value '1' for scaling out. Question 17 focuses more on the object of scaling up, in order to deepen the knowledge about the (policy) context of the pilot. In this way, the character of scaling up (e.g. replication of innovations, circulation of elements of the innovation or institutionalization of procedures) can be identified and considered in the broader analysis of the pilot contexts.

5.3 Condition variables: internal success

As described in section 2.6, this research distinguishes five conditions for the internal success of pilots. The range of potential conditions for success was broader, but in order to ensure researchability and research validity, the number of conditions in QCA research needs to be low. A large number of conditions leads to the individualization of each case. The fewer the number of conditions that potentially explain the reaching of a certain outcome, the closer the research comes to the 'core' elements of causal mechanisms (Berg-Schlosser et al., 2012). The conditions that are distinguished in this research all have an overarching character that connects the different operationalizable elements to one another. This is shown in table 5.1 below.

Dimension	Operationalizable elements (based on the literature review)
Position of the pilot	 Freedom to explore novel ideas (independence) Flexible applications of law and regulations
(pos_int)	Allowed to fail
Project design and scope	Clear boundaries in space
(pd_int)	 Clear boundaries in time Smaller scale than regime projects
	• Small scale \rightarrow reduced risks
Resource distribution (rd_int)	 Additional time for experimenting, monitoring and analysis Attention and manpower and expertise Additional budget
Participants (part_int)	 Boundary spanners Strong leading actors People willing to take risks/experiment People open to gaining new knowledge
Process design (proc_int)	 Collaborative process design Learning environment Monitoring Open dialogue

Table 5.1: Overview of conditions for internal success (Author, 2021)

The operationalizable elements that were distinguished from scientific literature, are implemented in the survey outline that is presented in appendix A.

5.4 Condition variables: external success

This research distinguishes nine conditions that potentially influence the external success of drought adaptation pilots. These conditions are presented in table 5.2 below, together with the operationalizable elements of each condition according to scientific literature. As was the case with the conditions for internal success, several conditions in this case were merged into overarching conditions to strengthen the researchability and the internal validity of the research (Berg-Schlosser et al., 2012).

Dimension	Operationalizable elements (based on the literature review)
Position of the pilot (pos_ext)	 Connection niche and regime Reporting to organisations during pilot; monitoring progress by regime The legal experimenting room in the pilot was representative for the legal room in everyday practice (legal representativeness) or a specific arrangement that provides extra room?
Project design (pd_ext)	 Representative for problem complexity Sufficient system understanding: generalizability Outcomes are trusted, valued and considered representative
Resource distribution (rd_ext)	 Regular budgets and rules Aimed at efficiency Representative distribution of resources and tasks Practicing with mobilising of regular means and funds
Participants (part_ext)	 Diverse range of actors involved Future users involved Engage potential criticasters and colleagues of the carrying organisations in the same learning process as pilot
Process design (proc_ext)	 Future application areas Strategies for uptake as part of the design Linking to actual societal debate or policy question

Table 5.2: Overview of conditions for external success (Author, 2021)

The operationalizable elements that were distinguished from scientific literature, are implemented in the survey outline that is presented in appendix A.

6. Data preparation for analysis

This chapter focuses on the first step of QCA analysis, as described by Rihoux and De Meur (2008): building a dichotomous data table. The preparation of the data started with the structuring and the organization of the data in an Excel-file. All cases with multiple responses were put together in the data-file. Also, the guestions were numbered and assigned a specific code that included the name of the variable that they referred to. Once this was done, the overall excel file was uploaded in SPSS to be able to conduct the reliability analyses. In order to calculate the reliability scores, the five answer options (fully agree, agree, neutral, disagree, fully disagree) were given a numeric score ranging from one to five. The external sides of the dimensions were mirrored when the focus was on the internal success, and vice versa. Afterwards, a Crohnbachs Alpha test was conducted. The dichotomization of the data was done by generating the medium score of the variables. Therefore, the primary data was decisive at first. Some cases initially scored a neutral value of 3,00. When this occurred often within one case, additional response was sought. This sometimes resulted in additional data, by which the case could be assigned either a '0' or a '1'. When this additional response could not be gathered, secondary data was used to assign a value. This reflects the methodological literature by Rihoux & De Meur (2008), who describe that dichotomization can be done by means of empirical or theoretical/case knowledge or a combination of both sources. This chapter will describe the dichotomization process for all conditions and outcomes.

6.1 Single and multiple response test

The survey resulted in 28 cases that could be included in the analysis. For most cases, multiple potential respondents were approached. However, for many cases, one survey was eventually filled in. This can be explained by the fact that various pilots were executed by the same employees of the waterboard authorities, agriculture organizations or other initiating actors, these are project leaders in the innovation and climate adaptation field. During conversations with a few of these project leaders, they indicated that they got the surveys multiple times, forwarded by other actors of the pilots because they thought that this project leader had the best view of the entire pilot process. However, it is still important to take into account possible implications of the limited response per case on the results of this research. In several cases, multiple surveys were filled in by different respondents. In order to assess to what extent the limited response per case influences the results of this research, these surveys were coded individually, in order to examine whether the survey results differ between the respondents. This comparison is presented in a table in appendix B.

Out of the 25 options where the given answers could individually lead to a differing score to be assigned, 1 option indicated a differing score. This accounts for four percent. The eventual score of this case could still be attributed by means of the mean score of the different answering options. In these cases with multiple answers, the neutral score also occurs rather often. This is, because after a first round of analysis, there was actively sought for more responses in cases where the neutral score occurred often. In these cases, the additional response helped to be able to score the presence of the different variables. As can be seen later in this chapter, not in every case additional response could be gathered. In these cases,

additional secondary data or additional descriptive data from the survey was decisive to determine the scores for the variables. The vast majority of the answers resulted in similar codes. While more data is always better, one could conclude that it is not essential to gather information from multiple respondents per case.

With regard to the outcome variable, the cases with multiple respondents also did not show a significant difference between individual responses. The views of the individual respondents sometimes differed between 'all goals were reached' and 'most of the goals were reached', but there were no cases in which one respondent saw the pilot as very successful and another respondent saw it as unsuccessful. With regard to the external success, most of the interpretations of successful uptake were aligned among respondents of the same case. However, in one case ('Bufferboeren Loosbroek'), one participant indicated that scaling up had not taken place and another respondent indicated some kind of scaling up because of a renewed and increased emphasis on the relationship between water and soil. Section 6.3 is about the coding considerations and will describe more elaborately how this contradictory response was dealt with.

6.2 Data preparation: internal success

The internal success was questioned by survey questions 13 and 14. First, question 13 directly asked about the extent to which previously stated goals were reached. Question 14 additionally elaborated on which type of goals were reached: new cooperation schemes, substantial goals or the gathering of new knowledge. The answers to both questions were compared in the process of categorizing the data, this was done as previously described in section 5.1. Based on the gathered data, twenty-three cases could clearly be defined as internally successful and were attributed the value '1'. For two cases, the respondents described that the previously stated goals were not met, these were attributed the value '0'. For the remaining three cases, the respondents described that the division between goals that were reached and goals that were not reached was equal. The pilot 'Blauwe Dienst -Waterconservering' resulted in new cooperation schemes and new knowledge. However, as described in the answers to questions eight and nine, this pilot was aimed at a contractual innovation. The new cooperation scheme and the gathering of new knowledge can be seen as side-goals, but the main innovation was not reached according to the survey answers. Therefore, this case was attributed a value '0'. Finally, the Pilot Stippelberg and the Pilot Beekloop Bergeijk only resulted in the gathering of new knowledge. The answers to questions eight and nine point out that the gathering of new knowledge about the problem was not one of the specific goals of the pilot. Therefore, and following the line of reasoning that was discussed in section 5.1, these two cases were attributed the value '0'.

6.3 Data preparation: external success

The external success was questioned by survey questions 15, 16 and 17. Questions 15 was guiding, because it directly asked whether there was some kind of effect as a result of the pilot. Based on this question, nine cases were categorized with the value '0', because respondents indicated that the pilots have not resulted in broader effects. Question 16 was decisive whether the remaining cases would be attributed a value of 0,5 or 1. When the pilot

only had effect within the carrying or facilitative organisations themselves, they were attributed the value '0,5', this refers to the scaling up category in the conceptual model of this research. This score is used only to indicate the different categories of external success for now. Eventually, these scores will be translated into '0' or '1', in the separate QCA analyses. When the pilots actually resulted in broader effect in other contexts outside the area of the carrying or facilitative organisations, and were scaled out, they were attributed the value '1'. Based on the gathered data, twelve cases were attributed the value '0,5' and five cases were attributed the value '1'. For two cases, the gathered data asked for additional research to be able to determine what value should be attributed. In the case of 'Boer Zoekt Water', the respondent answered the optional 'other' category of questions 15 with the following addition:

"Yes, as inspiration for other actors. Possibilities are examined to use the pilot's system for additional data gathering, so that the Waterboard Authority can potentially use it to improve its policy" (Respondent 'Boer Zoekt Water')

"More and more parties find each other and the system, also from other water authorities" (Respondent 'Boer Zoekt Water')

The extent to which these 'possibilities' that are being examined are already actively used by the waterboard authority does not become clear, also not after secondary data review. However, the respondent indicated that the pilot resulted in some kind of effect in the carrying organization(s) and also towards other organizations. Therefore, this case was attributed the value 1.

In the case of 'Bufferboeren', the respondent answered the optional 'other' category with:

"In principle: yes. Soil and the relation between soil and water is becoming more of an item and developments are underway in various rural areas. Whether that is due to this project....?" (Respondent 'Bufferboeren')

This answer only hints on a broader societal development, and does not refer to specific effects of the pilot itself. Additional secondary data review also gave no reason to assume that the pilot had direct effects on new pilots or policy changes in the organizational area or outside this area. Therefore, this case was attributed the value '0'.

6.4 Data preparation: pilot paradox conditions

6.4.1 Position of the pilot

The first dimension of the pilot paradox is the position of the pilot towards the broader regime. Based on the literature review, the position of the pilot as being a 'safe haven' at a distance from everyday practices contributes to the internal success of pilots. The external success, however, benefits more by a relatively strong connection between the pilot and the broader regime. In the survey, questions eighteen and nineteen focused on these respective two sides of this dimension of the pilot paradox. To prepare the data for analysis, the data was scored as described in table 6.1 below. This shows the perspective from the internal success. In order to double-check whether the scoring operation was executed correctly, the data was also scored the other way around. This should result in a data matrix where a value '0' for internal success means a value '1' for external success and vice versa.

Question 18	Value	Question 19	Value			
Fully agree	5	Fully agree	1			
Agree	4	Agree	2			
Neutral	3	Neutral	3			
Disagree	2	Disagree	4			
Fully disagree	1	Fully disagree	5			

Table 6.1: Scoring of answers for the position of the pilot

After the assigning values to the given answers, a reliability test was conducted in order to assess the internal consistency of the items. This reliability test of all the items included resulted in a score of 0,774 (see Appendix C), which is considered acceptable (Lavrakas, 2008). To check the internal consistency, several additional tests were carried out, but did not result in a higher consistency score. After this test, the mean score of each case was calculated. Based on this score, the cases were attributed either the value '0' or the value '1' from the viewpoint of internal success. This resulted in sixteen cases with a value '1' and eleven cases with a value '0'. One case had a mean score of three, which stands for the neutral category. This was: 'Pilot SIMMBA stuw'. In order to assign a value to this case, additional secondary data was used. The 'Pilot SIMMBA stuw' was a relatively small-scale innovation pilot in which a new smart weir was tested in one municipality. The smart weir was developed by an individual who works at the waterboard authority 'De Dommel', who came up with the idea by himself (Waterschap De Dommel, n.d.; Klimaatadaptatie Brabant, 2020). This suggests a relatively high amount of creative freedom. The pilot was also set up by this individual, together with one representative from the municipality in which the pilot was located. The secondary data does not suggest a strong connection between the pilot and the carrying organisations. Furthermore, the gathered descriptive primary data shows that the pilot was not embedded in any policy- or research programme and was also not a concrete succession of earlier developments (policy choices, prior pilots or living labs, prior policy programmes). This does not suggest a strict connection between the pilot and the carrying organisations. Based on this information, this case was attributed the value '1', viewed from the internal success perspective.

6.4.2 Project design

The second dimension of the pilot paradox is the project design. In terms of the process design of the pilot, the limited scale, and the presence of clear boundaries in space and time contributes to the internal success of pilots. On the other hand, the outcomes need to be representative, generalizable, and perceived as valuable and trustworthy. Data was gathered about these both sides. In the survey, questions twenty and twenty-one focused on these respective two sides of this dimension of the pilot paradox. The gathered data was also mirrored as described in the previous section. Afterwards, a reliability test was conducted. The results of this test are shown in appendix C, and it resulted in a reliability score of 0,763. The systematic addition and removal of items did not result in significantly higher scores. Therefore, all items were taken into further analysis. The average Likert-scores were

calculated and resulted in twelve cases in which the condition for internal success was more present and fifteen cases in which the condition for external success was more present. One case, the 'Blauwe Dienst – Waterconservering', resulted in a neutral score. Additional primary data that was gathered by the descriptive questions in the survey, shows that this pilot was conducted on a regional scale. More specifically: on several locations in a waterboard authority region. Moreover, secondary data also suggests that multiple locations were part of the pilot (Waterschap Rijn en Ijssel, n.d.). Also, the pilot ran for a relatively long time. The initial pilot ran for two years, but the 'blue services (blauwe diensten)' were concluded for a period of six years, after which the evaluation phase started. This long duration and the fact that it was carried out on multiple locations in a waterboard authority region suggest that in this case, the condition for external success (representative outcomes) was more present than the condition for internal success (limited scale and reduced risks).

6.4.3 Resource distribution

The third dimension of the pilot paradox is about the resource distribution of the pilot. The internal success of the pilots benefits from having additional resources aimed at reaching innovative outcomes. This can be financial resources, manpower or time for monitoring and analysis. However, to increase the chance to scale up the pilot, it is more beneficial to fit the pilot in a representative resource distribution aimed at efficiency. Data was gathered about these both sides. In the survey, questions twenty-two and twenty-three focused on these respective two sides of this dimension of the pilot paradox. A reliability test was conducted and pointed out that removing items did not result in a significant increase of the reliability score (0,746). Therefore, all items were taken into further analysis. The average Likert-scores were calculated and resulted in fifteen cases in which the condition for internal success was more present. No cases had a neutral score, so further secondary data was not necessary to score the cases I terms of this variable.

6.4.4 Participants

The fourth dimension of the pilot paradox is aimed at the composition of the participants of the pilot. The internal success of the pilot benefits the most of the presence of (a group of) individuals who are enthusiastic boundary spanners of the pilot. Furthermore, participants should be open for new knowledge and wanting to learn new things. The external success, however, benefits more from a wide range of different participants from different backgrounds, who are also involved in regular projects aimed at drought adaptation. This external representativeness is considered important for the ability of the pilot to eventually scale up. Data was gathered about these both sides. In the survey, questions twenty-four and twenty-five focused on these respective two sides of this dimension of the pilot paradox. A reliability test was conducted and pointed out a score of 0,618. This is rather low, but not unacceptable according to the general guidelines (Lavrakas, 2008). Additional analysis showed that removing items did not result in a significant increase of the reliability score. Therefore, all items were taken into further analysis. The average Likert-scores were calculated and resulted in twenty-one cases in which the condition for internal success was more present.

Two cases had a neutral score after this first round of scoring. These were: 'Blauwe Dienst -Waterconservering' and 'Slimme Stuw SAWAX'. In order to score these cases in a founded way, additional descriptive data from the survey was used. The pilot 'Blauwe Dienst -Waterconservering' was a cooperative pilot between local farmers and the water authority 'Rijn en Ijssel'. In this pilot, local farmers performed services for the water authority to help conservate freshwater in the region, for which they were compensated. The two most important actors were therefore the water authority and the local farmers. However, an additional advisory group was involved, containing the agricultural organization 'LTO Noord', other agriculture and nature organizations, land ownership organizations and landscape management organizations. Based on this broad variety of pilot participants, the focus of this pilot was considered to be on the external representativeness. Secondly, the pilot 'Slimme Stuw SAWAX' involved a testing location for a new type of smart weirs in which the following actors were involved: a local farmer, the municipality of Ommen, the water authority 'Vechtstromen', Vitens, RWS, the agricultural organization LTO, knowledge and consultancy parties KnowH20, Avallo and De Bakelse Stroom. Moreover, this pilot served specifically as a demonstration pilot where potential future users were also involved (Van Bakel, van den Eertwegh, van Deijl & Mensink, 2020). Based on this broad variety of pilot participants, the focus of this pilot was also considered to be on the external representativeness.

6.4.5 Process design

The final dimension of the pilot paradox is about the process design. On the one hand, the possibility to have an open dialogue within a tailor-made mutual learning environment is considered to be important for the possibility of the pilot to reach the desired innovation, and therefore to be internally successful. For the external success, however, a pilot should lie its focus on the identification of broader embedding potential. Or, in other words: the pilot should have its results ready for mainstreaming. Data was gathered about these both sides. In the survey, questions twenty-six and twenty-seven focused on these respective two sides of this dimension of the pilot paradox. A reliability test was conducted and pointed out a score of 0,635. This is again rather low, but not unacceptable (Lavrakas, 2008). Additional analysis showed that removing items did not result in a significant increase of the reliability score. Therefore, all items were taken into further analysis. The average Likert-scores were calculated and resulted in twenty cases in which the condition for internal success was more present and eight cases in which the condition for extern success was more present. No single-response cases resulted in a neutral score.

7. Results and analysis

In this chapter, the various steps of analysis will be discussed. In the first paragraph, a picture of the variety of cases is sketched in order to get a proper understanding of the research population. The second paragraph focuses on the first research question and provides a more in-depth review of the outcome variable of internal and external success. This was also partly discussed in the previous chapter, but this paragraph discusses more aspects of the process of scaling up. Paragraph three discusses the QCA analysis of the conditions for internal success. In this paragraph, all steps taken are explained and the results are interpreted. Subsequently, paragraphs four and five have the same function, but for the two QCA analyses for external success. All figures in this chapter that indicate no source, were created by the author.

7.1 Case characteristics

One of the main elements of QCA research is about the acquiring of sufficient case knowledge (Schneider & Wagemann, 2012). This was partly done by the previously executed creation of the database, but now that the actual research population is known and additional descriptive data is gathered by means of the survey, it is relevant to draw a picture of the various cases. Therefore, this section will dive deeper into for example the embeddedness of the pilots in research or policy programs, the nature of the pilots and previous developments that led to the organization of the pilots.

First, it is helpful to sketch a picture about the character of the innovations that are pursued in the different cases. Table 7.1 shows an overview of the types of innovations that were the central aims of the pilot projects. It should be noted that pilots often serve multiple goals. For example, the pilot 'Infiltratie Vinkelsestraat Vinkel' focused on sub-irrigation, but did this on a company broad level where water from a lower piece of land was transported and used for subirrigation on an elevated piece of land (Brabants Dagblad, 2017). However, the overview presented in table 7.1 focuses on the main innovation goals of the pilots.

Type of innovation (main goal)	Cases	Quantity
Reuse effluent for agricultural uses	Boer, Bier, Water; Hergebruik Effluent Haaksbergen; Waterfabriek Wilp;	3
Drip irrigation	Waterwijs Boeren Haarlo; More Crop per Drop Drenthe; More Crop per Drop Brabant	3
Level controlled drainage	Peil gestuurde drainage NLP; Peilbeheer op Maat	2
Subirrigation	Subirrigatie Mariapeel; Infiltratie Vinkelsestraat Vinkel;	2
Smart weirs	Slimme stuw SAWAX; Pilot SIMMBA stuw	2
Integrated water system management on company- level	Bufferboeren Loosbroek; Pilot Twickel; Demonstratiebedrijf Neppelenbroek; Wijs met Water; Proeftuin Oost-Stegeren; Boer Zoekt Water	6
Restructuring stream valleys in area pilots	Beekherstel Groote Molenbeek A73 Oost; DHZ pilot Pepinusbeek; Beekloop Bergeijk; Regge Mozaïek Enter	4
New cooperation mechanisms	Blauwe Dienst: Waterconservering; Boeren met Water (blauwe en groene diensten); Pilot Witteveensleiding; KlimaatKlaor?!	4
<i>Water retention (underground or above ground)</i>	Pilot Stippelberg; Blauwe Poort Laarbeek (waterhouderij);	2

Table 7.1: Main innovation goals of the pilots

Most of the pilots that were researched have been completed (78.6%), six cases (21,4%) are currently in the monitoring and evaluation phase. Two questions also focused on the composition of participants in the pilots. Figure 7.1 shows what kind of actors were the initiator of the various pilots. When analysing this figure, it stands out that in 63% of the cases, a (employee of a) governmental party was the initiator of the pilot. In most of these cases, it concerns a water authority. Several cases also indicated a role of a municipality or a province in the initiating phase. For example, the pilot 'Wijs met Water' was initiated by the municipality of Hof van Twente. The pilot 'Boeren met Water' was initiated by a coalition of the ministry and the province of Overijssel. The category with the second highest percentage is private companies. These private companies are mostly farmers who want to develop new sustainable water-management systems on a company level. This is for example the case in the pilot 'Boer Zoekt Water', where the farmer started a pilot to try out the effectiveness of sub-irrigation as a measure to deal with drought problems on his grounds (Boer Zoekt Water, n.d.). Private parties can also concern larger companies, as was the case in the pilot 'Boer, Bier, Water', which was initiated by Swinckels Familiy Brewers. Only a few pilots were initiated by knowledge institutes (Alterra, KWR), agricultural interest groups (LLTB, LTO), local foundations and semi-governmental actors (Brabant Water).



Figure 7.1: Who was the initiator of the pilots?

Additionally, it is interesting to have a look at the question: where does the pilot come from? As earlier described in the literature review, the context of the pilot can have an influence on the scaling up potential of the pilots (Vreugdenhil, 2010). When looking at the data, it stands out that for seventeen cases, the participants indicated that there was no direct link between a prior development (i.e. policy choice, prior pilot projects or policy programmes) within the organisation and the initiation of the pilot. This is a relatively large number, that indicates that many pilots are standing in itself, aimed at reaching innovation. For eleven cases, the respondents stated that the pilot was indeed a continuation of prior developments in the organization. In several cases, the pilot was a continuation of a previous pilot project. For example, the pilot 'Hergebruik Effluent RWZI' followed up on a previous pilot project that focused on how to reuse of effluent of a sewage treatment plant. In this case, the prior pilot focused more on the design of the project as a whole. One respondent stated that:

New research questions step by step, each pilot is an expansion of the prior pilot (Respondent Hergebruik Effluent RWZI)

So in this case, the pilots were used as testing environments in which step by step new elements were added. Another notable result is that when pilots are follow-ups of prior pilots, the character of the innovation does not necessarily need to be similar. For example, in the case of 'Waterwijs Boeren Haarlo', the prior pilot was aimed at the placement of weirs to retain more water in the area. The pilot 'Waterwijs Boeren Haarlo' itself, however, is aimed at testing the effectiveness of drip irrigation. Both pilots are focused on the same problem, but are different types of innovations. Finally, several pilots were part of broader area-wide projects. In this case, the pilots were used to test or to demonstrate a specific innovation.

Next to these prior developments, data was also gathered about the broader embeddedness of the pilots. Vreugdenhil (2010) discussed that the embeddedness of pilots in policies could potentially influence the degree of freedom of these pilots. Figure 7.2 shows that the largest share of cases was not specifically embedded in a research- or policy programme. This suggests a relatively high share of pilots that did have experimental freedom. Twenty-three percent of the pilots were embedded in a policy programme and fourteen were embedded in a research programme. Seventeen percent of the respondents opted for the 'other category'. In the additional open answer section, several additions were described. The pilots 'More Crop per Drop' and 'Boer Zoekt Water' were for example part of a subsidy scheme by the Deltaplan High Sandy Soils. The pilot 'Witteveensleiding' made use of a subsidy scheme of the Deltaplan Agricultural Water Management. Finally, 'Klimaat Klaor?!' was part of the Deltaplan Spatial Adaptation. Other pilots for which the respondents opted for the 'other' category were part of local action plans or the local bundling of measures.



Figure 7.2: Embeddedness of the pilots

Next to the context of the pilots, information was also gathered about the nature of the pilots itself. As earlier described in the literature review, pilots can have multiple natures and multiple goals ranging from technical to organizational and societal pilots. Figure 7.3 shows the results with regard to the nature of the pilots in this research. It is important to mention that the respondents could tick multiple boxes in this question. What stands out is that ninety percent of the cases had a technical character. Additionally, almost seventy percent of the drought adaptation challenge. Only four percent of the cases were described as 'financial'. This information suggests that the focus of the majority of the pilots was more on the technical effectiveness of certain measures, instead of the financial feasibility. More discussion on this topic will be featured later in this chapter.



Figure 7.3: The nature of the pilots

7.2 Analysis of internal and external success

The first sub-question of this research focuses on which pilots are successful and which ones are not. For a small part, this was already discussed in sections 6.2 and 6.3. With regard to the internal success, five cases are seen as not successful and twenty-three cases are regarded internally successful, which accounts for 82 percent. In light of the external success, ten cases did not result in broader effects. Twelve cases resulted in successful scaling up, having some kind of effect within the carrying organizations. Finally, six cases resulted in actual scaling out and had thereby also effect within other organizational contexts. In order to put these kinds of effects in context, additional questions in the survey were aimed at gathering information about what kind of effect the pilots had and what the object of scaling up or out was.

As described in the literature review in chapter two, pilots can result in different types of broader effects. The main categorization of Van Doren et al. (2018) makes a distinction in horizontal and vertical scaling up. Horizontal scaling up refers to the spatial growth of a pilot, resulting for example in other pilots or in inspiration for projects elsewhere. Vertical scaling up is about the extent to which a pilot results in changes in policy and more structural knowledge development processes (Van Doren et al., 2018). When looking at the gathered data, it stands out that sixty-seven percent of the cases that had broader effects, resulted in inspiration for other parties. This can be regarded as the least comprehensive or least tangible form of scaling up. It needs to be noted that respondents could tick multiple boxes under this question. Mostly, 'inspiration' was selected together with another answer. Still, the second and third category were also forms of horizontal scaling up. Only twenty-eight percent of the cases that resulted in broader effects resulted in vertical scaling up, in the form of policy change.





Additionally, data was gathered about the object of scaling up. This can be the innovation itself, an element of the innovation or the general ideas. Depending on the objective of the pilot, it can also be new cooperation mechanisms, certain procedures or methods. When analysing the gathered data, it stands out that in the majority of cases, the innovation itself or ideas were the object of scaling up (both 67% of the cases that resulted in broader effects). In most of the cases, the entire innovation was scaled up. Only in eleven percent, parts of the innovation were the object of scaling up. Also, cooperation mechanisms (56%) and procedures (50%) were found to be often occurring objects of scaling up, reflects the earlier findings about the collaborative character of many of the pilots. Measurement-, modelling- or design methods were seen as the objects of scaling up in thirty nine percent of the cases.

7.3 Analysis of conditions for internal success

The first QCA analysis that is conducted focuses on the conditions for internal success. For this, the data matrix that was composed in chapter six will be used. This analysis will be conducted from the internal success perspective. Therefore, the internal success variants of the dimensions will be used in the QCA operation (*pos_int, pd_int, rd_int, part_int, proc_int*). A '1' in this case, means a presence of the internal side of the dimension. A '0' means the presence of the external side of the dimension. The data matrix that is used to answer subquestion two, is presented below. A first look at the data matrix already shows that the vast majority of the cases is internally successful. This means that the QCA can result in relatively many possible subsets that lead to internal success. It is however still interesting to continue the QCA operation, in order to examine potential helpful subsets and try to make conclusions about the relative importance of the various dimensions for the internal success.

Case							Internal
nr.	Case name	pos_int	pd_int	rd_int	part_int	proc_int	succes
C1	Boer, bier, water	0	0	1	0	0	1
C2	Klimaat Klaor	1	1	0	1	1	1
	Subinfiltratie Effluent						
C3	Haaksbergen	1	1	1	1	0	1
C4	Bufferboeren Loosbroek	0	0	0	0	1	0
	Waterwijs Boeren						
	Gelderland (bevloeiing						
C5	haarlo)	1	1	0	1	0	1
	Pilot Waterfabriek						
C6	(Terwolde)	1	1	1	1	1	1
C7	Pilot peilgestuurde drainage	1	1	0	1	1	1
	More crop per drop (Noord-						
C8	Brabant)	0	1	0	1	0	1
C9	Pilot Grote Molenbeek	0	0	0	1	0	1
C10	Boer Zoekt Water	0	0	0	1	0	1
	Pilot Stroomgebied						
	pepinusbeek-vulensbeek-						
C11	middelsgraaf	0	0	0	1	1	0
C12	Proeftuin Oost Stegeren	0	0	1	1	0	1
	Blauwe dienst -						
C13	Waterconservering	0	0	0	0	1	0
	More crop per drop						
C14	(Drenthe)	0	0	0	1	1	1
C15	Pilot Stippelberg	0	0	1	0	1	0
C16	Pilot Beekloop Bergeijk	1	0	1	1	1	0
	Blauwe dienst - Boeren met						
C17	water	0	0	0	1	1	1
	Waterhouderij Beek en						
	Donk/Blauwe Poort						
C18	Laarbeek	1	1	1	0	1	1
C19	Sub-irrigatie Mariapeel	1	0	0	1	1	1
C20	Slimme stuw (SAWAX)	1	0	1	0	1	1
C21	Regge Mozaïek Enter	1	0	0	0	1	1
	Infiltratie Vinkelsestraat						
C22	Vinkel	1	1	1	1	1	1
C23	Pilot SIMMBA stuw	1	1	1	1	1	1
	Demonstratiebedrijf						
C24	Neppelenbroek	1	1	1	1	1	1
C25	Pilot Twickel	1	1	1	1	1	1
C26	Witteveensleiding	1	0	1	1	1	1
C27	Wijs met Water	1	0	1	1	1	1
C28	Peilbeheer op Maat	1	1	1	1	0	1

Table 7.2: Datamatrix for the internal success

7.3.1. Generating the truth table

Now it is time for step two of the QCA operation as described by Rihoux and De Meur (2008): the construction of a "truth table". This table shows all configurations (combinations of conditions) that were found in the data, as well as the relation to the outcomes. In essence, five configurations are possible in crisp set QCA (csQCA) (Rihoux & De Meur, 2008; Berg-Schlosser et al., 2009):

- 1. Configurations with a '1' outcome: that lead to internal success
- 2. Configurations with a '0' outcome, that lead to internal failure
- 3. Configurations with a '-' outcome, which means that the outcome is indeterminate
- 4. Configurations with a 'C' outcome. These are contradictory configurations that lead to a '0'outcome in some cases and to a '1' outcome in other cases.
- 5. Configurations with an 'L' or 'R' outcome, which are logically possible but have not been observed within the empirical cases.

The truth table for the internal success was generated by using Tosmana. This QCA software is especially useful for small to medium csQCA operations (Cronqvist, 2004). The truth table is presented in table 7.3 below.

Case numbers	pos- int	pd- int	rd- int	part- int	proc- int	Internal succes	Quantity
C9, C10	0	0	0	1	0	1	2
C1	0	0	1	0	0	1	1
C12	0	0	1	1	0	1	1
C8	0	1	0	1	0	1	1
C5	1	1	0	1	0	1	1
C3, C28	1	1	1	1	0	1	2
C4, C13	0	0	0	0	1	0	2
C11(0), C14(1),	0	0	0	1	1	C	3
C17(1)							
C15	0	0	1	0	1	0	1
C21	1	0	0	0	1	1	1
C19	1	0	0	1	1	1	1
C20	1	0	1	0	1	1	1
C16(0), C26(1),	1	0	1	1	1	C	3
C27(1)							
C2, C7	1	1	0	1	1	1	2
C18	1	1	1	0	1	1	1
C6, C22, C23, C24, C25	1	1	1	1	1	1	5

Table 7.3: Truth table for the internal success

First, it is interesting to manually analyse the truth table. With regard to the outcome (internal success), twelve possible configurations arise that can result in internal success. As described earlier, this was to be expected because of the relatively high share of internally successful pilots. Also, two configurations with a '0' outcome can be identified, these are C4 (Bufferboeren Loosbroek), C13 (Blauwe Dienst – Waterconservering) and C15 (pilot

Stippelberg). The subset with the highest quantity of cases resulting in a successful outcome is the subset where all conditions are present. This suggests a link between the presence of the conditions and the presence of a successful outcome in first sight. On the other hand, there is a large variety of subsets that also lead to a successful outcome. C1, C9 and C10 show that a case can also lead to internal success when only one condition is present. This suggests that there are potentially other (contextual) factors that influence a pilot's ability to become internally successful. Also, two contradictory configurations arise. These configurations both contain three cases. The contradictory cases will be dealt with in the following paragraph.

7.3.2. Resolving contradictory configurations

As shown in the truth table in table 7.3, two contradictory configurations were detected. The presence of contradictory cases is normal in the course of a csQCA, as this methodology has an iterative character and is about the dialogue between the researcher and the cases. These contradictory cases, however, need to be resolved or at least reduced in order to obtain more coherent data. Methodological literature suggests various strategies to deal with these contradictory configurations, ranging from the addition of conditions to the re-examination of the cases as a whole (Rihoux & De Meur, 2008; Berg-Schlosser et al., 2009).

What stands out in the contradictory configurations of these cases is that both configurations consist of three cases per subset, in which one case results in a '0' outcome and two cases in a '1' outcome. Also, in both cases, the case that results in a '0' outcome is an area pilot that is aimed at the restructuring of a stream valley in a dry rural area. A potential strategy to deal with these contradictory cases is to recode all contradictory configurations as '0' on the outcome (Rihoux & De Meur, 2008). However, this would not make sense in this case, because in both instances, two cases are successful in the specific subset. Therefore, a closer look into the nature of these cases might help. The two cases are both indicated as 'area pilots' by the water authorities, but are somewhat different than other cases. They do meet the criteria for case selection, but can be seen as 'borderline' cases, because of the relatively high scale and the integration of many area-wide goals, instead of testing one specific innovation to deal with regional drought. Rihoux & De Meur (2008) argue that such borderline cases can best be left out of the analysis, when they lead to contradictory configurations. This is also in line with the use of frequency criteria, as described by Berg-Schlosser et al. (2009). Therefore, both cases (C11 and C16) are omitted from further analysis. A new truth table was generated, which can be seen in table 7.4.

Case numbers	pos-	pd-	rd-	part-	proc-	Internal	Quantity
	Int	Int	Int	Int	Int	succes	
C9, C10	0	0	0	0	1	1	2
C1	0	0	0	1	0	1	1
C12	0	0	0	1	1	1	1
C8	0	0	1	0	1	1	1
C5	0	1	1	0	1	1	1
C3, C28	0	1	1	1	1	1	2
C4, C13	1	0	0	0	0	0	2
C14, C17	1	0	0	0	1	1	2
C15	1	0	0	1	0	0	1
C21	1	1	0	0	0	1	1
C19	1	1	0	0	1	1	1
C20	1	1	0	1	0	1	1
C26, C27	1	1	0	1	1	1	2
C2, C7	1	1	1	0	1	1	2
C18	1	1	1	1	0	1	1
C6, C22, C23, C24, C25	1	1	1	1	1	1	5

Table 7.4: Final truth table for the internal success

7.3.3. Sufficiency and necessity of single sets

The new truth table shows that the contradictory configurations are solved. Now the final truth table is generated, the consistency scores can be calculated for the single sets or individual variables. The consistency scores of the single sets can be calculated manually trough making use of the following formula:

"The number of cases with a '1' value on the condition AND a '1' value on the outcome, divided by the total number of cases with a [1] outcome value" (Rihoux & De Meur, 2008).

Following this calculation formula, the necessity scores of the variables are: Position of the pilot (pos_int): 0,696; Project design (pd_int): 0,522; Resource distribution (rd_int): 0,565; Participants: (part_int): 0,826; Process design (proc_int): 0,652. The participants dimension of the pilot paradox has the highest necessity score. In scientific literature, a debate consists about the threshold of necessity. This research commits to the generally most widely shared viewpoint of Schneider (2019) and Ragin (2000) that a single case or a subset must have a necessity consistency score of 0,9 or higher to be regarded as necessary condition. The case should have a sufficiency consistency score of 0,8 or higher to be regarded as sufficient for the outcome (Ragin, 2000; Fiss, 2011).

This means that the presence of a (a coalition of) leading boundary spanners and competent participants that are open to new knowledge seems to be a sufficient condition for internal success of drought adaptation pilot projects, based on the threshold of 0,8. However, it is not necessary, because other conditions can also produce internal success. The remaining four single sets are not considered sufficient or necessary to produce the desired outcome by themselves, based on their consistency scores. However, the consistency score still says something about the relative importance of the single conditions compared to the desired outcome of internal success. The position of the pilot, as being a safe haven at a distance from the carrying organization also seems to influence the internal success.

7.3.4. Boolean minimization and interpretation of results

The minimization of the data is done by using Tosmana. This software minimizes the configurations by using minimization algorithms. It is important to apply the minimization procedure to the '1' configurations and to the '0' configurations separately, because QCA does not expect perfect causal symmetry (Schneider & Wagemann, 2012; Rihoux & De Meur, 2008). First the minimization of the '1' configurations was executed. This resulted in various potential subsets, thus also various possible formulas. These are shown in Appendix D. As earlier described, this was to be expected, because of the high share of internally successful cases. Thus: many subsets potentially lead to internal success. This suggests that there is no all-encompassing formula for success when looking at the conditions that are part of this research. However, Tosmana offers the possibility to reduce the possible formulas further, by using simplifying assumptions. This can be done by taking the logical remainders into account in the analysis. These are configurations that were not identified by the gathered data. The software makes a simplifying assumption by assuming that if these configurations existed, they would also lead to a '0' outcome. The software only selects logical remainders that are necessary to obtain a shorter minimal formula and adds them to the observed cases (Rihoux & De Meur, 2008).

When including these logical remainders in the analysis, fourteen simplifying assumptions are made, and the following minimal formula is obtained. When the variable is presented in capital letters, it indicates a presence of the specific condition. When the variable is presented in regular letters, it indicates an absence of the condition.

POS_INT + PART_INT + proc_int

This formula contains three elements. It shows that internal success is more likely to happen in cases where the pilot is a safe haven at a distance from the carrying organizations, comprising of much creative and organizational freedom, or when the pilot comprises of a coalition of leading boundary spanners. A surprising result is found in the third element of this formula. It shows that in cases where the process design has a more open and collaborative character aimed at mutual learning, instead of being focused on the broader embedding potential, it more likely results in internal failure. This is at odds with the literature about the pilot paradox, which suggests that an open learning environment contributes to internal success.

The fact that only a few cases resulted in internal failure, perhaps makes it more interesting to look at which subsets were present in these cases, in order to say something about the combination of conditions that lead to internal failure. This follows methodological literature by Rihoux & De Meur (2008), which states that every QCA analysis should involve a minimization of both the '1' and the '0' outcomes, which means that for every QCA analysis the success ('1') and the failure ('0') should be analysed and follow a minimization procedure separately. This is because of the causal asymmetry principle of QCA. For this, the following minimal formula is generated:

PROC_INT	*	pos_int	*	pd_int	*	part_int
(C4,C13+C15)						

This formula shows that in cases where the pilot was: 1) designed together in an open learning environment, 2) was no safe haven at a distance 3) was not executed on a limited scale to reduce risks, and 4) not comprised of a coalition of leading boundary spanners resulted in internal failure. The combination of three dimensions of the pilot paradox influences the internal success as expected. When the position of the pilot, the project design and the composition of participants is too much focused on the external success (i.e. broad representativeness), this could become a barrier in the chance of reaching the desired innovation. One dimension, again, forms a contrary image when linking the results of this QCA back to the literature review. Following this formula, the presence of a collaborative process design within an open learning environment seems to hinder the chances to reach internal success, when combined with an absence of the three other conditions. It can, however, also mean that the absence of three of the conditions outweighs the presence of this condition.

When including the logical remainders for reduction, the formula looks as follows:

PROC_INT * pos_int * part_int (C4,C13+C15)

Here, two simplifying assumptions were executed. The project design dimension is excluded from the minimal formula, based on these simplifying assumptions. The cases for which this formula counts are 'Bufferboeren Loosbroek', 'Blauwe Dienst – Waterconservering' and 'Pilot Stippelberg'. All three pilots were different in nature and aim. Where 'Bufferboeren Loosbroek' focused on developing technical innovations in a cooperation scheme consisting of the water authority and the ZLTO, the pilot 'Stippelberg' was aimed at testing a technical innovation as part of a research programme by various knowledge institutes in cooperation with the water authority. The pilot 'Blauwe Dienst – Waterconservering' was not aimed at a technical innovation, but rather focused on a new contractual cooperation mechanism in which local farmers and property owners performed maintenance services for the water authority (Waterschap Rijn en Ijssel, 2011). All pilots were conducted in a cooperation scheme

with a relatively broad variety of stakeholders. 'Bufferboeren Loosbroek' was for example conducted within a coalition of Brabant Water, the water authority, the agricultural interest group ZLTO, the province and several knowledge organizations. Additionally, in the pilot 'Stippelberg', the water authority, province, several knowledge institutes, and the forest manager were part of the pilot's coalition. The result of the QCA analysis suggest that this connectedness of the pilot with the carrying organisations and the broad variety of stakeholders that was present, potentially obstructed the development of the innovation within the pilot. When many stakeholders have a direct say in the pilot, it is plausible that this slows down the overall process and delimits room for experimentation.

7.4 Analysis of conditions for external success: scaling up

The second QCA analysis that is conducted focuses on the conditions for external success. For this, the data matrix that was composed in chapter six will again be used. This analysis will be conducted from the external success perspective. Therefore, the external success variants of the dimensions will be used in the QCA operation (*pos_ext, pd_ext, rd_ext, part_ext, proc_ext*). A '1' in this case, means a presence of the external side of the dimension. A '0' means the presence of the internal side of the dimension. The data matrix that is used to answer sub-question two, is presented below. As one can see, the previously used scoring of value '0.5' for scaling up and '1' for scaling out is recoded into two variants. In other words: the data is dichotomized. Variant 1 sees scaling up and scaling out as external success and variant 2 only sees scaling out as external success. This paragraph will conduct the QCA for scaling up (variant one). The next paragraph will focus on scaling out (variant two). Through this separate approach, more nuanced data can be gathered about to what extent the different subsets influence the different types of uptake of the pilot's results.

Case nr.	Cases	pos_ ext	pd_ ext	rd_ ext	part_ ext	proc_ ext	Extern succes	Extern succes	Extern succes
								V1	V2
C1	Boer, bier, water	1	1	0	1	1	1	1	1
C2	Klimaat Klaor	0	0	1	0	0	0.5	1	0
C3	Subinfiltratie Effluent Haaksbergen	0	0	0	0	1	0.5	1	0
C4	Bufferboeren	1	1	1	1	0	0	0	0
	Loosbroek								
C5	Waterwijs Boeren Gelderland (bevloeiing haarlo)	0	0	1	0	1	0.5	1	0
C6	Pilot Waterfabriek (Terwolde)	0	0	0	0	0	0	0	0
C7	Pilot peilgestuurde drainage	0	0	1	0	0	0.5	1	0
C8	More crop per drop (Noord-Brabant)	1	0	1	0	1	0	0	0
C9	Pilot Grote Molenbeek	1	1	1	0	1	1	1	1
C10	Boer Zoekt Water	1	1	1	0	1	1	1	1
C11	Pilot Stroomgebied pepinusbeek- vulensbeek- middelsgraaf	1	1	1	0	0	0,5	1	0
C12	Proeftuin Oost Stegeren	1	1	0	0	1	1	1	1
C13	Blauwe dienst - Waterconservering	1	1	1	1	0	0.5	1	0
C14	More crop per drop (Drenthe)	1	1	1	0	0	0,5	1	0
C15	Pilot Stippelberg	1	1	0	1	0	0	0	0
C16	Pilot Beekloop Bergeijk	0	1	0	0	0	0	0	0
C17	Blauwe dienst - Boeren met water	1	1	1	0	0	0.5	1	0
C18	Waterhouderij Beek en Donk/Blauwe Poort Laarbeek	0	0	0	1	0	1	1	1
C19	Sub-irrigatie Mariapeel	0	1	1	0	0	0.5	1	0
C20	Slimme stuw (SAWAX)	0	1	0	1	0	0.5	1	0
C21	Regge Mozaiek Enter	0	1	1	1	0	1	1	1
C22	Infiltratie Vinkelsestraat Vinkel	0	0	0	0	0	0	0	0
C23	Pilot SIMMBA stuw	0	0	0	0	0	0.5	1	0
C24	Demonstratiebedrijf Neppelenbroek	0	0	0	0	0	0	0	0
C25	Pilot Twickel	0	0	0	0	0	0	0	0
C26	Witteveensleiding	0	1	0	0	0	0	0	0
C27	Wijs met Water	0	1	0	0	0	0	0	0
C28	Peilbeheer op Maat	0	0	0	0	1	0.5	1	0

Table 7.5: Datamatrix for the external success

7.4.1. Generating the truth table

The second step in the analysis is the construction of a truth table with all the configurations (combinations of conditions) that were found in the data, as well as the relation to the outcomes. The truth table for scaling up was generated by using Tosmana. The truth table is presented in table 7.6 below.

Case numbers	pos- Int	pd- Int	rd- Int	part- Int	proc- Int	Internal succes	Quantity
C6(0), C22(0), C23(1), C24(0), C25(0)	0	0	0	0	0	С	5
C3, C28	0	0	0	0	1	1	2
C18	0	0	0	1	0	1	1
C2, C7	0	0	1	0	0	1	2
C5	0	0	1	0	1	1	1
C16, C26, C27	0	1	0	0	0	0	3
C20	0	1	0	1	0	1	1
C19	0	1	1	0	0	1	1
C21	0	1	1	1	0	1	1
C8	1	0	1	0	1	0	1
C12	1	1	0	0	1	1	1
C15	1	1	0	1	0	0	1
C1	1	1	0	1	1	1	1
C11, C14, C17	1	1	1	0	0	1	3
C9, C10	1	1	1	0	1	1	2
C4(0), C13(1)	1	1	1	1	0	С	2

Table 7.6 Truth table for external success: scaling up

7.4.2. Resolving contradictory configurations

As shown in the truth table in table 7.6, three subsets were detected that contain contradictory configurations. These contradictory cases need to be resolved or at least reduced in order to obtain more coherent data. The first contradiction is found in the subset where all the external conditions are absent. Four cases result in a '0' outcome, which is to be expected according to the literature about the pilot paradox. However, one case results in a '1' outcome, even though all external conditions are absent. This is interesting, because it could potentially relativize the role of the pilot paradox for the scaling up potential of pilots. It is, however, just one case and the categorization of the data can also have influenced the occurring of this contradiction. The contradictory case (C23) is called 'Pilot SIMMBA stuw'. It

involves the testing of a smart weir that can be used for water retention purposes in dry periods. Scientific literature poses several options on how to deal with contradictory cases. One probabilistic strategy that could work in this case, is using frequency criteria. In this particular contradictory subset, four cases result in a '0' outcome and only C23 results in a '1' outcome. It could make sense to translate the most frequently travelled path for all the cases, which would mean that the outcome of C23 needs to be recoded into a '0' outcome (Rihoux & De Meur, 2008). However, because of the case-based character of the research, this probabilistic strategy alone would not be enough to alter the data. Another strategy opposed by Ragin (1987) suggests recoding all contradictory cases as '0' on the outcome value. In this way, the contradictory cases are seen as unclear and accepts fewer potential configurations. However, by accepting this, the consistency in the case-outcome relationship remains (Rihoux & De Meur, 2008). Through this reasoning, C23 was eventually coded '0' on the outcome value.

There is one subset that involves contradictory cases remaining. These are C4 and C13. For both cases, a complementary primary and secondary data review was executed. However, this did not result in answers to why these cases in the same subset lead to different outcome values. Based on the earlier executed literature review, one would expect that the outcome of this subset should be '1', because four of the conditions for external success are present in this subset. However, no empirical or secondary data about C4 ('Bufferboeren Loosbroek') suggests any direct scaling up of the pilot's results. For the purposes of consistency, the same strategy as before was chosen to deal with this contradictory configuration. Both cases were assigned '0' as the outcome value. In this way, the contradictory cases are seen as unclear and fewer potential configurations are accepted (Ragin, 1987). A new truth table was generated, which can be seen in table 7.7.

Column1	pos-	pd-	rd-	part-	proc-	Scaling	Quantity
	ext	ext	ext	ext	ext	up	
C6, C22, C23,	0	0	0	0	0	0	5
C24, C25							
C3, C28	0	0	0	0	1	1	2
C18	0	0	0	1	0	1	1
C2, C7	0	0	1	0	0	1	2
C5	0	0	1	0	1	1	1
C16, C26, C27	0	1	0	0	0	0	3
C20	0	1	0	1	0	1	1
C19	0	1	1	0	0	1	1
C21	0	1	1	1	0	1	1
C8	1	0	1	0	1	0	1
C12	1	1	0	0	1	1	1
C15	1	1	0	1	0	0	1
C1	1	1	0	1	1	1	1
C11, C14, C17	1	1	1	0	0	1	3
C9, C10	1	1	1	0	1	1	2
C4. C13	1	1	1	1	0	0	2

Table 7.7: Final truth table for external success: scaling out

7.4.3. Sufficiency and necessity of single sets

Now the final truth table is generated, the consistency scores can be calculated for the single sets or individual variables. The consistency scores of the single sets can be calculated trough the following formula:

"The number of cases with a '1' value on the condition AND a '1' value on the outcome, divided by the total number of cases with a [1] outcome value" (Rihoux & De Meur, 2008).

Following this calculation formula, the scores of the variables are: Position of the pilot (pos_ext): 0,444; Project design (pd_ext) 0,611; Resource distribution (rd_ext): 0,611; Participants: (part_ext): 0,278; Process design (proc_ext): 0,389. Based on these consistency scores, no single sets are sufficient or necessary for the outcome to happen, because no values lie above the threshold of 0,8 for sufficiency and 0,9 for necessity (Fiss, 2011; Ragin, 2000; Schneider, 2019). However, representativeness and generalizability of the outcomes (pd_ext) and the use of a representative resource distribution (rd_ext) both seem to have the most individual influence on the external success.

7.4.4. Boolean minimization and interpretation of the results

The minimization of the data is again done by the use of Tosmana, a software-program that minimizes the configurations by using minimization algorithms. It is important to apply the minimization procedure to the '1' configurations and to the '0' configurations separately, because QCA does not expect perfect causal symmetry (Schneider & Wagemann, 2012; Rihoux & De Meur, 2008). First the minimization of the '1' configurations was executed. This resulted in various potential subsets, thus also various possible formulas. These are shown in Appendix D. When including the logical remainders, the formulas are simplified more. This results in the following formulas:

pos-ext * RD-EXT +	pos-ext * PART- EXT +	rd-ext * PROC-EXT +	POS-EXT * PD-EXT * part- ext
(C2,C7+C5+C19+C21) Simplifying assumptions: 10	(C18+C20+C21)	(C1+C3,C28+C12)	(C9,C10+C11,C14,C17+C12)
pos-ext * RD-EXT +	pos-ext * PART- EXT +	pd-ext * PROC- EXT +	PD-EXT * RD-EXT * part-ext
(C2,C7+C5+C19+C21) Simplifying assumptions: 9	(C18+C20+C21)	(C1+C3,C28+C12)	(C9,C10+C11,C14,C17+C19)
pos-ext * PART- EXT +	pos-ext * PROC- EXT +	PD-EXT * PROC- EXT +	RD-EXT * part-ext
(C18+C20+C21)	(C3,C28+C5)	(C1+C9,C10+C12)	(C2,C7+C11,C14,C17+C19)

Simplifying assumptions: 9

What stands out when looking at these results, is that there is no pattern for successful scaling up that really are distinctive compared to other alternatives. As was the case with internal success, there is not a specific 'formula for success' that, when followed, will lead to externally successful pilot projects. There are, however, sub-patterns and results that are

rather interesting. First, the absence of condition *pos_ext* often can together with the presence of other conditions for success (PART_EXT, RD_EXT or PROC_EXT) potentially still be externally successful. This means that pilots that are safe havens at a distance from the carrying organizations can still successfully scale up, when they for example ensure a sufficient amount of external representativeness with regard to the participants of the pilot, as was the case for 'Blauwe Poort Laarbeek', 'Slimme Stuw Sawax' and 'Regge Mozaïek Enter'. All three cases show a relatively varied composition of stakeholders, while at the same time being experienced as a 'safe haven' without strict connections to the carrying organizations. A potential explanation for this phenomenon could be that the pilot is subdivided into phases, where the development of the innovation was done in an environment that is regarded as a 'safe haven', but the monitoring was done by also including other stakeholders and potential future stakeholders. Such a strategy is mentioned for the pilot 'Slimme Stuw Sawax', where the first steps of the development of the smart weir was conducted by technicians, partly connected to the Starting Up Business Challenge, after which the effectiveness of the measure was monitored (and later also demonstrated) together with local farmers (KnowH2O, n.d.; Waterschap Vechtstromen, n.d.). Also, pilots that are a safe haven at distance can still be successfully scaled up when combined with a representative resource distribution existing of regular budgets and means, or with an examination of potential future application areas or scaling strategies in the pilot design. It is important to make the remark that this is just a sub-pattern, and that the coverage and the number of cases in these subsets is rather low. However, it does say something about the internal relationship between the different dimensions of the pilot paradox.

When further examining the coverage scores of the various subsets that are mentioned in the minimal formulas, it stands out that the subsets with the highest coverage scores are the following:

POS_EXT * PD_EXT * part_extCoverage: 0,400PD_EXT * RD_EXT * part_extCoverage: 0,400RD_EXT * part_ext * proc_extCoverage: 0,400

The main result that stands out is that the external representativeness of participants of the pilot is absent in three of these subsets. This can mean several things. First, it can mean that there is no relation between the external representativeness of the participants in the pilot and the external success. This does, however, not necessarily have to be the case. It can also mean that the presence of a coalition of leading boundary spanners (the other side of the 'participants' dimension) also influences the scaling up potential of pilots. This reflects an insight that was discussed in the first literature review, namely that boundary spanners can also perform transformational and outgoing leadership functions and can thereby serve as ambassadors towards for example future users of the pilot (Van Meerkerk & Edelenbos, 2014; Van Buuren et al., 2016). In this way, they perform a dual role as connective actors. It can also mean that both sides of the paradox can apply at the same time. Based on the gathered data, the 'participants' dimension of the pilot paradox does not seem to be as clear-cut as described

in scientific literature. This gives room to the possibility that the paradoxical character in this dimension does not necessarily have to be as paradoxical as described, and that a pilot can contain a coalition of leading boundary spanners while at the same time offering enough space for sufficient external representativeness.

Furthermore, the combination of a project design that ensures representativeness with regard to the problem complexity and generalizable outcomes, together with a strong connection of the pilot towards the carrying organizations or a resource distribution that is built out of regular budgets and aimed at efficiency, leads to the most potential for successful scaling up. For this scaling up within the organization, the QCA thus proves that representativeness of the project is key. The fact that this representativeness requires a connection with the carrying organizations in terms of scaling up within the organization is understandable from a conceptual point of view.

The same Boolean minimization procedure was conducted for the various configurations that led to a '0' outcome. This results in the following formula when the logical remainders are not included for reduction:

POS-EXT * PD-EXT * PART-	pos-ext * rd-ext * part-ext * proc-	POS-EXT * pd-ext * RD-EXT *
EXT * proc-ext +	ext +	part-ext * PROC-EXT
(C4,C13+C15)	(C6,C22,C23,C24,C25+C16,C26,C27)	(C8)

POS_EXT * PD_EXT * PART_EXT * proc_ext	0.2500 **
pos_ext * rd_ext * part_ext * proc_ext	0.6667 **
POS_EXT * pd_ext * RD_EXT * part_ext * PROC_EXT	0.0833

When examining this generated minimal formula, it stands out that the subset with the largest coverage score indicates that when the external sides of four of the five dimensions of the pilot paradox are absent, scaling up will most likely be unsuccessful. These dimensions are the position of the pilot, the resource distribution, the composition of participants and the process design. This result is in line with the literature about the pilot paradox, because it shows that pilots that are a safe haven, consist of additional budgets for innovation, are composed by a small group of leading boundary spanners and designed as collaborative open learning environments, are less likely to be successful externally. In these cases, the focus lies to a large extent on reaching innovation, which comes at the expense of the representativeness and external generalizability of the results.

There are, however, two other subsets that can hinder successful scaling up. One subset has a low coverage score of 0.0833 and only counts for one case. This can therefore not be regarded as a pattern in the data. The other subset suggests that a presence of the external sides of the position, project design and participants dimension, together with an absence of the external side of the process design dimension, can lead to external failure. When examining the cases that are part of the subsets, it stands out that these are exactly the same cases that formed the formula for internal failure. In these cases, the desired innovation was not reached, or the previously stated goals and knowledge questions were not answered. This can explain this contradictory result of this QCA procedure. The fact that these cases were externally unsuccessful may be because the innovation was assessed as not successful, or at least not successful enough to be an effective measure to deal with the regional drought problem. When this conclusion is drawn in such a pilot, no broader effect will be pursued. In this way, the conditions for external success can be present, but the fact that internal success (i.e. the previously stated innovation-goal) was not reached results in no further scaling up. In principle, it is a possibility that an internally unsuccessful pilot still leads to successful scaling up in terms of new cooperation mechanisms or new pilot procedures. However, the participants in these cases indicated that there is an increased attention for the relation between soil and water in terms of dry rural areas. The respondent could, however, not attribute this effect directly to the pilot.

The previous minimal formula can be minimized more by including the logical remainders, this did not result in new remarkable results, the minimal formula is presented in appendix D.

7.5 Analysis of conditions for external success: scaling out

In order to search for (combinations of) conditions that influence the scaling out potential of pilot projects, another QCA analysis is performed which only sees successful scaling out as '1' outcomes and all other forms of effect as '0' outcomes. The previously mentioned '0.5' category, is in this QCA therefore attributed the value '0'. For this analysis, the same data matrix as in paragraph 7.4 will be used.

7.5.1. Generating the truth table

First, the truth needs to be generated, in order visualize all the possible configurations and recognize contradictory configurations. This truth table is shown in table 7.8 below.

Column1	pos_ext	pd_ext	rd_ext	part_ext	proc_ext	Scaling out	Quantity
C6, C22,	0	0	0	0	0	0	5
C23, C24, C25							
C3, C28	0	0	0	0	1	0	2
C18	0	0	0	1	0	1	1
C2, C7	0	0	1	0	0	0	2
C5	0	0	1	0	1	0	1
C16, C26,	0	1	0	0	0	0	3
C27							
C20	0	1	0	1	0	0	1
C19	0	1	1	0	0	0	1
C21	0	1	1	1	0	1	1
C8	1	0	1	0	1	0	1
C12	1	1	0	0	1	1	1
C15	1	1	0	1	0	0	1
C1	1	1	0	1	1	1	1
C11, C14,	1	1	1	0	0	0	3
C17							
C9, C10	1	1	1	0	1	1	2
C4, C13	1	1	1	1	0	0	2

Table 7.8: Final truth table for the external success: scaling out

As can be seen in the truth table, five subsets were identified that resulted in successful scaling out, comprising of six cases in total. Eleven subsets of cases were found that resulted in no scaling out ('0' outcome). No cases resulted in contradictory configurations with regard to the scaling out.

7.5.2. Sufficiency and necessity of single sets

Because of the fact that no contradictory configurations were present in the truth table, this step can be skipped and the calculation of the consistency scores for the single sets can be executed directly. For this, the same formula as used before will be used. Following this calculation formula, the necessity scores of the variables are: Position of the pilot (pos_ext): 0,667; Project design (pd_ext): 0,833; Resource distribution (rd_ext): 0,500; Participants:

(part_ext): 0,500; Process design (proc_ext): 0,667. The participants dimension of the pilot paradox has the highest necessity score. This means that, when looking at the single sets, a project design that focuses on obtaining results that are representative and generalizable can be seen as a sufficient condition for successful scaling out, based on the sufficiency threshold as described in methodological literature (Fiss, 2011; Ragin, 2000; Schneider, 2019). The remaining single sets cannot be regarded as sufficient or necessary. The position of the pilot as being connected to the carrying organizations and a process design that focuses on potential future application areas have the second highest consistency scores.

7.5.3. Boolean minimization and interpretation of the results

The minimization of the data is again done by the use of Tosmana (Cronqvist, 2004). First, the minimization of the configurations that have a '1' outcome was executed. This results in the following formula:

POS_EXT * PD_EXT * rd_ext * PROC_EXT +	POS_EXT * PD_EXT * part_ext * PROC_EXT_+	pos_ext * pd_ext * rd_ext * PART_EXT * proc_ext_+	pos_ext * PD_EXT * RD_EXT * PART_EXT * proc_ext
(C1+C12)	(C9,C10+C12)	(C18)	(C21)

When examining this formula, it stands out that the data still is rather varied. This may be because of the relatively low amount of cases where scaling out took place. The subset POS_EXT * PD_EXT * part_ext * PROC_EXT has the highest coverage score (0,5). This pattern can be recognized at the cases 'Grote Molenbeek', 'Boer Zoekt Water' and 'Proeftuin Oost Stegeren'. A combination of the connection with carrying organisations, a representative project design and the identification of broader embedding potential contributed to successful scaling out in these cases. Rather remarkable is that the participants dimension was more focused on the internal side in these cases. This is a similar pattern as previously seen in paragraph 7.4, at the QCA analysis of conditions for successful scaling up. As previously described, this can mean several things. It can hint at a potential role for boundary spanners in the external success. But it can also imply that both sides of the paradox are present at the same time. The fact that all three cases have a medium Likert-score ranging between 3,17 and 3,50, which is just above the average score of 3,00, puts this aspect more into perspective. Also, when including the logical remainders in the analysis, the participants dimension disappears from the subset with the highest coverage (0,67) and seems important in two other cases, as can be seen in the following formula:

pd_ext * PART_EXT + PD_EXT * PROC_EXT + pos_ext * RD_EXT * PART_EXT (C18) (C1+C9,C10+C12) (C21)

Because of this, it is best to be careful with making general statements about the role of the participants dimension in light of its role with regard to the scaling out potential. What can be concluded, based on these minimal formulas, is that a combination of a representative project design that focuses on gaining generalizable outcomes and a process design that is aimed at identifying potential embedding potentials and includes strategies for broader uptake, positively influences the chances of pilot projects to successfully scale out. This combination

of conditions makes sense when looking at it conceptually. When the involved actors in a pilot have developed a successful innovation and they want to transfer this innovation to other contexts, the innovation needs to be transferable and not too context specific. This requires that the pilot produces representative outcomes that can be trusted and valued by external actors. A small, context-specific pilot where the project design was mainly focused at reducing risks will most likely be assessed less representative by potential external future users. Following this line of reasoning, the process design needs to identify these potential external future actors and also designing strategies on how the successful pilot can potentially be scaled out.

The final step in this QCA procedure is to execute the same Boolean minimization procedure for the '0' configurations. Because of the large share of cases that were unsuccessful in terms of scaling out (78,6%), rather extensive minimal formulas were generated. The following minimal formula can be generated when the logical remainders are not included for reduction:

pos_ext * pd_ext * part_ext +	pos_ext * par proc_ext +	t_ext * POS_ RD_E	EXT * PD_EXT XT * proc_ext	* pd_ext * RD_ + * part_ext * + PROC_EXT +	EXT PD_EXT* rd_ext* - PART_EXT* proc_ext
(C2,C7+C3,C28+C5+C6, C22,C23,C24,C25)	(C2,C7+C6,C22, ,C25+C16,C26,C	C23,C24 27+C19) (C4,C	13+C11,C14,C17)	(C5+C8)	(C15+C20)
pos_ext * pd_ext * p	oart_ext +	POS_EXT * PD_EXT * PART_EXT * proc_ext +	pd_ext * RD_EXT * part_ext * PROC_EXT +	PD_EXT * RD_EXT * part_ext * proc_ext +	pos_ext * PD_EXT * rd_ext * proc_ext
(C2,C7+C3,C28+C5+C6, 5)	C22,C23,C24,C2	(C4,C13+C15)	(C5+C8)	(C11,C14,C17+C19)	(C16,C26,C27+C20)

What stands out is that there is a broad variety of potential subsets that lead to unsuccessful scaling out. In the first formula, two subsets have the highest coverage score. These are:

pos_ext * pd_ext * part_ext	coverage: 0,454
pos_ext * part_ext * proc_ext	coverage: 0,500

These two subsets indicate that in many cases where scaling out was unsuccessful, there was a combination of an absence of a connection with the carrying organizations and an absence of the external representativeness of the participants composition. These pilots operated more at a distance from the carrying organizations and were guided by a small coalition of leading boundary spanners. In one subset, these two conditions were complemented with a project design that was not enough focused on representativeness, but in fact more designed on a limited scale. In the other subset, these two conditions were complemented by a process design that was less focused on the broader embedding potential, but more at being collaboratively set up and ensuring an open learning environment. These results are in line with the expectations that resulted from the literature view. Furthermore, this additional minimization of the '0' outcome provides more clarity when it comes to the role of the participants dimension, that was discussed earlier in this paragraph.

The abovementioned subsets suggest that the absence of the external side of this dimension does in fact influence the scaling up potential of pilots.

Another remarkable point is that the resource distribution dimension is not present in these subsets. In fact, some of the other subsets suggest that in several cases, a presence of a resource distribution that was aimed at efficiency and composed of regular budgets, still can lead to unsuccessful scaling out when combined with other conditions. The following subsets provide a result that contradicts with the expectations that resulted from the literature review:

POS_EXT * PD_EXT * RD_EXT * proc_ext Coverage: 0.227

The subset with the highest coverage is present in cases C4,C13, C11,C14,C17. Two of these cases, C4 ('Bufferboeren Loosbroek') and C13 (Blauwe Dienst - Waterconservering), also resulted in contradictory result in the previous QCA analysis about conditions for scaling up. These cases were both internally unsuccessful, which can explain this contradictory result. C11 ('Pilot Stroomgebied Pepinusbeek') was also internally unsuccessful. However, C14 ('More Crop per Drop Drenthe') and C17 ('Blauwe Dienst – Boeren met Water') were internally successful, but somehow did not result in successful scaling out, even though three of the conditions for external success were present in these cases. When looking more closely into these two pilots, it stands out that both pilots are aimed at reaching a technical innovation within the water system of an agricultural company or several agricultural companies in a region. These company-oriented pilots could be seen as context specific, however the respondents did indicate that the project design was focused on generating representative and generalizable outcomes. The pilots were also both conducted in a coalition of various stakeholders, including local and regional governments, knowledge institutes and executors (blauweengroenediensten.nl, 2008; Delphy, 2021). Based on the minimal formula above, only the external side of the process design was absent in these pilots. This means that the process design was collaboratively designed in an open learning environment, but did not focus enough on the identification of potential future application areas outside the pilot. This suggests that this procedural dimension is rather important for the scaling out of pilot projects. Even when most of the other external conditions are present, the lack of an identification of future application areas and the lack of examining potential scaling up strategies can obstruct the scaling out potential of pilot projects in some cases, especially in these technical company-oriented pilot projects.

Following the QCA guidelines of Rihoux & De Meur (2008), the analysis for the '0' outcome was also conducted with the inclusion of logical remainders. Because of the high number of cases that did not result in successful scaling out, the number of potential minimal formulas was also rather high. These formulas are therefore not shown in the report itself, but can be found in appendix D. When analysing these formulas, it stands out that the subsets 'pos_ext * part_ext', 'part_ext * proc_ext' and 'pd_ext * part_ext' occur the most (with coverage scores of respectively 0,636, 0,636 and 0,500). This reflects the earlier findings about the combination of these conditions and their role with regard to the scaling out potential for pilot projects.

8. Conclusions and discussion

8.1 Answering the research questions

This research aimed to gain knowledge about the specific conditions or combinations of conditions that influence the internal- and external success of drought adaptation pilots in the Dutch dry rural areas. The central theoretical construct that was tested was the pilot paradox, which states that there is a contradictory tension between the conditions for internal- and external success (Van Buuren et al., 2016; Van Buuren et al, 2018). The research was aimed at testing this pilot paradox by using Qualitative Comparative Analysis (QCA), which enables researchers to search for patterns in the data, while at the same time also considering the context of the pilots. The research was organised around the following main question:

What conditions influence the internal and external success of drought-adaptation pilots in dry rural areas in The Netherlands?

In order to answer the main question, the research was subdivided in several sub-questions, that will be answered below.

1. Which pilots are regarded as internally and/or externally successful and what is the nature of these drought-adaptation pilots in the Dutch context?

This research focused on twenty-eight pilot projects in the context of drought adaptation in the Dutch dry rural areas. Out of these twenty-eight cases, twenty-three were regarded internally successful. Five cases did not reach their previously stated innovation goals and were therefore seen as internally unsuccessful. With regard to the external success, thirty-six percent resulted in no broader effects, forty-three percent resulted in successful scaling up and twenty-one precent also resulted in successful scaling out. This research also took into account the different types of broader effects. What stood out, is that most pilot scaled up in a horizontal way. This means that there was spatial growth of the pilots in terms of new pilots, projects or the pilot inspired involved actors and external actors. Only twenty-eight percent of the pilots resulted in actual policy change, which is seen as vertical scaling up.

Based on the gathered descriptive data, several conclusions can also be drawn. First, the pilots that were part of this research had multiple innovation goals. These goals varied from the testing of new technical irrigation measures, new cooperation mechanisms and the more large-scale realization of water retention areas. What stands out, is that in many pilots, multiple (societal) goals were combined in the pilots. Secondly, the nature of the pilots was found to be mainly technical. Also the cooperative and the societal nature reoccurred in many cases. This shows that the drought problem is seen as a societal-wide problem on which various actors have to work cooperatively. However, only a few cases were initiated by societal or private sector actors. In eighty-nine percent of the cases, a governmental actor was the initiator of the pilots. Interestingly, the highest share of pilot projects was not embedded (yet) in programmes or bundling of pilots. This can be because of the relatively new character of the drought issue. Approximately one quarter of the pilots were embedded in a policy programme.

2. Which conditions contribute to the internal success of the drought adaptation pilots? The first QCA analysis into the internal success pointed out that in order to reach internally successfully pilots, a position that is a safe haven at a distance from the carrying organizations or the presence of a coalition of leading boundary spanners that support innovation and can connect different actors within the pilot contributes to the potential of pilots to reach internal success. It also showed that in cases where pilots were no real safe havens, boundary spanners were less present and the project design was not designed on a limited scale and with clear boundaries, internally success was not reached. Another remarkable point was that the process design dimension showed a contradictory result. This dimension was about the potential paradoxical relationship between a process design that is collaboratively set up within a mutual learning environment to reach the internal innovation and a process design that focuses on the broader identification of future embedding potential for the external success. This QCA analysis therefore showed that a process design that is focused on this external identification does not appear to form a barrier for the scaling up potential of pilots, and can even contribute to the internal success.

3. Which conditions contribute to the scaling up of the drought adaptation pilots? The second QCA analysis into the scaling up of pilot projects pointed towards a rather dispersed image as a result of the minimization process of the '1' outcome. However, when analysing more closely, patterns could be recognized. Based on the subsets with the highest coverage scores, the position of the pilot as being connected to the carrying organisations, the project design that focused on external representativeness and generalizability, and a resource distribution aimed at efficiency and designed out of regular budgets together seemed to have the most potential influence on the outcome. The participants dimension seemed to have a rather contradictory role, which suggests that boundary spanners can in some cases also play a role for the external success of pilots. This was a phenomenon that also resulted from the literature review. The QCA analysis also showed that a combination of the absence of 1) a connection with the carrying organisations, 2) the absence of a resource distribution consisting of regular budgets, 3) an absence of external representativeness of the participants composition and 4) the absence of a process design that includes the identification of the broader embedding potential can explain why pilots fail to scale up.

4. Which conditions contribute to the scaling out of the drought adaptation pilots?

The final QCA analysis focused on the conditions that potentially influence the scaling out potential of drought adaptation pilots. Only a few cases resulted in actual scaling out, which suggests that this is rather difficult to achieve. At first, the minimal formula for the '1' outcome also showed a rather dispersed image. However, when analysing the QCA outcomes more closely, patterns were indeed recognized within the conditions. The combination of a project design that focused on external representativeness and generalizable results and a process design aimed at identifying broader embedding potential contribute to the scaling out of drought adaptation pilots. In summary, this means that there is a requirement that focuses on the object of the pilot, which needs to be representative for the problem complexity and generalizable into other contexts. The other factor is that for successful scaling out to other organisations or organisational areas, this future application areas first need to be identified

by involved actors in the initial pilot. In other words, these involved actors need to cast their eyes on the outside world, instead of staying focused on the internal process of the pilot itself.

The minimization of the '0' outcomes showed that an absence of the combination of the external side of four of the dimensions can hinder the scaling out of pilots. First, pilots that are not that connected to the carrying organisations, have a project design that is not focused on the external representativeness and do not have an externally representative composition of involved participants have less chance to successfully scale out the innovation. Secondly, pilots that are not that connected to carrying organisations, do not have an externally representative participants composition and are not focusing their process design on identifying potential future application areas also are less likely to result in successful scaling out.

5. To which extent do the conditions for internal success collide with the conditions for external success?

Now that the various QCA analyses are carried out, it is interesting to analyse whether the paradoxical relationship between the conditions for internal success and external success is present in case of drought adaptation pilots in Dutch dry rural areas. With regard to the internal success, the position of the pilot as being a safe haven and the presence of a coalition of leading boundary spanners proved to contribute to internal successfully pilots. From the minimal formula of the '0' outcomes in this QCA, one can conclude that the combination of the position of the pilot as safe haven, a project design on a limited scale with reduced risks and the presence of boundary spanners also influences the chance to reach internal success. When comparing this to the results of the QCA for scaling up, it stands out that the main conflicting character lies in the first two dimensions of the pilot paradox. While the minimal formula for scaling up showed various alternative potential subsets, the subset with the highest coverage included the presence of both these conditions. This means that for successful scaling up, the pilot should be positioned in connection to the carrying organisations and the project design should focus on generating externally representative and generalizable results that have more chance to be assessed as trustworthy and valuable by these carrying organisations.

Furthermore, when comparing the results of the QCA for internal success with the QCA for scaling out, it stands out that the main conflict also lies in the project design of the pilot. For the internal success of the pilot, the pilot should operate on a limited scale in order to reduce risks within clear boundaries. However, for the potential to scale out the results of the pilot, it should produce outcomes that are representative for the problem complexity and can be generalized to other contexts, combined with a process design that searches for these other contexts. Secondly, there is also a potential conflict within the position of the pilot and the participants dimension. The QCA for internal success proved that the pilot should be conducted as a safe haven at a distance with a presence of leading boundary spanners. However, the minimization of the '0' outcomes for the scaling out potential showed that this could result in unsuccessful scaling out, when also combined with an internally focused project design or process design.

All in all, the abovementioned patterns suggest that a conflict in several dimensions of the pilot paradox is present in the case of drought adaptation pilots in Dutch dry rural areas. However, other dimensions prove to be less conflicting. This suggests that there are potential pathways to break through the paradox. In the next section, these pathways will be further examined in a qualitative way.

8.2 Discussion: searching for pathways and putting the results in perspective

These potential pathways arise when taking on a comparative view towards the various QCA analyses that were conducted. First, these potential pathways can be sought in looking at the gathered patterns and comparing the influence of the dimensions on both the internal and the external success. The QCA analysis into the internal success showed that the process dimension had a contrary effect on the internal success in many of the cases. The other analyses into the conditions for external success showed that an absence of a process dimension that is aimed at identifying the broader embedding potential, together with other dimensions, could hinder successful scaling up and scaling out. Also, the resource distribution dimension appeared only in the minimal formula for successful scaling up. This dimension is about the question whether a pilot should contain a representative resource distribution or should gather additional resources for innovation. Based on a comparison of the various QCA dimensions, one can conclude that a drought adaptation pilot where the resource distribution is constructed out of regular resources aimed at efficiency and a process design that is aimed at identifying the broader embedding potential, does not appear to inhibit the chances to reach internal success within the pilot. This can provide a potential pathway to realize external success, without hindering the internal success of the pilot.

Secondly, as this research has shown, the different dimensions can also play a role in different phases of the pilot. This was shown in the QCA analysis for scaling up, which showed various subsets in which the pilot was not regarded as a safe haven overall, but did result in successful scaling up. This time dimension can potentially be used to overcome the paradoxical character of the various dimensions. In several cases in the pilot, primary and secondary data suggests that the pilot was divided into various phases. It needs to be noted that these were primarily technical pilots, where the development of the innovation itself was executed in a safe haven environment, after which the testing and monitoring was done by including various stakeholders and potential future users, in order to enlarge the chances for the pilot to scale up. Additional research into this time dimension within the process of pilots.

Third, the participants dimension showed another contradictory character. The presence of a coalition of leading boundary spanners was considered important for the internal success. However, the minimal formula that resulted from the QCA analysis into scaling up also hinted on a potential role for boundary spanners in terms of the external success. This viewpoint also resulted from the broader literature review. This is when these boundary spanners do not only serve as connecting and leading actors within the pilot, but also play a role as ambassador on the outside of the pilot. However, the minimal formulas for a '0' outcome in
terms of scaling up and scaling out also indicate that a certain extent of external representativeness of the participants composition can hinder successful scaling up and scaling out when combined with other dimensions.

All in all, the QCA analyses in this research sometimes showed a scattered image, especially the analysis that looked into the potential conditions for the scaling up of pilots. Where potential pathways for success were identified in this research, a large-scale applicable formula for success was not identified. This can partly be the result of the variety of types of innovations that were taken into consideration when generating the database. As was described in the results chapter, a relatively high variety of types of cases was taken into analysis. The scope of this research was demarcated by the geographic location and the soil type (Dutch dry rural areas) instead of using specific innovation types. Perhaps, including a smaller variety of different types of innovations results in clearer patterns. However, the number of pilots focused on drought adaptation that are finished or in the evaluation phase is still rather low. This means that this kind of follow up research into the conditions for success of different types of innovations probably can best be carried out in a few years' time.

The fact that some minimal formulas showed this scattered image, can be a result of the fact that many cases were internally successful, and many cases did not result in successful scaling out. This makes it difficult to identify patterns in the data, because there are more potential combinations of conditions that lead to a certain outcome. However, this can also be a result of the condition selection. Rihoux & De Meur (2008) argue that, when a QCA analysis leads to a long minimal formula, the researcher should reconsider his or her choice of conditions. This was done in this research, by systematically trying different combinations of conditions. It did, however, not result in a clearer pattern. The scattered image in some minimal formulas can also mean that there is another influential condition, that was not included in the QCA analysis (Rihoux & De Meur, 2008; Berg-Schlosser et al., 2009). Therefore, additional research potential pathways to overcome the paradox.

8.3 Recommendations and further research

By means of conducting several QCA analyses that focused on the conditions for internal and external success of drought adaptation pilots in Dutch dry rural areas, this research has shown that the pilot paradox can be quantitatively recognized in the case of drought adaptation pilots on Dutch dry sandy soils. However, the data sometimes also showed rather dispersed patterns and did not result in a specific 'formula for success' to overcome the pilot paradox. The qualitative comparison of the minimal formulas that resulted from the various QCA analyses did, however, identify potential pathways that can help to overcome this paradox. Based on the conducted research, several recommendations can be made. These are both substantive recommendations, but also new questions that arose as a result of this research.

When new drought-adaptation pilots are being set up, it is important to consider the potential paradoxical character of the conditions for internal success and external success, and the implications that this can have on the pursued goals of the pilot. In this research, the patterns

in the data suggest that elements of this paradox play a role in the case of drought adaptation pilots in Dutch dry rural areas. Initiators of new pilots should paint a clear image of the main goals of the pilot. This is, not only the substantive goals, but also what they want to accomplish with the pilot and what the desired broader effect of the pilot is. Additional to this, initiators of pilots should always be aware of the specific context of the pilot and the potential role that this plays in terms of the potential for scaling up. When this picture is sketched, a strategy needs to be designed on how to achieve these goals. The initiator has the ability to strategically include specific elements in the pilot, to make sure that the pilot reaches its desired effect in terms of scaling up. For example, when the goal is to scale out a pilot to other organisations, this research showed that it is important to focus the project design on generating representative and generalizable results and identify potential future application areas. This can for example be done by strategically introducing meetings with potential future users somewhere in the pilot's process and making sure that the pilot has a processing time that is considered representative for the problem complexity. However, this can on its turn inhibit the chances to reach internal success. It can be tried to compensate this, by making sure that there is sufficient experimental freedom during the pilot. Another strategy is to include various phases in the pilot, that differ in character. All in all, this requires a comprehensive balancing act by the initiator(s). This research showed that there is no universal formula for success in this respect.

The research also raised new questions or insights for which further research is required. First, it can be helpful to further research the potential role that boundary spanners can play in terms of the external success. This research hinted on such a potential role, that can prove to be an important element in the strategy for pilots to both reach internal and external success. Secondly, additional research into the process of scaling up in the context of drought adaptation pilots in Dutch dry rural areas is recommended. This research showed that only a small percentage of the pilots resulted in policy change, or vertical scaling up. This is interesting and raises the question whether this was also the goal, or if there are barriers that obstruct the way to vertical scaling up in this context. Finally, further research into the strategic use of different phases in a pilot's design can be helpful to be able to identify more concrete pathways towards success.

Bibliography

Aase, T. H., Chapagain, P. S., & Tiwari, P. C. (2013). Innovation as an expression of adaptive capacity to change in Himalayan farming. *Mountain research and development, 33*(1), 4-10.

Bakel, J. van., Eertwegh, G. van den., Deijl, D. van. & Mensink, J. (2020). SAWAX: Slimme stuw voor optimaal waterbeheer [report]. Retrieved on 29-06-02021 from https://edepot.wur.nl/541886

Basurto, X., & Speer, J. (2012). Structuring the calibration of qualitative data as sets for qualitative comparative analysis (QCA). *Field Methods*, *24*(2), 155-174.

Bauer, A., & Steurer, R. (2014). Multi-level governance of climate change adaptation through regional partnerships in Canada and England. *Geoforum*, *51*, 121-129.

Berg-Schlosser, D., De Meur, G., Rihoux, B., & Ragin, C. C. (2009). Qualitative comparative analysis (QCA) as an approach. In B. Rihoux & C.C. Ragin (Eds.) *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (pp. 1-18). Thousand Oaks: SAGE Publications Inc.

Berkhout, F., Angel, D., & Wieczorek, A. J. (2009). Asian development pathways and sustainable socio-technical regimes. *Technological Forecasting and Social Change, 76*(2), 218-228.

Blauweengroenediensten.nl (n.d.). Blauwe en Groene Diensten: Pilots > Boeren met water [internet source]. Retrieved on 25-06-2021 from

https://www.blauweengroenediensten.nl/nivo%202/CPilots_BMW-1.htm

Boer Zoekt Water (n.d.). Klimaatadaptatie met sub-irrigatie: hoe zit dat? [internet source]. Retrieved on 07-07-2021 from http://www.boerzoektwater.nl/

Bosch, S. van den. & Rotmans, J. (2008). Deepening, Broadening and Scaling up: a Framework for Steering Transition Experiments [essay report]. Retrieved on 21-03-2021 from https://repub.eur.nl/pub/15812/

Boscherini, L., Chiaroni, D., Chiesa, V., & Frattini, F. (2010). How to use pilot projects to implement open innovation. *International Journal of Innovation Management*, *14*(06), 1065-1097.

Brabants Dagblad (2017, april 10th). Opening infiltratievoorziening Vinkelsestraat in Vinkel [news article]. Retrieved on 29-06-2021 from https://www.bd.nl/s-hertogenbosch/opening-infiltratievoorziening-vinkelsestraat-in-vinkel~a472eabb/

Broek, J. van den., Elzakker, I., Maas, T., & Deuten, J. (2020). *Voorbij Lokaal Enthousiasme - lessen voor de opschaling van living labs*. Den Haag: Rathenau Instituut.

Buuren, A. van., Ellen, G. J., & Warner, J. F. (2016). Path-dependency and policy learning in the Dutch delta: toward more resilient flood risk management in the Netherlands? *Ecology and Society*, *21*(4).

Buuren, A. van., Vreugdenhil, H., Verkerk, J. V. P., & Ellen, G. J. (2016). Beyond the pilot paradox How the success conditions of pilots also hinder their up-scaling in climate governance. In Workshop: Beyond experiments: Understanding how climate governance innovations become embedded 25th– 27th April [conference paper]. Retrieved from

http://nws.eurocities.eu/MediaShell/media/Paper_Pilot_Paradox.

Buuren, A. van., Vreugdenhil, H., Van Popering-verkerk, J., Ellen, G.J., Van Leeuwen, C. & Breman, B. (2018). The Pilot Paradox: Exploring Tensions between Internal and External Success Factors in

Dutch Climate Adaptation Projects. In Turnheim, B., Kivimaa, P & Berkhout, F. (Red.), *Innovating Climate Governance* (pp. 145-163). Cambridge, United Kingdom: Cambridge University Press.

Cronqvist, L. (2004, 16-17 sept.). *Presentation of TOSMANA. Adding multi-value variables and visual aids to QCA* [conference paper]. COMPASS Launching Conference, Leuven, Belgium. http://www.compasss.org/wpseries/Cronqvist2004.pdf

Cuéllar-Gálvez, D., Aranda-Camacho, Y., & Mosquera-Vásquez, T. (2018). A model to promote sustainable social change based on the scaling up of a high-impact technical innovation. *Sustainability*, *10*(12), 4532.

Davis, J. (2004). Scaling up urban upgrading. International Development Planning Review, 26(3).

Delphy (2021). More Crop per Drop Drenthe [internet source]. Retrieved on 26-06-2021 from https://delphy.nl/research/more-crop-per-drop-drenthe/

De Moor, K., Berte, K., De Marez, L., Joseph, W., Deryckere, T., & Martens, L. (2010). User-driven innovation? Challenges of user involvement in future technology analysis. *Science and Public Policy*, *37*(1), 51-61.

Dijk, M., De Kraker, J., & Hommels, A. (2018). Anticipating constraints on upscaling from urban innovation experiments. *Sustainability*, *10*(8), 2796.

Doren, D. van., Driessen, P. P., Runhaar, H., & Giezen, M. (2018). Scaling-up low-carbon urban initiatives: Towards a better understanding. *Urban Studies*, *55*(1), 175-194.

Douthwaite, B., Kuby, T., van de Fliert, E., & Schulz, S. (2003). Impact pathway evaluation: an approach for achieving and attributing impact in complex systems. *Agricultural systems*, *78*(2), 243-265.

Engle, N. L. (2011). Adaptive capacity and its assessment. *Global environmental change, 21*(2), 647-656.

Fiss, P.C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal, 54*(2), 393-420.

Franzel, S., Cooper, P., & Denning, G. L. (2001). Scaling up the benefits of agroforestry research: lessons learned and research challenges. *Development in practice*, *11*(4), 524-534.

Friborg, O., Martinussen, M., & Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences, 40*(5), 873-884.

Friedrich, P., Karlsson, A., & Federley, M. (2013). Report 2.1: Boundary conditions for successful Urban Living Labs [report]. Retrieved on 16-03-2021 from

https://www.buildup.eu/en/practices/publications/boundary-conditions-successful-urban-living-labs-suburbanlab-report-21.

Gillespie, S. (2004). Scaling up community-driven development: a synthesis of experience [discussion paper]. Retrieved on 24-03-2021 from https://ageconsearch.umn.edu/record/60312/

Hargadon, A. B., & Douglas, Y. (2001). When innovations meet institutions: Edison and the design of the electric light. *Administrative science quarterly*, *46*(3), 476-501.

Hartmann, A., & Linn, J. F. (2008). Scaling up: a framework and lessons for development effectiveness from literature and practice. *Wolfensohn Center for Development Working Paper*, (5).

Hasluck, C. 2000. Early lessons from the evaluation of the new deal programmes, Research Report No. 49 [review paper]. Retrieved from http://www.opengrey.eu/item/display/10068/570747

Hoogma, R., Kemp, R., Schot, J., & Truffer, B. (2002). *Experimenting for sustainable transport*. Milton Park, United Kingdom: Taylor & Francis.

Hughes, S., Yordi, S., & Besco, L. (2020). The role of pilot projects in urban climate change policy innovation. *Policy Studies Journal*, *48*(2), 271-297.

Hunter, S. T., Cassidy, S. E., & Ligon, G. S. (2012). Planning for innovation: A process oriented perspective. In Mumford, M. (Ed). *Handbook of organizational creativity* (pp. 515-545). Cambridge, US: Academic Press.

Ingram, J., Maye, D., Kirwan, J., Curry, N., and Kubinakova, K. (2015). Interactions between niche and regime: An analysis of learning and innovation networks for sustainable agriculture across Europe. *Journal of Agricultural Education and Extension, 21*(1), 55–71.

Jeuken, A., Beek, E. V., Duinen, R. V., van der Veen, A., Bocalon, A., Delsman, J., ... & Zuurbier, K. (2012). Balancing supply and demand of fresh water under increasing drought and salinisation in the Netherlands [report]. Retrieved on 28-10-2020 from https://library.wur.nl/WebQuery/wurpubs/485321.

KnowH2O (n.d.). SAWAX – Slimme Adaptieve Waterbeheer Extender [internet source]. Retrieved on 07-07-2021 from https://www.knowh2o.nl/projecten/sawax-slimme-adaptieve-waterbeheer-extender/.

Kennisportaal Klimaatadaptatie (n.d.). Regionale klimaat-adaptatiestrategieën [internet source]. Retrieved on 26-10-2020 from https://klimaatadaptatienederland.nl/overheden/nas/regionaal/

Klimaatadaptatie Brabant (2020, June 17). Simpel en goedkoop water vasthouden met nieuwe SIMMBA-techniek [internet source]. Retrieved on 14-05-2021 from https://www.klimaatadaptatiebrabant.nl/k/n442/news/view/3146/2775/simpel-en-goedkoop-watervasthouden-met-nieuwe-simmba-techniek.html

Lavrakas, P. J. (2008). Encyclopedia of survey research methods. London: Sage publications.

Legewie, N. (2013). An Introduction to Applied Data Analysis with Qualitative Comparative Analysis. *Forum: Qualitative Social Research*, *14*(3). https://doi.org/10.17169/fqs-14.3.1961

Lintsen, H. (2002). Two centuries of central water management in the Netherlands. *Technology and culture*, *43*(3), 549-568.

Lumbricus (2021). Onderzoeksprogramma voor een klimaatrobuust bodem- en watersysteem [internet source]. Retrieved on 29-03-2021 from https://www.programmalumbricus.nl/

Martin, S., & Sanderson, I. (1999). Evaluating public policy experiments: measuring outcomes, monitoring processes or managing pilots?. *Evaluation, 5*(3), 245-258.

Massa, S., & Testa, S. (2008). Innovation and SMEs: Misaligned perspectives and goals among entrepreneurs, academics, and policy makers. *Technovation*, *28*(7), 393-407.

McFadgen, B. & Huitema, D. (2016). Are all experiments created equal? A framework for analysis of the learning potential of policy experiments in environmental governance. *Journal of Environmental Planning and Management*, 1-20. Doi: 10.1080/09640568.2016.1256808

Meerkerk, I. van., & Edelenbos, J. (2014). The effects of boundary spanners on trust and performance of urban governance networks: findings from survey research on urban development projects in the Netherlands. *Policy Sciences*, *47*(1), 3-24.

Meijerink, S., & Stiller, S. (2013). What kind of leadership do we need for climate adaptation? A framework for analyzing leadership objectives, functions, and tasks in climate change adaptation. *Environment and Planning C: Government and Policy*, *31*(2), 240-256.

Meur, G. de., Rihoux, B. & Yamasaki, S. (2009). Addressing the Critiques of QCA. In B. Rihoux & C.C. Ragin (Eds.). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (pp. 1-18). Thousand Oaks: SAGE Publications Inc.

Ministerie van Infrastructuur en Waterstaat (2019). Deltaprogramma Zoetwater: De oogst van de klimaatpilots [report]. Retrieved on 02-03-2021 from https://library.wur.nl/WebQuery/hydrotheek/2291957r.nl.

Miranda, S., Tavares, P., & Queiró, R. (2018). Perceived service quality and customer satisfaction: A fuzzy set QCA approach in the railway sector. *Journal of Business Research*, *89*, 371-377.

Muro, M., & Jeffrey, P. (2008). A critical review of the theory and application of social learning in participatory natural resource management processes. *Journal of environmental planning and management*, *51*(3), 325-344.

Naber, R., Raven, R., Kouw, M., & Dassen, T. (2017). Scaling up sustainable energy innovations. *Energy Policy, 110*, 342-354.

Nadin, V. (2007). The emergence of the spatial planning approach in England. *Planning, Practice & Research, 22*(1), 43-62.

Nair, S., & Howlett, M. (2015). Scaling up of policy experiments and pilots: a qualitative comparative analysis and lessons for the water sector. *Water resources management, 29*(14), 4945-4961.

Nam, T., & Pardo, T. A. (2011). Smart city as urban innovation: Focusing on management, policy, and context. In Estevez, E. & Janssen, M. (Eds.). *Proceedings of the 5th international conference on theory and practice of electronic governance* (pp. 185-194). New York: Association for Computing Machinery. ISBN: 978-1-4503-0746-8

Navarro, S., Llinares, C., & Garzon, D. (2016). Exploring the relationship between co-creation and satisfaction using QCA. *Journal of Business Research, 69*(4), 1336-1339.

Oliver, P. (2010). The student's guide to research ethics. United Kingdom: McGraw-Hill Education.

Özerol, G., Troeltzsch, J., Larrue, C., Lordkipanidze, M., Browne, A. L., de Boer, C., & Lems, P. (2016). Drought awareness through agricultural policy: multi-level action in Salland, The Netherlands. In H. Bressers, N. Bressers & C. Larrue (Red.), *Governance for Drought Resilience* (pp. 159-179). Basel: Springer.

Philip, S. Y., Kew, S. F., van der Wiel, K., Wanders, N., & van Oldenborgh, G. J. (2020). Regional differentiation in climate change induced drought trends in the Netherlands. *Environmental Research Letters*, *15*(9), 094081.

Popper, K. R. (1963). Science as falsification. Conjectures and refutations, 1(1963), 33-39.

Provincie Gelderland (2019). Actieplan Klimaatadaptatie [policy document]. Retrieved on 26-10-2020 from https://edepot.wur.nl/545669

Provincie Noord-Brabant (n.d.) Deltaplan hoge zandgronden [map]. Retrieved on 26-10-2020 from https://www.brabant.nl/onderwerpen/water/water-en-klimaat/deltaplan-hoge-zandgronden.

Rae, C., Kerr, S., & Maroto-Valer, M. M. (2020). Upscaling smart local energy systems: A review of technical barriers. *Renewable and Sustainable Energy Reviews, 131*, 110020.

Ragin, C.C. (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C.C. (2000). Fuzzy set social science. Chicago: University of Chicago Press.

Rijksoverheid (2010). Deltaprogramma 2011: Werk aan de delta [internet source]. Retrieved on 26-10-2020, from

https://www.deltaprogramma.nl/deltaprogramma/documenten/publicaties/2010/09/21/deltaprogramm a-2011.

Rijksoverheid (2020). Deltaprogramma 2021: Versnellen en intensiveren om doelen 2050 te halen [internet source]. Retrieved on 26-10-2020 from

https://www.rijksoverheid.nl/actueel/nieuws/2020/09/15/deltaprogramma-2021-versnellen-en-intensiveren-om-doelen-2050-te-halen

Rodima-Taylor, D. (2012). Social innovation and climate adaptation: Local collective action in diversifying Tanzania. *Applied Geography, 33*, 128-134.

Rooy, P. van, & L. Sterrenberg (2000) *Het blauwe goud verzilveren: Integraal waterbeheer en het belang van omdenken.* The Hague: Rathenau Instituut.

Ryghaug, M., Ornetzeder, M., Skjølsvold, T. M., & Throndsen, W. (2019). The role of experiments and demonstration projects in efforts of upscaling: an analysis of two projects attempting to reconfigure production and consumption in energy and mobility. *Sustainability*, *11*(20), 5771. doi:10.3390/su11205771

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public administration*, *80*(1), 1-22.

Samoff, J., Dembélé, M., & Sebatane, E. M. (2013). Scaling up by focusing down: Creating space and capacity to extend education reform in Africa 1. In L Tikly & A.M. Barett (Red.). *Education Quality and Social Justice in the Global South* (pp. 121-138). London: Routledge.

Schneider, C. Q. (2019). Two-step QCA revisited: the necessity of context conditions. *Quality & Quantity, 53*(3), 1109-1126.

Schneider, C. Q., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis.* Cambridge: Cambridge University Press.

Schumpeter, J.A. (1934). *The Theory of Economic Development.* Harvard University Press, Cambridge, MA.

Shi, L., Chu, E., & Debats, J. (2015). Explaining progress in climate adaptation planning across 156 US municipalities. *Journal of the American Planning Association*, *81*(3), 191-202.

Sengers, F., Wieczorek, A. J., & Raven, R. (2019). Experimenting for sustainability transitions: A systematic literature review. *Technological Forecasting and Social Change*, *145*, 153-164.

Silverman, D. (2013). Doing qualitative research. London: SAGE Publications Ltd.

Smithers, J., & Blay-Palmer, A. (2001). Technology innovation as a strategy for climate adaptation in agriculture. *Applied Geography, 21*(2), 175-197.

STOWA (n.d.). Klimaatadaptatie in de praktijk KLIMAP [internet source]. Retrieved on 26-10-2020 from https://www.stowa.nl/onderwerpen/klimaatadaptatie/zoetwatertekort-droogte/klimaatadaptatie-de-praktijk-klimap.

STOWA (2020). Lumbricus gaat, KLIMAP komt [internet source]. Retrieved on 29-03-2021 from https://www.stowa.nl/nieuws/lumbricus-gaat-klimap-komt

Taylor, B. M., Harman, B. P., & Inman, M. (2013). Scaling-Up, Scaling-Down, and Scaling-Out: Local Planning Strategies for Sea-Level Rise in N ew S outh W ales, A ustralia. *Geographical Research*, *51*(3), 292-303.

Thiel, S. van. (2014). *Research methods in public administration and public management: an introduction*. London: Routledge.

Turner, J. R., & Müller, R. (2003). On the nature of the project as a temporary organization. *International journal of project management, 21*(1), 1-8.

Thomann, E., & Maggetti, M. (2020). Designing research with qualitative comparative analysis (QCA): Approaches, challenges, and tools. *Sociological Methods & Research, 49*(2), 356-386.

Unie van Waterschappen (2018). Waterschapskaart 2018 [image]. Retrieved on 26-10-2020 from https://www.uvw.nl/publicatie/waterschapskaart/

Varis, M., & Littunen, H. (2010). Types of innovation, sources of information and performance in entrepreneurial SMEs. *European Journal of Innovation Management*, *13*(2), 128-154.

Vidaurre, R., Stein, U., Browne, A., Lordkipanidze, M., Furusho, C., Goedeking, A., ... & Homann, C. (2016). Eifel-Rur: Old Water Rights and Fixed Frameworks for Action. In H. Bressers, N. Bressers & C. Larrue (Red.), *Governance for Drought Resilience* (pp. 67-82). Basel: Springer.

Vreugdenhil, H. (2010). Pilot Projects in Water Management: Practicing Change and Changing Practice [dissertation]. Retrieved on 04-03-2021 from http://repository.tudelft.nl/islandora/object/uuid:91bda74b-cf5d-4caf-8159-1e57650291cc?collection=research Vreugdenhil, H., Slinger, J., Thissen, W., & Rault, P. K. (2010). Pilot projects in water management. *Ecology and Society*, *15*(3).

Vreugdenhil, H., Taljaard, S., and Slinger, J. H. (2012). Pilot projects and their diffusion: A case study of integrated coastal management in South Africa. *International Journal of Sustainable Development*, *15*(1–2), 148–172.

Waterforum (2020, June 16th). KLIMAP-project: naar een regionale uitrol klimaatadaptieve maatregelen voor hooggelegen zandgronden [internet source]. Retrieved on 08-04-2021 from https://www.waterforum.net/klimap-project-naar-een-regionale-uitrol-klimaatadaptieve-maatregelen-voor-hooggelegen-zandgronden/

Waterschap Aa en Maas (2021, March 4). Provincie en waterschappen werken samen aan natuur- en wateropgaven in Brabant [internet source]. Retrieved on 24-03-2021 from https://www.aaenmaas.nl/actueel/nieuws/2021/maart/provincie-waterschappen-werken-samen/

Waterschap De Dommel (n.d.). Waterschap ontwikkelt slimme Simmba stuw [internet source]. Retrieved on 14-05-2021 from https://www.dommel.nl/waterschap-ontwikkelt-slimme-simmba-stuw

Waterschap Peel en Maasvallei (2010). Eindrapportage Nieuw Limburgs Peil [policy document]. Retrieved on 26-10-2020 from https://edepot.wur.nl/168921

Waterschap Rijn en IJssel (n.d.). Meer efficiency én betrokkenheid [internet source]. Retrieved on 14-05-2021 from https://www.wrij.nl/waterbeheerplan/strategie-beleid/uitdagingen/samenwerkeninwoners/#item_4316

Waterschap Vechtstromen (n.d.). SAWAX: de 'slimme' stuw [internet source]. Retrieved on 07-07-2021 from https://edepot.wur.nl/427925

Wester, K. L. (2011). Publishing ethical research: A step-by-step overview. *Journal of Counseling & Development, 89*(3), 301-307.

Wieczorek, A. J., Raven, R., & Berkhout, F. (2015). Transnational linkages in sustainability experiments: A typology and the case of solar photovoltaic energy in India. *Environmental Innovation and Societal Transitions, 17*, 149-165.

Winden, W. van. (2016). Smart city pilot projects, scaling up or fading out? Experiences from Amsterdam. Paper presented at Regional Studies Association Annual Conference, Graz, Austria. Retrieved from

https://www.narcis.nl/publication/RecordID/oai:hbokennisbank.nl:amsterdam_pure%3Aoai%3Apure.h va.nl%3Apublications%2F76b2f91f-8783-4862-b0a0-c88c82dea08c.

Winden, W. van., & Buuse, D. van den. (2017). Smart city pilot projects: Exploring the dimensions and conditions of scaling up. *Journal of Urban Technology*, *24*(4), 51-72.

WUR (n.d.). Hoge Zandgronden in Nederland, met zowel Oost- als Zuid Nederland [image]. Retrieved on 29-03-2021 from https://www.wur.nl/nl/project/Lumbricus-het-kennisprogramma-voor-een-klimaatrobuust-bodem-en-watersysteem.htm

Appendix A) Outline of the survey

- 1. What is the name of the pilot?
- 2. What is the location of the pilot?
- 3. What was the starting date of the pilot?
- 4. What was the ending date of the pilot?
- 5. Who was the initiator of the pilot?
- 6. Which parties were involved in the pilot?
- 7. In which phase is the pilot currently?
 - Exploration
 - Design
 - Implementation
 - Monitoring and evaluation
 - Finished
- 8. How would you describe the goals of the pilot?
 - Better understand a certain problem
 - Develop or test an innovation
 - Solve a problem
 - Show an innovation to the outside world
 - Improve the policy implementation
 - Show a certain action to the outside world
 - Distract attention away from other problems
 - Other
- 9. How would you describe the character of the pilot?
 - Technical
 - Societal
 - Conceptual
 - Cooperative
 - Financial
 - Contractual
 - Organisational
 - Other
- 10. How would you describe the geographical level of the pilot?
 - Company-level: one (agricultural) company
 - Locally: one or several locations within a municipality
 - Regionally: several locations within a water authority region
 - Regionally: several locations within a province
 - Interregionally: several locations in various regions or provinces
 - Other
- 11. Is the pilot part of a programme or bundling pilots?
 - No
 - Yes: living lab/innovation platform
 - Yes: research programme
 - Yes: policy programme
 - Yes: the pilot is part of a bundling pilots aimed at the same problem
 - Other

12. Is the pilot a follow-up of previous developments with regard to the same problem?

- Yes
- No

B) if yes: please describe these developments.

13. To what extent were the previously stated goals of the pilot reached?

No goals were reached	Most goals were not reached	The division between reached and not reached was equal	Most goals were reached	All goals were reached

14. Please indicate to what extent you agree with the following statements:

The pilot resulted in cooperation between actors that did not cooperate before the pilot

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

The results of the pilot formed a substantive answer to the previously stated knowledge questions

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

The pilot resulted in new knowledge about the problem

Strongly disagree	Disagree	Neutral	Agree	Strongly agree	

15. Is there at this moment diffusion of the pilot(s results)?

- [] No
- [] Yes, in the form of new pilot(s)
- [] Yes, in policy
- [] Yes, in regular projects
- [] Yes, as inspiration for others
- [] Other:

16. If yes, on what scale did the diffusion take place?

- [] Within the area of the local municipality
- [] Within the area of the waterboard authority
- [] Within the area of the province
- [] From local scale towards a regional scale
- [] From local scale towards national scale
- [] Other:

17. If yes, what is being diffused?

- [] The innovation
- [] Parts of the innovation

- [] Ideas
- [] Measurement-, modelling or design-methods
- [] Collaboration/partnerships
- [] Procedures
- [] Other:

The following questions all had a five-point Likert scale ranging from strongly disagree to strongly agree

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

- The was much freedom to try out new ideas, independent from the mother organisations of the participants
- The pilot was allowed to fail
- In the pilot, more flexible rules were present than in regular projects
- 19. Please indicate to what extent you agree with the following statements:
 - The pilot remained connected to the mother organisations of the participants
 - Participants needed to report their progress to the mother organisations
 - The legal experimenting room was representative for the laws and regulations in regular projects
- 20. Please indicate to what extent you agree with the following statements:
 - The pilot was executed on a limited scale to reach innovation
 - The pilot was clearly demarcated in (geographical) space and time
 - The pilot was executed on a limited scale to reduce risks
- 21. Please indicate to what extent you agree with the following statements:
 - The pilot was representative for the problem complexity
 - The results of the pilot are valued by the carrying organisations
 - The outcomes of the pilot can be easily generalized to a larger programme or project
- 22. Please indicate to what extent you agree with the following statements:
 - During the pilot, additional budget for innovation and creativity became available
 - During the pilot, additional manpower and expertise became available
 - During the pilot, additional time to experiment, monitor or analyse became available
- 23. Please indicate to what extent you agree with the following statements:
 - The financial means that became available during the pilot were part of regular budgets
 - The division of resources and tasks is aimed at efficiency
 - The division of resources and tasks is representative for regular projects with regard to drought
 - The pilot included practicing with mobilising regular resources, outside the additional innovation resources
- 24. Please indicate to what extent you agree with the following statements:
 - The pilot consists of a relatively small group enthusiastic front-runners
 - I would describe the participants as 'open to new knowledge and taking on risks'
 - The pilot included strong leadership by one or a few actors

- 25. Please indicate to what extent you agree with the following statements:
 - A wide scale of potential stakeholders was involved in the pilot
 - Potential future users were involved in the pilot
 - Colleagues and potential criticasters from the carrying organisations were involved in the same learning process (i.e., by receiving updates or joining meetings)
- 26. Please indicate to what extent you agree with the following statements:
 - I would describe the pilot as an open learning environment
 - The various actors collaboratively designed the process of the pilot
 - Participants could discuss beyond the traditional role distribution
- 27. Please indicate to what extent you agree with the following statements:
 - During the pilot, potential future application locations/areas were searched
 - The outcomes of the pilot were fed back to actual policy questions
 - Strategies for scaling up were part of the design of the pilot

Appendix B) Single and multiple response test

Name of the pilot	PROCESS	PARTICIPANTS	RESOURCE DISTRIBUTION	PROJECT DESIGN	POSITION
F2AGRI Boer Bier Water	2.83	2.83	3.14	2.67	2.5
Boer, Bier, Water (F2AGRI)	1.83	2.67	3.29	2.17	3
Boer, Bier, Water	2.5	3	3.43	2	3
KlimaatKlaor?!	3.17	3.33	2.57	3.33	3.33
Klimaat Kloar ?!	3.33	3	2.71	3	3.17
Klimaat Klaor?!	3.5	2.83	2.43	3.17	3
Hergebruik effluent voor de landbouw via subirrigatie	2.83	3.33	3.43	3.17	3.5
Hergebruik effluent RWZI	3	3.5	3.14	3.17	3.83
Bufferboeren	3.33	2.67	2.43	2.83	2.5
Bufferboeren	3.17	2.83	2.14	2.83	2.67
Waterwijs Boeren Haarlo	2.83	3.17	2.14	4.5	4
bevloeiing in waterwingebied Haarlo	2.33	3.67	1.86	5	4.33

Dimension (questions)	All items	Without .1	Without .2	Without .3	Without .4 (reversed .1)	Without .5 (reversed .2)	Without .6 (reversed .3)	Without .7 (reversed .4)
Position of the pilot (18 and 19 reversed)	0,774	0,724	0,743	0,759	0,741	0,751	0,722	
Project design (20 and 21 reversed)	0,763	0,694	0,757	0,712	0,696	0,750	0,752	
Resource distribution (22 and 23 reversed)	0,746	0,619	0,656	0,664	0,763	0,768	0,772	0,706
Participants (24 and 25 reversed)	0,618	0,620	0,616	0,639	0,418	0,398	0,583	
Process design (26 and 27 reversed)	0,635	0,645	0,620	0,582	0,592	0,558	0,533	

Appendix C) Reliability test results

Appendix D) Overview of the minimal formulas that were not included in the results and analysis chapter

procprocint * int * PDpos-INT * pos-int * pd-**PROC-INT *** POS-INT * PD-INT * PROC-INT * POS-INT * int * rd-int int * rd-int * POS-INT * pd-PART-INT + RD-INT + * pd-int int + PART-INT PAR * RD-INT Т-INT + + (C2.C7+C3.C28+C5+C6.C22.C23) (C6.C22.C23.C24.C25+C18+C20) (C19+C20+C21+C2 (C1+C1 (C5+C (C9,C10+C14, C17) ,C24,C25) +C26,C27)6,C27) 2) 8) procint * pos-int * proc-int pos-PROC-INT * POS-INT * PD-INT * PROC-INT * POS-INT * int * pd-int * rdrd-int * POS-INT * pd-PART-INT + RD-INT + int * PARTpd-int PARTint + * RD-INT INT + INT + (C2,C7+C3,C28+C5+C6,C22,C2 (C6,C22,C23,C24,C25+C18+C20 (C19+C20+C21+C2 (C1+C1 (C8+C9,C (C9,C10+C14, 3,C24,C25) +C26,C27) 6,C27) 10) C17) 2) procint * proc-int PROCpos-**PROC-INT** * pos-int INT * pd-POS-INT * PD-INT * PROC-INT * POS-INT * int * rd-int * int * rd-int POS-INT * pd-PART-INT + RD-INT + pd-int PART-* PARTint + * RD-INT + INT INT + (C2,C7+C3,C28+C5+C6,C22,C23 (C6,C22,C23,C24,C25+C18+C20 (C19+C20+C21+C2 (C1+C1 (C8+C9,C (C14,C17+C ,C24,C25) +C26,C27)6.C27) 10) 19) 2) Scaling up without logical remainders POS-EXT * pos-ext * rdpos-ext * PDpos-ext * pdpos-ext * pd-ext POS-EXT * PD-EXT * PD-EXT * rdext * PART-EXT * RDext * RD-EXT part-ext * EXT * procext * PROC-EXT * proc-RD-EXT * part-ext + PROC-EXT + part-ext + EXT + ext + ext (C2, C7+C5)(C3,C28+C5) (C18+C20) (C19+C21) (C1+C12) (C9,C10+C11,C14,C17) POS-EXT * pos-ext * PDpos-ext * rdpos-ext * pd-ext POS-EXT * PD-EXT * pos-ext * RD-PD-EXT * rdext * PART-EXT * RD-* part-ext * EXT * part-ext * EXT * proc-EXT * procext * PROC-RD-EXT * part-ext + PROC-EXT + proc-ext + EXT + ext + ext (C9,C10+C11,C14,C17) (C19+C21) (C1+C12) (C2,C7+C19) (C3, C28+C5)(C18+C20) POS-EXT * pos-ext * rdpos-ext * PDpos-ext * RDpos-ext * pd-ext POS-EXT * PD-EXT *

Internal success without logical remainders

PD-EXT * rd-

ext * PROC-

EXT +

(C1+C12)

EXT * part-ext *

proc-ext +

(C2, C7+C19)

part-ext *

PROC-EXT

(C3, C28 + C5)

+

EXT * PART-

EXT * proc-

(C20+C21)

ext

ext * PART-

EXT * proc-

ext +

(C18+C20)

RD-EXT * part-ext +

(C9,C10+C11,C14,C17)

Scaling out (VALUE 0) with logical remainders

			POS_E	X		noc ovt *
pos_ext * part_e	ext +		T * pd_ext +	POS_E proc_e	XT * xt +	PD_EXT * rd_ext
(C2,C7+C3,C28+C5 19)	5+C6,C22,C23,C24,C25+C	16,C26,C27-	+C (C8)	(C4,C13-)	+C11,C14,C17+	C15 (C16,C26,C27+C20)
Simplifying Assum	ptions 11					
			POS EX			
pos_ext * part_e	ext +		T * [—] pd_ext +	POS_EX ⁻ proc_ext	Γ* +	PD_EXT * rd_ext * proc_ext
(C2,C7+C3,C28+C5 +C19)	5+C6,C22,C23,C24,C25+C	16,C26,C27	(C8)	(C4,C13+C1 5)	1,C14,C17+C1	(C15+C16,C26,C27+C2 0)
Simplifying Assum	ptions 10					
pos_ext * part_e	ext +		POS_EX proc_ext	T * +	pd_ext * RD_EXT	pos_ext * + PD_EXT * rd_ext
(C2,C7+C3,C28+C5 +C19)	5+C6,C22,C23,C24,C25+C	16,C26,C27	(C4,C13+C 5)	11,C14,C17+	+C1 (C2,C7+C5)	6+C8 (C16,C26,C27+C2 0)
Simplifying Assum	ptions 11					
pos_ext * part_e	ext +	l	POS_EXT	* +	pd_ext * RD_EXT +	PD_EXT * rd_ext * + proc_ext
(C2,C7+C3,C28+C5 7+C19)	5+C6,C22,C23,C24,C25+C	16,C26,C2(C4,C13+C11 15)	,C14,C17+C	(C2,C7+C5+C 8)	(C15+C16,C26,C27+C 20)
Simplifying Assum	ptions 10					
			∨⊤ *			pos_ext *
pos_ext * part_e	ext +	POS_E	xi [*] t +	pd_ext * p	oart_ext +	PD_EXT * rd_ext
(C2,C7+C3,C28+C5 6,C26,C27+C19)	5+C6,C22,C23,C24,C25+C	1 (C4,C13+ 7+C15)	C11,C14,C1	(C2,C7+C3, 24,C25+C8)	C28+C5+C6,C2	2,C23,C (C16,C26,C27+ C20)
Simplifying Assum	ptions 9					
			- *			PD_EXT *
pos_ext * part_e	ext +	proc_ext	' ₊ p	d_ext * pa	rt_ext +	rd_ext * proc_ext
(C2,C7+C3,C28+C5 16,C26,C27+C19)	5+C6,C22,C23,C24,C25+C	(C4,C13+C 17+C15)	11,C14,C (0 C	C2,C7+C3,C2 24,C25+C8)	28+C5+C6,C22,0	C23, (C15+C16,C26,C2 7+C20)
Simplifying Assum	ptions 8					
pos_ext * part_e	ext +		POS_EX ⁻ proc_ext	Γ* +	pd_ext * PROC_E +	pos_ext * XT PD_EXT * rd ext
(C2,C7+C3,C28+C5 +C19)	5+C6,C22,C23,C24,C25+C	16,C26,C27	(C4,C13+C1 5)	11,C14,C17+	C1 (C3,C28+C5)	
Simplifying Assum	ptions 13					
		-		*	pd ext*	
pos_ext * part_e	ext +	Р Р	proc_ext +	F	PROC_EXT	proc_ext
(C2,C7+C3,C28+C5 27+C19)	5+C6,C22,C23,C24,C25+C	16,C26,C ((1	C4,C13+C11 5)	,C14,C17+C	(C3,C28+C5+C 8)	(C15+C16,C26,C27+C 20)
Simplifying Assum	ptions	8				
POS_EXT * proc_ext +	pd_ext * part_ext	+	part_ext *	* proc_ext	+	pos_ext * PD_EXT * rd_ext

(C4,C13+C11,C14,C1 7+C15)	(C2,C7+C3 C24,C25+C	8,C28+C5+C C8)	C6,C22,C23,	(C2,C7+C6 16,C26,C2	6,C22,C23,0 7+C19)	C24,C25	5+C11,C14	4,C17+	-C (C16,C26,C27+ C20)
Simplifying Assumpt	ions 8								
POS_EXT * proc_ext +	pd_ext * p	oart_ext	+	part_ext *	proc_ext	+			PD_EXT * rd_ext * proc_ext
(C4,C13+C11,C14,C 17+C15)	(C2,C7+C3, C24,C25+C	C28+C5+C 8)	6,C22,C23,	(C2,C7+C6,C C16,C26,C2	C22,C23,C2 7+C19)	24,C25+	C11,C14,0	C17+	(C15+C16,C26,C2 7+C20)
Simplifying Assumption	ions 6								
POS_EXT * proc_ext +	pd_e PRO +	xt * C_EXT	part_ext *	proc_ext	+				pos_ext * PD_EXT * rd_ext
(C4,C13+C11,C14,C17 15)	7+C (C3,C2)	28+C5+C8	(C2,C7+C6, 7+C19)	C22,C23,C24	4,C25+C11,	C14,C1	7+C16,C2	6,C2	(C16,C26,C27+C2 0)
Simplifying Assumption	ions 12								
POS_EXT * proc_ext +	pd_ex PROC +	dt* C_EXT p	eart_ext * p	proc_ext -	+			PC * p	0_EXT * rd_ext roc_ext
(C4,C13+C11,C14,C17 C15)	7+ (C3,C2 8)	8+C5+C ((C	C2,C7+C6,C 27+C19)	22,C23,C24,0	C25+C11,C	14,C17-	+C16,C26,	, (C1 20)	5+C16,C26,C27+C
Simplifying Assumption	ions 10								
pd_ext * part_ext	+	part_e	ext * proc_	_ext +			POS_E * RD_E * PART_I T +	XT XT F r EX f	PD_EXT * d_ext * proc_ext
(C2,C7+C3,C28+C5+C 4,C25+C8)	C6,C22,C23,	,C2 (C2,C7 ,C26,C	+C6,C22,C2 27+C19)	23,C24,C25+0	C11,C14,C1	7+C16	(C4,C13)	(+	C15+C16,C26,C27 -C20)
Simplifying Assumption	ions 7								
pd_ext * PROC_EXT par +	t_ext * pr	oc_ext +	ŀ			POS_ RD_E PART +	EXT * XT * _EXT	PD_ proc	EXT * rd_ext * _ext
(C3,C28+C5+C8) (C2, C19	,C7+C6,C22))	2,C23,C24,C	C25+C11,C1	4,C17+C16,C	C26,C27+	(C4,C1	3)	(C15+)	-C16,C26,C27+C20
Simplifying Assumption	ions 10								