Dissertation

Proximity and collaborative knowledge production in the water sector

Pieter W. Heringa

KWR Watercycle Research Institute

Rathenau Instituut

The **Rathenau Instituut** promotes the formation of political and public opinion on science and technology. To this end, the Institute studies the organization and development of science systems, publishes about social impact of new technologies, and organizes debates on issues and dilemmas in science and technology.

KWR Watercycle Research Institute

"Bridging science to practice" – that is the mission of KWR Watercycle Research Institute. Our shareholders are the ten Dutch water companies. KWR researchers work at the interface of science, business and society. Their strength lies in their ability to translate scientific knowledge into practical and implementable solutions for end-users in the Dutch and European water sectors. KWR has developed a solid reputation as a top-level innovation accelerator and international network builder.

Proximity and collaborative knowledge production in the water sector

Rathenau Instituut
Anna van Saksenlaan 51

Postal address:
Postbus 95366
2509 CJ Den Haag

Telefoon: +31 (0)70 342 15 42
E-mail: info@rathenau.nl
Website: www.rathenau.nl

Publisher: Rathenau Instituut
Lay-out: Boven de Bank
Coverphoto: Getty Images
Printing: Drukkerij Quantes, Rijswijk

This book is printed on FSC certified paper.

ISBN/EAN: 978-90-77364-61-1

Preferred citation:
Heringa, P.W., Proximity and collaborative knowledge production in the water sector, Den Haag, Rathenau Instituut 2014

This dissertation is based on research the author has done as an employee of the Rathenau Instituut. This text is the author's PhD thesis. The views in this dissertation are not necessarily the same as those of the Rathenau Instituut.

Proximity and collaborative knowledge production in the water sector

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 23 januari 2015 om 12.30 uur
door Pieter Willem HERINGA
Master of Science
geboren te Borsele.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. W. van Vierssen

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. W. van Vierssen,	Technische Universiteit Delft, promotor
Prof. dr. L. Hagdorn,	Vrije Universiteit Amsterdam
Prof. dr. R. Tijssen,	Universiteit Leiden
Prof. dr. K. Frenken,	Universiteit Utrecht
Prof. dr. H.H.G. Savenije,	Technische Universiteit Delft
Dr. L.K. Hessels,	Rathenau Instituut
Dr. M.W. van der Zouwen,	KWR Watercycle Research Institute
Prof. dr. N.C. van de Giessen,	Technische Universiteit Delft, reservelid

Dr. L.K. Hessels en Dr. M.W. van der Zouwen hebben als begeleider in belangrijke mate aan de totstandkoming van het proefschrift bijgedragen.

ISBN 978-90-77364-61-1

Content

1 Introduction

Between January 2011 and June 2014, companies worldwide spent more than 84 billion dollars on improvements in their processes to conserve, manage or obtain water (Clark, 2014). Even in the Netherlands, considered as a global reference for water management, there are increasing concerns about water quality and the resilience of freshwater ecosystems (OECD, 2014b). This illustrates the need for innovations and knowledge production in the water sector. The processes of innovating and producing knowledge are said to be hindered by institutional fragmentation and badly managed multi-level governance (Thomas and Ford, 2005; OECD, 2011). The multiplicity of actors involved, each with their own motivations and stakes, creates a strong need for joint decision-making and also for collaborative knowledge production (OECD, 2011).

In general, knowledge production is increasingly seen as an inherently collaborative process. There has been a growing interest in collaborative knowledge production in both the scientific literature and the policy discourse. The literature agrees on the benefits of collaboration in knowledge intensive processes (Hagedoorn et al., 2000; Katz and Martin, 1997). Despite this, there are still many questions regarding the factors that shape the patterns of collaboration. Several topical themes in research, as well as developments in research policy, illustrate the relevance of these questions. Below, I elaborate on two of them: the role of geographical distance and organisational diversity in collaborative knowledge production processes.

The idea that geographical distance matters for collaborative knowledge production has intrigued scholars for a long time, and many research policies aim to affect its role. Hagstrom (1965) found that even within a building, the propensity for daily interaction quickly decreases if the distance between researchers increases. In Europe in particular there is a strong interest in the role of geographical distance in collaborative knowledge production. This is clear in both the scientific literature (Hoekman et al., 2010; Chessa et al., 2013) and in policy. The latter has a specific focus on crossing geographical borders. National research policies promote international collaboration and exchange in knowledge production, and the European Union (EU) has stipulated its goals for the development of a European Research Area (ERA), where the circulation of knowledge is not hindered by geographical borders (Delanghe et al., 2009). The concept of the ERA is accompanied by specific policy instruments, such as European Technology Platforms, Joint Programming Initiatives and European Innovation Partnerships. These instruments influence the dynamics and configuration of collaborative knowledge production, for example by promoting collaboration across national borders. Such European policy instruments are particularly relevant for the water sector, for which there are relatively many dedicated instruments, such as the European Technology Platform called Water Supply and Sanitation Technology Platform (WssTP), the Joint Programming Initiative "Water

challenges for a changing world" (Water JPI), and the European Innovation Partnership on Water (EIP-W). Alongside generic instruments such as the EU Framework Programmes (FPs), which also include relevant programmes for the water sector, it is reasonable to assume that these instruments have a considerable influence on the configuration of collaborative knowledge production in the water sector in general and on the influence of geographical proximity in particular.

Along with this interest in geographical distance, the literature also pays increasing attention to Triple Helix collaboration. This is the conceptual idea that collaborations between partners from diverse organisational backgrounds promote and drive innovation (Etzkowitz and Leydesdorff, 2000). The idea is elaborated upon in a large body of scientific literature (Hessels and Van Lente, 2008), and has also been implemented in research policies at both national and international level (Nieminen and Kaukonen, 2001; Potì and Reale, 2007). The European Framework Programmes, for example, include incentives to stimulate research organisations to involve other stakeholders in producing knowledge, and also to stimulate firms to involve public research organisations in precompetitive innovation processes (Delanghe et al., 2009).

These two elements – geographical distance and organisational diversity in research collaboration – are often intertwined in practice. Many research policies incorporate incentives to stimulate both. Also, conceptually speaking, geographical distance and organisational diversity can be considered to be different dimensions of a generic phenomenon. In both cases, two actors perceive a certain kind of distance (or proximity) between them, and the effort required to overcome this distance may influence the propensity of the actors to collaborate. This is acknowledged in the body of literature on proximity in economic geography and innovation studies, which started with the notion of geographical proximity, but has since incorporated many other dimensions, including organisational, social and cognitive proximity (Boschma, 2005). The central premise is that those in closer proximity have a tendency to collaborate as it is easier and less costly to do so. However, it has also been found that too great a proximity may lead to lock-in and lower the innovation potential of collaboration (Knoben and Oerlemans, 2006; Boschma, 2005; Nooteboom et al., 2007). I will elaborate on the concept of proximity more extensively when discussing the conceptual framework below.

1.1 Scientific contribution of this thesis

The innovative contribution of this thesis lies to a considerable extent in the fact that it combines ideas from the existing literature in a novel way and applies the theoretical concepts to a specific empirical context. I have identified a niche in the literature that concerns a combination of three elements. All three are described in the literature on proximity in the fields of innovation studies and economic geography, although they also have connections with other bodies of literature. First, there is a need for empirical studies on proximity that are based on applied fields of research with strategic relevance. Most studies to date

analyse patterns in basic sciences, such as nanotechnology (Autant-Bernard et al., 2007; Cunningham and Werker, 2012), biotechnology (Ter Wal, 2009), or computing and life sciences (Weterings and Ponds, 2009). Others examine a cross-section of all sciences (Hoekman et al., 2010; Bouba-Olga et al., 2012). There are indications that the effect of proximity is different in applied research and knowledge production, also because applied research often involves organisationally more diverse actors (e.g. d'Este and Iammarino, 2010). This difference will be the strongest in fields with research questions that are contextualised to local conditions. Second, the literature suggests that the effect of proximity differs across different levels of geographical aggregation. Katz (1994), for example, showed that geographical proximity has an effect in Canada, the UK and Australia, but the effect is much stronger in the UK than in Canada or Australia. Such differences across countries raises the question of how the geographical effect within a nation compares to the effect at transnational or subnational level. Thus far, there seems to be no analysis of proximity in a particular research field that combines several geographical levels. Third, recent studies have demonstrated the relevance of analysing several dimensions of proximity simultaneously to gain more insight into potential interactions (Ter Wal, 2009; Cunningham and Werker, 2012). It is suggested, for example, that geographical proximity functions as an auxiliary dimension to others (Boschma, 2005; Ter Wal, 2009). To understand this mechanism of proximity better, it is necessary to know more about how the dimensions of proximity interact with each other.

My thesis contributes to filling the gaps in this niche. First, I have chosen the water sector as the object of study. Water is a strategic research field, where a multiplicity of actors from a wide variety of organisational backgrounds produce new knowledge: this not only includes 'traditional' knowledge producers such as universities and public research institutes, but also governmental organisations and NGOs (OECD, 2011; Frijns et al., 2013). Moreover, water is an applied field of research, where knowledge is often applied to and contextualised in specific local conditions and circumstances. However, water is rarely the object of study in literature on research networks and the dynamics of knowledge production. There is an extensive body of literature on the water sector and its knowledge production, but it pays little attention to research networks and the dynamics of research collaboration (with the exceptions, for example, of Frijns et al. (2013), and Pahl-Wostl et al. (2007) on social learning and collaborative knowledge production). Second, I combine two geographical scopes. The first two studies in this thesis are at the national level, while the latter two concern the transnational level. This combination of studies at different levels but dealing with a single research field may offer new insights into the relevance of geographical scale. Third, the various dimensions of proximity are studied simultaneously in the multivariate models. This allows better insight into the effect of each dimension in itself, and into interaction between the dimensions. In this way, for example, I can further examine whether several dimensions of proximity act as substitutes for each other. An additional, fourth element that gives added value to this thesis

is that I have triangulated my findings across different types of data and methods of analysis. This is crucial to obtain a comprehensive overview of the effect of proximity, because the different data types have their inherent biases; for example, they focus on specific outcomes of research or on actors from specific organisational backgrounds.

This thesis is not only relevant to scholars in the field of science studies. Insights into the dynamics of a heterogeneous and applied research field such as water may offer guidance to policymakers in designing instruments and regulations that steer collaborative knowledge production. It may also be of use to the actors involved in developing networks and selecting collaborators. The thesis provides an empirical basis to further develop European science and innovation policy and may also assist national policymakers to design a strategy for participation in the larger European programmes. This relevance to policy will be further elaborated upon later in this chapter.

1.2 Research questions

This thesis revolves around two key elements: collaborative knowledge production and the proximity of collaborators. Knowledge production is a collaborative activity, not only between individuals within an organisation but also across organisations. My central assumption is that the selection of collaborators and the resulting outcomes of collaboration are not random, but are influenced by driving factors. In this thesis I focus on the factor of proximity: the degree to which potential or actual collaborators lie close to each other in a specific dimension. I elaborate on four such dimensions: geographical, organisational, social and cognitive proximity. The main question of the thesis is:

How is collaborative knowledge production in the water field influenced by the geographical, organisational, social and cognitive proximity of the actors involved?
I have unravelled this rather broad question into four specific sub-questions that each inquire about a specific aspect of the main question. To determine the factors that influence network formation and the choice of collaborators, I analyse how the European research network has developed over time. More specifically, I investigate what characteristics lead an actor to acquire a central position in the network. Focusing on the factor of proximity, I analyse its effect on the propensity of actors to collaborate, and its effect on the reported outcomes of collaboration. This leads to four sub-questions:

1. How has the configuration of the European water knowledge production network developed over time?
2. What explains the variation in the centrality of the actors in the network?
3. What is the influence of proximity on the propensity of actors to collaborate in knowledge production in an applied field of research, water?
4. What is the influence of proximity on the reported outcomes of collaborative knowledge production?

1.3 Conceptual framework

To answer the questions posed above, I will first present a conceptual framework. Four concepts are central in addressing the research questions formulated above: collaborative knowledge production, network configuration and centrality, research outcomes and, last but not least, proximity. Each of these concepts is grounded in and builds upon an existing body of literature. Below, I elaborate on each concept.

1.3.1 Collaborative knowledge production

The most central conceptual element is collaborative knowledge production. While the concept is similar to the more frequently used term 'research collaboration', it has a slightly broader meaning. I prefer to use the concept of 'collaborative knowledge production' because 'research collaboration' may have the connotation of scientific research alone and I am convinced that non-academic knowledge production is crucial to the advancement of research fields, and hence that a broad definition is needed in order to obtain a comprehensive understanding of proximity mechanisms (Aguiléra et al., 2012).

Research is increasingly considered to be an inherently collaborative process (Katz and Martin, 1997). Collaboration can manifest itself in many different ways, ranging from a division of labour or access to research equipment to transmission of know-how (Laudel, 2001; Katz and Martin, 1997). General developments in the research process, such as the increasing complexity of research problems and differentiation into specialised research fields, have further encouraged greater specialisation by actors and thus greater interdependency and collaboration (Melin, 2000). Researchers and other knowledge producers collaborate for various reasons, such as gaining access to resources (Melin, 2000), accumulating reputation, improving the efficiency of research, or for learning and other personal purposes (Beaver, 2001; Van Rijnsoever and Hessels, 2011). I have defined collaborative knowledge production as a situation in which two or more actors are involved in a process of advancing their knowledge of an identified subject or topic.

1.3.2 Network configuration and centrality

In this thesis I consider the totality of all collaborations in a dataset to be one knowledge production network. The concepts employed to analyse the characteristics of the network under consideration stem from the literature on social network analysis (Barabási et al., 2002; Newman, 2004). Such networks can be analysed at several levels of aggregation, from individuals to countries; I have chosen to analyse the network at the organisation level. Large research organisations are broken down further into coherent entities, such as faculties of a university. An important question at this actor level is thus what factors explain how well a specific actor is embedded in the network. This embeddedness is made operational with the notion of centrality. In this thesis, centrality is measured in such a way that it incorporates both the direct access of a participant to other

participants (weighted degree centrality) and the function of the actor as a link in the shortest connection between two other participants (eigenvector centrality). Previous studies have revealed that institutional characteristics of individual actors help explain their centrality in the network. Generally speaking, organisations in higher education and research tend to have a high degree of centrality in knowledge production networks, while SMEs usually have low centrality (Protogerou et al., 2010). The geographical position of an actor also has an effect on their centrality in a network (Foddi and Usai, 2013). However, the importance of such characteristics differs across research fields (Heller-Schuh et al., 2011).

1.3.3 Research outcomes

Regardless of what a collaboration precisely entails, it is generally supposed to result in outcomes, such as new insights, products or technologies. In my conceptual framework I focus on knowledge-intensive outcomes. To date, the literature has mainly addressed the explicit outcomes of collaboration. Explicit knowledge is clearly articulated and codified (such as a publication), while tacit knowledge is based on action, experience, or involvement in a specific context (Alavi and Leidner, 2001). Most studies are based on the analysis of co-authorships (Chessa et al., 2013, Hoekman et al., 2010) and/or co-patents (Ter Wal, 2009; Chessa et al., 2013). A more extensive overview of this literature is provided in Chapter 3. However, there are many other relevant outcomes of collaboration. Innovations are often not patented but protected in other ways or shared openly. Non-profit organisations in particular tend not to capture the knowledge produced in patents or publications but rather share it through non-scientific publications or personal communication. Attention to other outcomes besides patents and publications is particularly relevant in my empirical field of study, the water sector, because in this sector patenting is rarely used as a method to secure intellectual property, even among profit organisations, and many non-profit organisations are involved in collaborative knowledge processes (MinEZ, 2014). There is, however, very little empirical literature on the effect of proximity on less explicit outcomes of collaboration. This is an important gap, given the disputes in the literature on the importance of face-to-face meetings and mutual trust in the transmission of tacit knowledge (Weterings and Ponds, 2009; Aguiléra et al., 2012). In chapter 3 I distinguish six forms of collaboration outcomes and I demonstrate that the degree of proximity indeed matters for the reported outcomes of a collaboration, with different effects for explicit and tacit outcomes.

1.3.4 Proximity

Proximity is a crucial factor in shaping collaboration. Previous studies have shown that, although the proximity between an actor and a potential collaborator is probably not the key reason for collaboration, it does influence the propensity to establish an actual collaboration (Hoekman et al., 2010, Bouba-Olga et al., 2012). This is not only the case for proximity in the geographical sense but also in other dimensions, such as socially, organisationally and cognitively (Boschma, 2005;

Knoben and Oerlemans, 2006). All four of these dimensions share one conceptual rationale: proximity contributes to knowledge production and innovation because it reduces uncertainty and transaction costs and enhances and eases coordination among collaborators (Boschma, 2005).

In this thesis, the findings are triangulated across different types of data. The relevance of such triangulation is explained in the methods section of this chapter. The data type also determines to some extent how the dimensions of proximity can be operationalised. Depending on the specific research question and the inherent limitations of the data, I have selected dimensions of proximity for each chapter. Below I introduce the four dimensions analysed throughout the thesis. The specific operationalisations for each analysis will be introduced in more detail in the respective chapters.

Geographical proximity

Geographical proximity is the first dimension of proximity described in the literature, and it is also most similar to the literal sense of 'proximity'. I have defined it as the shortest possible physical distance between the locations of two actors ('as the crow flies'). Geographical proximity is considered to ease learning and innovation, as it facilitates informal communication and direct exchange of knowledge (Boschma, 2005). It has been suggested that it sometimes functions as a substitute or complement to other dimensions of proximity (Rallet and Torre, 1999; Boschma, 2005). Previous empirical research indeed shows that collaborations occur more frequently at smaller geographical distances (Hoekman et al., 2010; Bouba-Olga et al., 2012). While this finding has been corroborated at several spatial levels and in various research fields (e.g. Balland, 2012; Broekel and Boschma, 2012), spatial configurations do differ across research fields (Barber and Scherngell, 2013). Some literature suggests that the effect of geographical distance decreases over time, as modern communication technologies make it easier to communicate across long distances at much lower costs (Scherngell and Lata, 2013). However, Hoekman et al. (2010) have shown that the tendency to work only with geographically proximate partners has not decreased over time. The tendency to publish articles with collaborators from the same administrative region slightly decreased in Europe in the period 2000-2007, but if one counts the share of international collaborations rather than international papers, the tendency for domestic collaboration is constant over time (Leydesdorff and Wagner, 2008; Hoekman et al., 2010; Frenken et al., 2009).

The fact that geographical proximity increases the propensity to collaborate does not imply that proximate collaborations are more relevant or valuable for knowledge production. Using survey data at the national level, Weterings and Ponds (2009) showed that most collaborations are geographically proximate, yet the collaborations at greater distances are perceived as more valuable and they also more often entail knowledge exchange on technical questions.

Organisational proximity

Organisational proximity refers to the degree of similarity between two actors in their internal aims, incentives and routines (Boschma, 2005; Aguiléra et al., 2012). Organisational proximity is considered to promote collaboration because it can reduce the uncertainty and opportunism that may be involved in knowledge creation. For organisationally proximate partners it is easier to protect interests in, for example, intellectual property and to arrange the division of rewards for the knowledge generated (Boschma, 2005). The definitions and operationalisations of organisational proximity in the existing literature are much more diverse than those of geographical proximity, and the findings are less unequivocal. Broekel and Boschma (2012) found a positive relationship between organisational proximity and the creation of a knowledge network among organisations, but they found no relationship between organisational proximity and innovative performance. The effect on network formation is in line with the findings of Balland (2012). Cunningham and Werker (2012) found a more indirect effect, showing that academic parties that collaborate with other academic parties are cognitively less proximate than collaborations where one or all of the partners are from non-academic organisations.

Social proximity

Social proximity can be defined as the degree to which collaborations are embedded in social connections between actors, following for example from earlier collaboration (Aguiléra et al., 2012).

Such social embeddedness eases and supports collaborative knowledge production and exchange (Broekel and Boschma 2012). The empirical evidence on the effect of this dimension of proximity is mixed. Some studies show that social proximity is the strongest predictor of the proximity dimensions studied, with a positive effect (Autant-Bernard, 2007; Ter Wal, 2009). One multivariate study found a positive effect, also when controlling for geographical and cognitive proximity (Broekel and Boschma, 2012). However, another multivariate study found that direct collaborators with one partner are not more likely to collaborate with each other – another proxy for social proximity (Balland, 2012).

Cognitive proximity

Cognitive proximity refers to similarities in the knowledge base of collaborators. A smaller cognitive distance makes it easier to understand each other and process gathered information efficiently. At the same time, having some cognitive distance ensures access to new knowledge (Nooteboom, 1999). Here, as well, the empirical evidence regarding the effect on collaboration is not unequivocal. Based on patent data, Cantner and Meder (2007) found that higher cognitive proximity contributes to the probability of two actors collaborating. Ter Wal (2009), also based on patent data, found the same positive effect in a univariate model, but this changed into a negative effect when controlling for geographical and social

proximity in a multivariate model. Broekel and Boschma (2012) found that higher cognitive proximity correlated with lower innovative performance of the actors.

1.4 Positioning of the chapters

In relation to the concept of proximity and knowledge production, three streams of literature can be identified, each analysing a specific aspect of collaborative knowledge production (Aguiléra et al., 2012):

1. How proximity explains the choice of collaboration partners and network formation (e.g. Autant-Bernard et al., 2007; Balland 2012).
2. How proximity explains the processes of knowledge production and knowledge sharing in research collaborations (Boschma 2005; Weterings and Ponds, 2009).
3. How proximity to collaborators explains the innovative performance of the collaborating organisations (Nooteboom et al., 2007; Broekel and Boschma, 2012).

It follows from the research questions formulated above that this thesis makes contributions to the first and second streams of literature. Chapters 2 and 5 of this thesis contribute to the first stream, demonstrating how several dimensions of proximity influence the patterns of collaboration in knowledge production. Chapter 4 also relates to this first stream, but provides a novel perspective. Where most literature in this stream is based on the analysis of dyadic relationships, Chapter 4 broadens this scope to other levels of aggregation. It is based on the analysis of characteristics of individual actors, and of the resulting network as a whole. The patterns of collaboration are thus not analysed in terms of proximity, but rather in terms of the geographical and organisational backgrounds of actors, and how these influence their centrality in the network. Chapter 3 fits in the second stream, demonstrating how the dimensions of proximity in collaboration relate to the reported outcomes of the collaboration. In summary, I will analyse two relationships between three concepts: Chapters 2, 4 and 5 discuss the relationship between proximity and the selection of collaborators, while Chapter 3 assesses the relationship between proximity and outcomes of collaboration. Moreover, the chapters differ not only in the specific element of collaborative knowledge production that is analysed, but also in their spatial scope and the type of data used. Table 1 provides an overview.

Table 1 Specific characteristics of the chapters (data sources between brackets)

Dependent variable	Collaboration	Outcomes
Spatial scope		
National (the Netherlands)	Chapter 2 (publications)	Chapter 3 (survey data)
Transnational (Europe)	Chapter 4 (EU projects) Chapter 5 (EU projects)	

Rathenau Instituut

1.5 Water as an empirical field

My main research question deals with collaborative knowledge production in the water sector. As explained above, one aspect of the niche I have identified in the literature is the relative lack of studies on applied research fields with strategic relevance. My second criterion for selecting this research field concerned the diversity of actors, which facilitates the analysis of the different dimensions of proximity: organisational, geographical and cognitive. Below, I explain how the water sector fulfils these criteria.

First of all, the water field can be characterised as an applied field of research with high strategic relevance. An important aspect of water research – like many environmental sciences – is the need to adapt and contextualise knowledge to local circumstances, such as specific environmental conditions. The strategic character of water research follows from its large societal relevance. A deeper understanding of water and water management is required to address environmental challenges that can have great consequences on a planetary scale (Rockström et al., 2009). While the past century may have witnessed a massive leap towards universal water provision in the most developed countries, even there water provision remains a challenge for the future. This was also acknowledged in the recently launched research programme of the European Union, Horizon 2020. It addresses seven grand societal challenges, water playing a prominent role in two of them: 'Climate action, resource efficiency and raw materials', and 'Food security, sustainable agriculture, marine and maritime research and the bio-economy'.

Second, the water sector is an interesting case with respect to the analysis of proximity and collaborative knowledge production because it includes actors from a wide variety of organisational backgrounds. They range from universities to commercial consultants, and from NGOs to governmental organisations, and many of them are involved in knowledge production (Blankesteijn, 2011; Frijns, 2013). The involvement of governmental organisations is inherent to the large impact of water policies on public health and public space. This was well illustrated by the work of John Snow (1855), who proved that cholera was spread

through contaminated drinking water. This made him the founding father of modern epidemiology, and it also formed the starting point for policies to improve the quality of drinking water from public pumps. Nevertheless, the variety in organisational backgrounds has been a hurdle to collaborative knowledge production. Most of the OECD countries indicate that they face a 'policy gap', caused by a high degree of fragmentation of responsibilities in implementing water policy. Even within the group of governmental organisations alone there is often a wide variety of actors, crossing various geographical levels and sometimes policy areas. This leads to segmentation and complicates collaboration (OECD, 2011). Overcoming such segmentation is a persistent challenge even for a country such as the Netherlands, which is considered a global reference for water management (OECD, 2014b).

The perceived fragmentation has led to calls for a more integrative approach in water management. There has been a general paradigm shift from the government as the single and exclusive authority for managing natural resources to a multi-stakeholder approach where many different stakeholders, each with their own institutional backgrounds, all participate in and contribute to the management of the resources. For research and knowledge production, this translates into participatory approaches, where actors from different backgrounds develop new knowledge together (Pahl-Wostl et al., 2007; Frijns et al., 2013). Such integration has also been termed the 'socialisation of water management', as it turns an autonomous, inward-oriented sector into a more open sector, interacting with related fields, policy included (Van der Brugge, 2009). The shift to an integrated approach is an ongoing process and far from complete (Biswas, 2004). Actors are said to suffer from myopia with regard to technology and innovation, which is sometimes caused or reinforced by regulatory and policy frameworks (Thomas and Ford, 2005). The sector does not sufficiently overcome the distances between different fields of expertise and is sometimes said to have closed networks (Van der Brugge, 2009). In terms of proximity: organisational proximity is expected to have an effect on the propensity of actors to collaborate.

Third, the water sector is interesting because of its geographical configuration. To begin with, the sector is largely bound within national systems (EIP, 2014). Almost everywhere in the world, water management is organised within national territories; in the case of federated nations, often even at the level of the states. It is only within these national systems that water management is organised according to local environmental conditions such as watersheds or catchment areas (Van Ast, 2000; Thomas and Ford, 2005; EIP, 2014). The challenges in the water sector, however, are typically specific to local environmental conditions, but not bound by administrative and cultural borders. Even within nations, water management often has complex, multi-level and fragmented structures. The Netherlands for example has a multi-level governance system with a relatively high degree of decentralisation. Decision-making is a joint responsibility of the central government, authorities at the regional level (both provinces and water

boards), municipalities and other stakeholders (Brouwer, 2013). Such configurations form an interesting case for the analysis of geographical proximity in collaborative knowledge production.

Fourth, the water sector has a high cognitive diversity. Water management consists of many different specialisms: water quality management and water quantity management, different types of water use (urban, agricultural, environmental, recreational, etc.), surface water and ground water, and water supply and wastewater collection and treatment. Each requires a specific expertise and knowledge base (e.g. Van Ast, 2000; Biswas, 2004). Although much attention is being paid to the concept of integration in water management, there are still many issues to be resolved, both conceptually and in practice. At least to some extent this is a problem of cognitive distances; there are many experts involved and their knowledge bases have become more specialised and hence narrower over time (Biswas, 2004). In other words, cognitive proximity is an important factor in collaborative knowledge production in the water sector.

1.6 Policy relevance

Apart from contributing to the scientific literature on proximity and collaboration, this thesis also aims to provide evidence and guidance to policies for collaborative knowledge production in the water sector and beyond. Collaborative knowledge production and the dimensions of proximity analysed in this thesis feature very prominently in the research policies of the past decades, despite the fact that the word 'proximity' itself is rarely mentioned. To illustrate the relevance of this research to such policies, here I will sketch their recent history, with a focus on European and Dutch policies, as these form the direct background to the empirical analysis.

Traditionally, policies on knowledge production have been organised in national systems. However, over the last few decades, in Europe, in addition to the national systems, also a European system has developed. This is relevant to the main research question because European knowledge and research policies have always focussed on collaboration and in many cases they have paid particular attention to the geographical and organisational background of actors. The emergence of a European system began with the establishment of European knowledge institutions in the 1950s, such as the Joint Research Centre (JRC) and CERN, and it has become more important ever since (Trondal, 2002). In other words, science and research have been part of European integration since the beginning. The first two European Communities, the European Coal and Steel Community (1951) and the European Atomic Energy Community, Euratom (1957), both incorporated joint research facilities, thus providing incentives for collaboration with peers across national borders. Since about 1965, there have been ongoing discussions about the principles of a broader European research policy, particularly in research areas where a shared approach was considered necessary, with countries cooperating in establishing centres and institutes such as the

European Space Agency but also for example the European Centre for Medium-Range Weather Forecasts (Tindemans, 2009).

Somewhat later, another successful pillar of cooperation was established with the creation of the Framework Programmes. This started relatively modestly, along the lines of existing Euratom programmes and the ESPRIT programme for information technology. At first there were only a few partners from industry, mainly large firms with their own R&D departments and established relationships with universities and public research institutes. However, over time the programmes have come to involve increasing numbers of firms and research partners. This has also created strong pressure to increase budgets. Moreover, the number of objectives and research areas covered has dramatically increased (Caracostas and Muldur, 2009). The Framework Programmes have a non-territorial approach in the sense that the funding is granted through competitive calls with content-based criteria. However, with the launch of the most recent FP in 2014, Horizon 2020, measures have been announced to widen participation. This includes 'teaming', where excellent research institutions are coupled with low performing R&D regions to create or upgrade institutions in that region, and 'twinning', where an emerging institution is linked to internationally leading institutions in the same field elsewhere (EC, 2014a).

A relatively new European policy instrument with particular relevance to proximity and research collaboration is the European Innovation Partnership (EIP) initiative, part of the Innovation Union, which has the specific objective of bringing together actors across sectors and borders. It aims to overcome fragmentation and involve all actors in the innovation chain, from those conducting basic research to end users. It is not a funding instrument, but it is meant to drive the alignment of priorities, the leveraging of other instruments and the formation of new partnerships. The EIPs are challenge-driven, and the second EIP since the introduction of the instrument specifically concerns water (EC, 2014b).

Meanwhile, policies at the national level have also incorporated more incentives for collaboration. Since the late 1970s, innovation has become much more of a focal point in science and technology policy (Velzing, 2013), with the concept of 'innovation policies' also coined. The Netherlands was one of the first European countries to turn these ideas into broad innovation policies (MinWB, 1979), with many other countries following suit. The introduction of innovation policies also led to increasing interest in the question of how governmental organisations can stimulate other actors to participate in innovative processes (Kuhlmann, 2001; Caracostas and Muldur, 2001). The concept of 'enabling technologies' brought to the fore the idea that governments and other actors can and should support and facilitate innovation in industry and other sectors, and it has shown that new knowledge can pertain to many different sectors.

This in turn has also led to reforms in many institutions and organisations involved in producing knowledge. Universities have become more entrepreneurial, working in closer collaboration with firms. This is illustrated by the establishment of science parks and changes to funding models at universities, as well as new management models (Clark, 1998; Etzkowitz, 2003). In France, for example, this has led to a much more prominent place for SMEs in innovation processes, while reducing the role of large programmes centred around large firms, favoured by 'old' industrial policy. Interestingly, it has also led to the emergence of new regional concentrations of innovation and research, such as Grenoble and Sophia Antipolis (Tindemans, 2009; Ter Wal, 2009).

This focus on innovation policy has also created greater awareness of the wide range of policy instruments that affect innovative performance. These instruments not only include activities such as the promotion of knowledge production by firms, but also support for interaction and collaboration between different actors, environmental regulations and improvements in the functioning of capital markets for investment in knowledge production (Tindemans, 2009).

In the second half of the 1980s, Dutch innovation policy gradually evolved into technology policy. The focus was on stimulating R&D, with a shift to stimulating promising new technology areas rather than backing sectors with proven strengths (Velzing, 2013). In the late 1980s and the beginning of the 1990s, a focus on collaboration returned, and policy was oriented especially towards stimulating collaborative knowledge production and shared innovation by firms and research institutes working together. The latter were also stimulated to adopt entrepreneurial approaches. The new policies were inspired by the concepts of national and regional innovation systems (Lundvall et al., 2002) and clusters (Porter, 1998). An important common premise in these concepts is the relevance of local and regional collaboration in knowledge production, also in an era of globalisation. This resulted in two focal points in innovation policies: first, stimulating collaboration between actors from diverse organisational backgrounds (universities, firms, governments, etc.); and, second, a more systemic approach, with more attention being paid to the analysis of the logic of these actors and of conditions that may support and promote innovation (Smits & Kuhlmann, 2004).

In the Netherlands, such policies received a strong boost with the dedicated use of natural gas revenues to improve knowledge infrastructures. This included funding instruments to build consortia of organisations across different organisational backgrounds. A series of temporary intermediary organisations were created, with the specific aim to stimulate the development of such consortia and promote public-private partnerships (e.g. Hessels and Deuten, 2013). The most recent development is the introduction of the 'top sector' policies, with the focus remaining on stimulating collaboration and interaction between actors across diverse organisational backgrounds, notably governmental organisations, public research institutes, universities and firms. There is also increasing attention being

paid to the connection between national and European policies (Velzing, 2013; OECD, 2014a).

Although this sketch of research policy over the past decades is far from exhaustive, it shows very clearly that collaborative knowledge production and proximity have had a prominent position. It is thus all the more remarkable that there are large blind spots in our knowledge of the effect of proximity on collaborative knowledge production. For the Dutch context, it has been concluded more generally that research and innovation policy remains insufficiently based on evidence and the empirical evaluation of existing practices (Velzing, 2013). This thesis contributes to filling that gap as it provides additional evidence on the effect of proximity on collaborative knowledge production.

1.7 Methods, data and structure

The main research question concerns the influence of proximity on collaborative knowledge production. This influence may manifest itself in many different ways, depending on the form of collaborative knowledge production that is being examined. One of the contributions of this thesis to the existing literature on proximity is that it triangulates findings across different data sources that contain different forms of collaboration: survey data on both formal and informal collaborations between individuals, publication data on co-authorships, and data from the EU Framework Programmes on joint project participation. This is important because each data type has its own inherent limitations regarding the operationalisation of the proximity dimensions, and in some cases the methods of analysis. Moreover, as will be shown in Chapter 3, the effect of proximity differs across different types of outcomes of collaboration; most data sources measure only those collaborations that result in a specific type of outcome. The use of a variety of different data types also allows for a comprehensive overview of the role of proximity at different levels of aggregation (individuals and organisations).

Most existing studies on the role of proximity in research collaboration are based on patent or publication data. Data of these types have the advantage that they allow for the analysis of very large datasets that are more or less readily available. However, they do not provide a full picture of knowledge production or innovation. Another disadvantage of the use of patent and publication data without any other additional data sources is that many relevant partners involved in knowledge production do not tend to publish or patent their findings. In particular, not-for-profit organisations and governmental organisations are underrepresented in such datasets. This is all the more relevant because, as shown in Chapters 2 and 5, there is a difference in the effect of proximity for knowledge producers (such as universities and research institutes) and knowledge users (such as firms and governmental organisations).

The sub-questions on the configuration of the research network and the centrality of the actors are addressed using project participation data from projects in the

EU Framework Programmes (Chapter 3). The set contains 2963 projects funded in FP 1-7, with starting dates between 1985 and 2010 (the latest updates in the dataset are from March 2010). This dataset is analysed using methods from Social Network Analysis.

The sub-question concerning the influence of proximity on the patterns of collaborative knowledge production is addressed using both the FP dataset mentioned above and a set of publication data (Chapters 5 and 2, respectively). The FP dataset is analysed with multivariate logistic regression models. The publication dataset consists of 2247 publications published between 2006 and 2008, retrieved from the Web of Science. It is filtered for affiliations in the Netherlands. This data is analysed using a gravity model, which is tested in a negative binomial regression analysis.

The sub-question concerning the influence of proximity on the reported outcomes of collaboration is addressed using a dataset based on a survey among members of the Royal Dutch Water Network (Chapter 3). In total, 618 questionnaires were filled out and returned. Respondents were asked to provide information on their collaboration with three random alters. This yielded complete information on 1020 individual collaborations. This data was analysed with Mann-Whitney tests, odd ratios and a multivariate multinomial logistic regression model.

Actors can be analysed at different levels of aggregation: to gain a full understanding of collaboration in knowledge production, it is important to analyse the behaviour of both individuals and organisations. Chapter 3 looks at the individual level, Chapters 2, 4 and 5 at the level of organisations (and sub-organisations).

In addition to the aggregation level of the actors themselves, the collaboration patterns can also be analysed at various levels of aggregation:

- the level of the actor: how the characteristics of an actor are related to his collaborative behaviour (Ferru, 2010; Heller-Schuh et al., 2011);
- the dyadic level: how characteristics of the relationship between two actors influence their collaboration (Cunningham and Werker, 2012; Bouba-Olga et al., 2012);
- the system level: what factors explain the entire network that emerges as a sum of all dyadic relationships (Heller-Schuh et al., 2011; Chessa et al., 2013).

Chapter 4 of this thesis takes both the perspective of the actor and the system, while Chapters 2, 3 and 5 focus on the dyadic level. Where appropriate, I elaborate upon implications for the other levels.

Table 2 provides an overview of the spatial scale, the data type, the method of analysis, the level of aggregation of the analysis and the level of aggregation of

the actors of each study in this thesis. Chapter 2 analyses the role of geographical and organisational proximity in the collaboration patterns that appear in co-authorships of scientific publications in the Netherlands. Chapter 3 examines the impact of geographical, social, organisational and cognitive proximity on the knowledge outputs of collaboration based on a survey among professionals in the Dutch water sector. Chapter 4 describes the configuration of the knowledge production network on water in Europe and analyses characteristics that influence the centrality of actors in that network. Chapter 5 studies the role of proximity in collaborations that result from EU funding. Finally, Chapter 6 brings together the conclusions of all of the empirical chapters and discusses the implications for future research and policy in more detail.

Table 2 Overview of spatial scale, data type, method of analysis and aggregation level of the empirical chapters.

	Chapter 2	Chapter 3	Chapter 4	Chapter 5
Spatial scale	National (Netherlands)	National (Netherlands)	Transnational (Europe)	Transnational (Europe)
Type of data	Publication data (Web of Science)	Survey data from own questionnaire	Participation data Framework Programmes	Participation data Framework Programmes
Method of analysis	Negative binomial regression	Correlation tests (Mann-Whitney) Multivariate logistic regression	Social network analysis (centrality measures)	Multivariate and bivariate logistic regression
Level of aggregation of analysis	Dyadic	Dyadic	Actor and system	Dyadic
Level of aggregation of actors	Organisations	Individuals	Organisations and sub-organisations	Organisations and sub-organisations

Rathenau Instituut

1.8 References

Aguiléra, A., V. Lethiais & A. Rallet (2012). Spatial and Non-spatial Proximities in Inter-firm Relations: An Empirical Analysis. *Industry and Innovation*, 19(3), 187-202.

Alavi, M. & D. E. Leidner (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, 107-136.

Balland, P.A. (2012). Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies*, 46(6), 741-756.

Barabási, A. L., H. Jeong, Z. Néda, E. Ravasz, A. Schubert & T. Vicsek (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical mechanics and its applications*, 311(3), 590-614.

Barber, M. J. & T. Scherngell (2013). Is the European R&D network homogeneous? Distinguishing relevant network communities using graph theoretic and spatial interaction modelling approaches. *Regional Studies*, 47(8), 1283-1298.

Beaver, D. D. (2001). Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, 52(3), 365-377.

Biswas, A. K. (2004). Integrated water resources management: a reassessment: a water forum contribution. *Water international*, 29(2), 248-256.

Blankesteijn, M. L. (2011). *Tussen wetten en weten: de rol van kennis in water-beheer in transitie*. Universiteit van Amsterdam, Amsterdam.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional Studies*, 39(1), 61-74.

Bouba-Olga, O., M. Ferru & D. Pépin (2012). Exploring spatial features of science-industry partnerships: A study on French data. *Papers in Regional Science*, 91(2), 355-375.

Broekel, T. & R. Boschma (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, 12(2), 409-433.

Brouwer, S. (2013). *Policy entrepreneurs and strategies for change: The Case of Water Management in the Netherlands*. Vrije Universiteit, Amsterdam.

Brugge, R. van de (2009). *Transition dynamics in social-ecological systems: the case of Dutch water management*: Erasmus University Rotterdam.

Cantner, U. & A. Meder (2007). Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination*, 2(1), 45-65.

Caracostas, P. & U. Muldur (2001). The Emergence of a new European Union Research and Innovation Policy. In P. Laredo & P. Mustar (Eds.), *Research and Innovation Policies in the New Global Economy*. Cheltenham, UK: Edward Elgar Publishing.

Chessa, A., A. Morescalchi, F. Pammolli, O. Penner, A. Petersen & M. Riccaboni (2013). Is Europe Evolving Toward an Integrated Research Area? *Science*, 339(6120), 650-651.

Clark, B. R. (1998). *Creating Entrepreneurial Universities: Organizational Pathways of Transformation. Issues in Higher Education*: ERIC.

Clark, P. (2014, July 14, 2014). A world without water. *Financial Times*. Retrieved from http://www.ft.com/intl/cms/s/2/8e42bdc8-0838-11e4-9afc- 00144 feab7de.html#slide0

Cunningham, S. W. & C. Werker (2012). Proximity and collaboration in European nanotechnology. *Papers in Regional Science*, 91(4), 723-742.

Delanghe, H., U. Muldur & L. Soete (2009). *European science and technology policy: Towards integration or fragmentation?*: Edward Elgar Publishing.

EC. (2014a). *Enabling synergies between European Structural and Investment Funds, Horizon 2020 and other research, innovation and competitiveness-related Union programmes: Guidance for policy-makers and implementing bodies*. Brussels: European Commission.

EC. (2014b). *Outriders for European Competitiveness: European Innovation Partnerships (EIPs) as a Tool for Systemic Change*. Brussels: European Commission.

EIP. (2014). *Barriers and bottlenecks for innovation in the water sector.* Brussels: European Innovation Partnership on Water.

Este, P. d' & S. Iammarino (2010). The spatial profile of university-business research partnerships. *Papers in Regional Science,* 89(2), 335-350.

Etzkowitz, H. (2003). Research groups as 'quasi-firms': the invention of the entrepreneurial university. *Research Policy,* 32(1), 109-121.

Etzkowitz, H. & L. Leydesdorff (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy,* 29(2), 109-123.

Ferru, M. (2010). Formation process and geography of science–industry partnerships: the case of the University of Poitiers. *Industry and Innovation,* 17(6), 531-549.

Foddi, M. & S. Usai (2013). Regional knowledge performance in Europe. *Growth and Change,* 44(2), 258-286.

Frenken, K., S. Hardeman & J. Hoekman (2009). Spatial scientometrics: Towards a cumulative research program. *Journal of Informetrics,* 3(3), 222-232.

Frijns, J., C. Büscher, A. Segrave & M. van der Zouwen (2013). Dealing with future challenges: a social learning alliance in the Dutch water sector. *Water Policy,* 15(2), 212-222.

Hagedoorn, J., A. N. Link & N. S. Vonortas (2000). Research partnerships. *Research Policy, 29*(4), 567-586.

Hagstrom, W. (1965). *The Scientific Community.* New York: Basic Books.

Heller-Schuh, B., M. Barber, L. Henriques et al. (2011). JRC scientific and technical reports: 'Analysis of Networks in European Framework Programmes' (1984-2006).

Hessels, L. K. & J. J. Deuten (2013). *Coördinatie van publiek-privaat onderzoek: van variëteit naar maatwerk.* SciSA rapport nr 1328. Den Haag: Rathenau Instituut.

Hessels, L. K. & H. van Lente (2008). Re-thinking new knowledge production: A literature review and a research agenda. *Research Policy,* 37(4), 740-760.

Hoekman, J., K. Frenken & R. J. W. Tijssen (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy,* 39(5), 662-673.

Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics,* 31(1), 31-43.

Katz, J. S. & B. R. Martin (1997). What is research collaboration? *Research Policy,* 26(1), 1-18.

Knoben, J. & L. Oerlemans (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews,* 8(2), 71-89.

Kuhlmann, S. (2001). Future governance of innovation policy in Europe—three scenarios. *Research Policy,* 30(6), 953-976.

Laudel, G. (2001). Collaboration, creativity and rewards: why and how scientists collaborate. I*nternational Journal of Technology Management,* 22(7), 762-781.

Leydesdorff, L. & C.S. Wagner (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics,* 2(4), 317-325.

Lundvall, B.Å., B. Johnson, E. S. Andersen & B. Dalum (2002). National systems of production, innovation and competence building. *Research Policy,* 31(2), 213-231.

Melin, G. (2000). Pragmatism and self-organization: research collaboration on the individual level. *Research Policy,* 29(1), 31-40.

MinEZ. (2014). *Implementatie van de visie op het toegepast onderzoek*. Den Haag: Ministerie van Economische Zaken.

MinWB. (1979). *Innovatienota- Het overheidsbeleid inzake technologische vernieuwing in de Nederlandse samenleving*. Den Haag: Tweede Kamer, zitting 1979-1980, 15855 no 1.

Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America,* 101(Suppl 1), 5200-5205.

Nieminen, M. & E. Kaukonen (2001). Universities and R&D networking in a knowledge-based economy. A glance at Finnish development. *Sitra Reports series,* 11.

2 The effect of proximity on research collaboration in a small country[1]

2.1 Introduction

The phenomenon of collaboration in scientific research has long attracted considerable attention. Since the early work of Smith (1958) and De Solla Price and Beaver (1966), there has emerged a vast community of scholars interested in measuring collaboration, discerning underlying patterns, and finding drivers of collaboration. Collaboration is associated with various benefits, such as higher citation impact (Frenken, Ponds & Van Oort, 2010), knowledge transfer, equipment sharing, and network formation (Gazni et al., 2012). As a result, it has become commonplace for research policy to encourage collaboration among researchers and institutions (Katz & Martin 1997; Melin, 2000).

Research collaboration has a strong spatial component. It is known that people are more likely to collaborate with geographically proximate partners, both at the micro level of a single building (Allen 1977) and at the macro level of very large countries, like the USA, or entire continents (Katz 1994; Hoekman et al. 2010). In addition, knowledge production and innovative activity are geographically clustered (Malecki 2010). These insights have inspired policies that actively encourage research co-location, for example in science parks, to promote knowledge exchange and spill-overs and to share large facilities. On the other hand, the EU actively promotes collaboration across long geographical distances, on the premise that a larger radius improves the chance of finding relevant collaborators for shared knowledge production.

For both types of policy it is crucial to understand what happens at different spatial levels. We know comparatively little about collaboration at the national level, even though science and innovation policy is primarily a national concern, even in the EU; a large proportion of collaborations in Europe takes place within countries (Hoekman et al., 2010; Chessa et al., 2013); and national borders draw hard boundaries around regulatory environments and markets for (semi)public goods. What we do know about proximity at the national level is mostly based on large Anglo-Saxon countries, particularly the USA and UK (e.g. Smith and Katz 2000; d'Este and Iammarino 2010). The functioning of proximity may be different in small countries where every location can be reached within a few hours. When geographic boundaries to face-to-face communication are low, co-location may be unnecessary and spatial patterns of collaboration may not reflect the effect of distance. How does geographic proximity work in a small country?

---

1 This chapter has been submitted - in slightly different form - to *Tijdschrift voor Economische en Sociale Geografie (TESG)*

In this chapter, we examine patterns of research collaboration in the Netherlands, a small country. Our focus is on collaboration in water-related research. Large parts of the water sector are organised along regional lines (e.g. water boards, water distribution areas). In addition, water-related research tends to involve applied research adapted to local conditions.

This article is structured as follows. In section 2.2 we give an overview of the literature on drivers of collaboration, with specific attention to the role of geographical proximity. In section 2.3 we describe the methodology used to construct a dataset on collaboration and statistically test the effect of distance. In section 2.4 we present our results. In section 2.5 we give our conclusions and discuss the implications for future research.

2.2 Current literature

Knowledge production is increasingly a collaborative activity (Katz & Martin 1997). Knowledge producers collaborate for a variety of reasons, including access to resources (Melin 2000), accumulating reputation , efficiency and effectiveness of research, learning, and personal purposes (Beaver 2001; Van Rijnsoever & Hessels 2011). Collaborative behaviour is shaped by conditions in the science system, such as funding patterns (Bozeman & Corley 2004) and the increasing specialisation and professionalisation of science (Beaver, 2001). Such conditions have a direct effect on the motivations for collaboration, for example by changing the availability of resources.

Distance has an indirect effect on collaborative behaviour. Motivations to collaborate remain the same, but the probability of actual collaboration diminishes with distance. For example, Bozeman and Corley (2004) observe that most researchers tend to work with people in their own research group rather than with distant collaborators. Measures to reduce distance include geographic co-location, for example in a science park, building research infrastructures, and lowering social, cultural, linguistic or political barriers. (Acedo et al. 2006; Katz 1994; Katz & Martin 1997).

2.2.1 The impact of geographical proximity

Gaining a better understanding of the role of physical distance in collaborative knowledge production is important. First, collaboration has many benefits. Some of these benefits may be larger for collaboration across longer distances. The odds of finding partners with a supplementary knowledge base with whom new knowledge can be produced, increases with the radius of search. Second, significant investments have been made to stimulate long-distance collaboration. One of the aims of the Framework Programmes of the European Union is to encourage collaboration across Member States. Simultaneously, other investments promote co-location, for example in science parks and clusters of innovation. Co-location is associated with collaboration and knowledge spill-overs (Breschi and Lissoni 2003). More insight in the relevance and effects of distance can improve the rationale behind such investments (Hoekman et al. 2010).

Being geographically proximate promotes collaboration. Most collaborations are initiated in informal settings, and geographical proximity facilitates such settings. Geographical proximity lowers coordination and transaction costs in collaboration (e.g. Katz 1994; Boschma 2005). However, too much geographical proximity may hinder the processes of knowledge production. If actors in a region become inward-looking, the result may be geographical lock-in, weakening the learning capacities of the actors (Boschma 2005).

Various researchers have conducted empirical studies on the role of geographical proximity, with mixed results. An overview of recent studies on science-industry collaborations can be found in Bouba Olga et al. (2012). The effect of geographical proximity is analysed at many different spatial levels, from research collaboration within a building (Hagstrom 1965) to collaboration across the globe (Matthiessen et al. 2010). The earliest studies were done at the microlevel. Hagstrom (1965), Allen (1977), and Kraut and Egido (1988) found that the probability of communicating between potential collaborators declines sharply as distance increases. This also holds for researchers who collaborate already, and the effect remains if it is controlled for similarities in the organisational background and thematic specialisation of collaborators. Research collaboration between countries is affected by linguistic, historical and cultural factors (Narin et al. 1991), and by the distance between countries, more than by thematic and socio-economic similarities (Andersson & Persson 1993; Nagpaul 2003). In a study of collaboration among the hundred largest cities in the world, Matthiessen et al. (2010) found that this is influenced by geographical proximity. Collaboration patterns between regions in Europe are influenced both by physical distance and national borders (Hoekman et al. 2010).

Studying collaboration within nations, Katz (1994) finds that in inter-university collaboration the frequency of collaboration between domestic universities declines exponentially with distance between the partners. However, in a sample of articles in economic top journals where authors had at least one affiliation in the US, Sutter and Kocher (2004) find that none of their geographical variables – spatial distance; being in the same state or in an adjacent state – is significant in explaining the collaboration pattern.

A popular hypothesis in the literature on the relevance of geographical distance is that the importance of distance decreases over time, as modern infrastructures enable researchers to overcome the barrier caused by distance (e.g. Merino & Rubalcaba 2012). Empirical studies suggest otherwise. In the UK, the average distance between collaborators decreased in the life sciences between 1981-1983 and 1992-1994 , but it remained more or less stable over time in the natural sciences, engineering, and multidisciplinary research (Smith & Katz 2000). Havemann et al. (2006) found no effect of geographical distance in a sample of German immunological institutes, if controlled for collaborations the same town, and this did not change in the time span 1992-2002.

Geographical proximity works differently in different areas (Aldieri 2011; Smith & Katz 2000; Weterings & Ponds 2009). For example, it has been shown to have a much stronger effect in engineering than in basic science (d'Este & Iammarino, 2010). Geographical proximity may also vary by spatial scale. Thus far, studies have either examined the supranational level (for example a continent) or the very local level (one organisation or building). What happens at the intermediate level – in a region or nation – remains underresearched. The national level is highly relevant. Despite long-standing efforts to build a European Research Area, science in Europe is still largely organised in national systems (Hoekman et al. 2010; Chessa et al. 2013). Studies that analyse a country usually take a relatively large one, such as Germany (Havemann et al. 2006); the UK (Katz 1994; Smith & Katz 2000; d'Este & Iammarino 2010); Canada (Katz 1994); Australia (Katz 1994); or the USA (Sutter & Kocher 2004). It is important to study small countries as well, because the research collaboration patterns of small countries is different from that of large countries, both for collaboration within and outside the national borders (Ukrainski et al., 2014).

2.2.2 Other dimensions of proximity

Originally, the focus of the literature was on geographical proximity. The idea has later been extended to other dimensions, such as social, cognitive and organisational proximity (Boschma 2005; Knoben & Oerlemans 2006). An early empirical example is the study of Egido and Kraut (1988) who controlled for organisational and thematic similarities to analyse the effect of geographical proximity. Some argue that these other dimensions may be more important than geographical proximity (Breschi & Lissoni 2003; Sternberg 2007).

2.2.3 Study design

We hypothesise that there is a difference between the functioning of geographic proximity in a small country and a large country or supranational region. To empirically test that hypothesis we study collaboration in water-related scientific research in the Netherlands.

The Netherlands is a small European country where distances between cities are relatively short compared to the US or Germany. This may influence people's perception of distance and of the effort required to overcome the barrier of distance. In addition, high-tech economic activity is not spread across the country according to labour market characteristics or localised economies of agglomeration but by the availability of research institutes (Van de Panne & Dolfsma 2003). The Netherlands is also the subject of another study on proximity in a small country by Ponds et al. (2007), who show that geographical proximity indeed plays a role in science-based industries. Water research is an applied research field where most knowledge is contextualised for local conditions (e.g. geology). This may imply that actors have an additional incentive to search proximate collaborators. On the other hand, they may search for someone working in similar conditions rather than for someone who is geographically

proximate. Moreover, water research involves actors from many different organisational backgrounds (Thomas & Ford 2005; EIP 2014). This allows us to examine the role of organisational proximity – defined for example in terms of incentives and routines (Boschma 2005; Aguiléra et al. 2012) – in addition to geographical proximity. Organisational similarity is said to facilitate collaboration by lowering uncertainty (Boschma, 2005). Differences in organisational background may be more important in applied research than in basic research (d'Este & Iammarino, 2010) and more so in industrial R&D than in public research (Scherngell & Barber, 2009). Empirical findings on the effect of organisational proximity vary. Balland (2012) finds a direct and positive effect, while Broekel and Boschma (2012) find no effect, and Ponds et al. (2007) find an indirect effect in that geographical proximity is stronger if organisational proximity is lower.

2.3 Methodology

Research collaboration involves the working together of researchers to achieve a common goal. This definition gives no indication of how closely researchers should work together for it to be considered collaboration. It is not easy for an external person to assess who should be counted as collaborators (Katz & Martin 1997). A common proxy is to use the names of co-authors on scientific papers that contain the results of research. We follow Melin and Persson (1996) in stating that the assumption that significant research collaboration will generally lead to co-authoring is realistic, as contributors will want to claim priority.

One of the advantages of using co-authorships as a proxy for collaboration is that it is comparatively easy to construct datasets large enough for quantitative analysis. However, it is not unproblematic to study collaboration based on co-authorship data. Not all forms of collaboration result in joint papers, while the mere fact that several people are listed as authors does not imply that they actually collaborated during the research phase (Laudel 2002). There have been a few attempts to quantify the extent of these limitations. In a small-scale study at Umeå University, Melin and Persson (1996) found that less than 5% of the authors indicated having experienced situations in which collaborative work did not result in co-authored articles. Where it did not, the main reason was that the contribution was considered too minor. Laudel (2002) finds that whether or not a collaboration results in co-authored articles depends on what the collaboration entails. Almost all collaborations based on division of labour resulted in co-authorship; in the exceptional cases where it did not this was most likely because the collaborative work failed to produce publishable results. All other forms of collaboration (such as shared access to research equipment) were rarely rewarded with co-authorship. It is hence likely that our dataset mainly contains information on (successful) collaborations involving a division of labour.

When analysing collaboration at an organisational level double affiliations present an additional problem. Scientists are often affiliated to more than one institute. It is questionable whether or not this should be counted as collaboration

from a conceptual perspective. From a practical perspective it is often impossible to avoid including dual affiliations. Until recently, the Web of Science did not link authors to their affiliations, but instead provided separate lists of authors and research addresses per article. Authors from different countries may all have an affiliation with one specific institute, and list only that institute when writing an article together. The reverse may also happen if one researcher has more affiliations in distant places (see also Katz & Martin 1997; Wagner & Leydesdorff 2005). These are important limitations to keep in mind when assessing the role of geographical distance in collaboration patterns.

2.3.1 Level of analysis

Co-authorships are primarily organised at the individual level. However, it makes sense to analyse patterns of co-authorship at higher levels of aggregation as well, for example at the level of research groups, departments, institutions, regions or countries. Most policies regarding research collaboration aim at such higher levels of aggregation (Katz & Martin 1997). Given that it is impossible to retrieve individual affiliation data and that our main focus is on the relevance of geographical distance, we conduct the analysis at the organisational level.

2.3.2 Data retrieval

In principle, data on co-authorship can be retrieved from any extensive bibliographic database. We used Thomson Reuters Web of Science (WoS), which is believed by many to be the most reliable source for a comprehensive survey of co-authored publications (e.g. Wagner & Leydesdorff 2005; Melin & Persson 1996). By including the WoS conference proceedings indexes, we improve coverage of the technical sciences. We have not distinguished between types of output (journal articles, letters, reviews, proceedings) as we are interested in connections between people and not in the scientific status of those connections (see Wagner & Leydesdorff 2005).

A topic search was carried out using the search terms “drinking water”, “water treat*” and “desalinat*” for the period 1969-2008. Experts were consulted to validate the initial dataset and adjust the search terms. Five journals published the largest part of these publications: Desalination, Water Research, Environmental Science & Technology, Water Science and Technology, and the Journal of the American Water Works Association. All articles published in these journals were downloaded. The keywords mentioned in the articles were used to develop a more refined set of keywords for topic search. Based on this topic search the final set of publications was generated. To keep the amount of data within workable limits, only the publications from 2006 to 2008 were used. From the final dataset, we extracted only those publications that contain affiliations from the Netherlands. All publications that involve international collaboration have been excluded to ensure that we measure only the effects of geographic proximity within national boundaries. The result is a set of 2,227 publications from 307 organisations, representing 646 co-authorship links between organisations from the Netherlands.

The number of publications per organisation has been used as a proxy for the size of the organisation's production of water-related knowledge.

All publications in the dataset contain details on the affiliations of the authors. The organisations in the affiliation addresses have been given a unique name and reference code that harmonises the address information provided by the WoS. The institutional affiliations have been accepted at face value. No attempt has been made to exactly reconstruct the organisational structure of universities and research institutes. Organisations were classified into seven types: universities, medical research centres, (semi)public research organisations, consultancies, firms, governmental bodies or other.

2.3.3 Distance matrix

Addresses were used to determine the town where the organisation is located. Since the analysis is carried out at organisation level, publications belonging to organisations that have locations in several places have been assigned a single location, namely the town that most frequently occurred in its articles. The 307 institutes are located in 97 towns throughout the Netherlands. These locations have been georeferenced and projected on a map (Figure 1). Latitudes and longitudes were used to calculate great circle distances between every combination of locations and produce a distance matrix.

2.3.4 Counts

This distance matrix is linked to a matrix containing all possible combinations of co-authorships between organisations. To analyse the effect of proximity, we need to measure the number of pairings an organisation has with other organisations. We count two-way collaborations. For example, if a paper lists four affiliations A, B, C and D, it involves six two-way collaborations A-B, A-C, A-D, B-C, B-D and C-D, each with a value of one, regardless of the number of authors belonging to any of the four organisations (for further elaborations on this counting technique, see Katz 1994). We use integer counting: if authors from institute A and authors from institute B write four articles together, then the link A-B has a value of 4. It is important to include pairs with zero co-authored publications. Excluding zero-flows from the analysis implies a loss of information on absent interactions. The explanatory variables may help explain why some organisations do not co-author at all (Eichengreen & Irwin 1998; Havemann et al. 2006).

2.3.5 Model

We use a gravity model to analyse the effect of proximity on research collaboration. Analogous to Newton's gravitational law, the gravitational force between two entities can be explained by the mass of these entities and the distance between them. The gravity model forms the core of a large body of literature on spatial interaction models (Murray 2010). It is also well suited to analyse spatial patterns in research collaboration (Beckman 1994). There are three reasons for using a gravity model to estimate the effect of physical distance. First, including

the size (mass) of the collaborating organisations in the model makes it much more realistic: organisations that produce more publications will naturally have a higher number of co-authorships. Second, the multiplicative nature of a gravity model has proven to provide a better fit to empirical data than additive (linear) models for many different phenomena. Third, the gravity model can easily be enriched by adding other relevant variables (Sutter & Kocher 2004).
The model is specified as follows:

(1) $co_{ij} = \beta_0 d_{ij}^{\beta_1} m_i^{\beta_2} m_j^{\beta_3}$

where co_{ij} is the number of co-authorships between institute i and j, d_{ij} is the geographical distance between the organisations, m_i and m_j represent the total water-related scientific output (mass) of organisations i and j. The unknown parameters to be estimated are β_0 to β_3.

It used to be very common to use a log-additive version of the model for analytical convenience (Sen & Smith 1995). Estimating the log-additive version using ordinary least squares (OLS) regression is not straightforward. The underlying reason is Jensen's inequality: E (ln (y)) ≠ ln (E (y)) (Santos Silva & Tenreyro 2006). Regression of a log-linear model will produce estimates of the logarithm of the error term μ_{ij}, not of μ_{ij} itself; the antilogarithms of these estimates are biased estimates of μ_{ij}. Ignoring that problem leads to systematic under-prediction of large values of the dependent variable (Flowerdew & Aitkin 1982). A second problem concerns the distribution of the error term, which implies that the values of co_{ij} are log-normally distributed around the estimate. However, co_{ij} measures binary co-authorships, which are nonnegative integers, hence their distribution will not be log-normal.

Last but not least, the logarithm of zero is not defined. Deleting all observations with zero co-publications leaves out important information on absence of collaboration. This produces biased results, particularly if the zero-valued observations are non-randomly distributed, which as Figure 3 shows is the case in our dataset (Eichengreen & Irwin 1998; Burger et al. 2009). This problem is often circumvented by adding a small positive number to all observations. However, in our dataset many observations have a value of zero and the exact value of the added constant has a considerable impact on the coefficients and on the explanatory power of the model. It can even be shown that any desired parameter estimate can be generated by adapting the value of the added constant (Flowerdew & Aitkin 1982; King 1988).

These problems can be overcome by assuming a different distribution. Recall that each observation of co_{ij} is a non-negative integer, and hence co_{ij} can be considered as having a discrete probability distribution. If there is a (small) constant probability P_{ij} that organisation *i* and *j* co-author a publication (and if co-authorships can be assumed to be independent of each other), then the

number of co-publications of *i* and *j* follows a Poisson distribution. The probability that *i* and *j* have exactly *k* co-publications is

$$(2)\quad P(co_{ij} = k) = \frac{e^{-\lambda_{ij}}\lambda_{ij}^{k}}{k!}$$

where λ_{ij} denotes the mean of the distribution (for more detail, see Flowerdew and Aitkin 1982).

One of the basic assumptions of the Poisson distribution is that the variance of the dependent variable is equal to its mean. Our dataset is characterised by overdispersion: the variance (.159) is much larger than the mean (.027). There is a distribution in the Poisson family – negative binomial distribution –that allows for variance higher than the mean. The expected value of co_{ij} will remain the same as in a Poisson model, but the variance has one more free parameter: it is a function of the conditional mean λ_{ij} and a dispersion parameter α. The dispersion parameter can model between-subject heterogeneity and solve the problem of overdispersion (Burger et al. 2009).

The probability mass function of a negative binomial model is:

$$(3)\quad P(co_{ij} = k) = \frac{\Gamma\,(co_{ij}+\alpha^{-1})}{co_{ij}!\,\Gamma\,(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\lambda_{ij}}\right)^{\alpha^{-1}}\left(\frac{\lambda_{ij}}{\alpha^{-1}+\lambda_{ij}}\right)^{co_{ij}}$$

where co_{ij} again is the number of coauthorships, Γ denotes the gamma function, α is the dispersion parameter, and λ denotes the conditional mean. This estimator is also known as a negative binomial pseudo-maximum likelihood model (NBPML).

2.4 Results and analysis

2.4.1 Mapping co-authorships

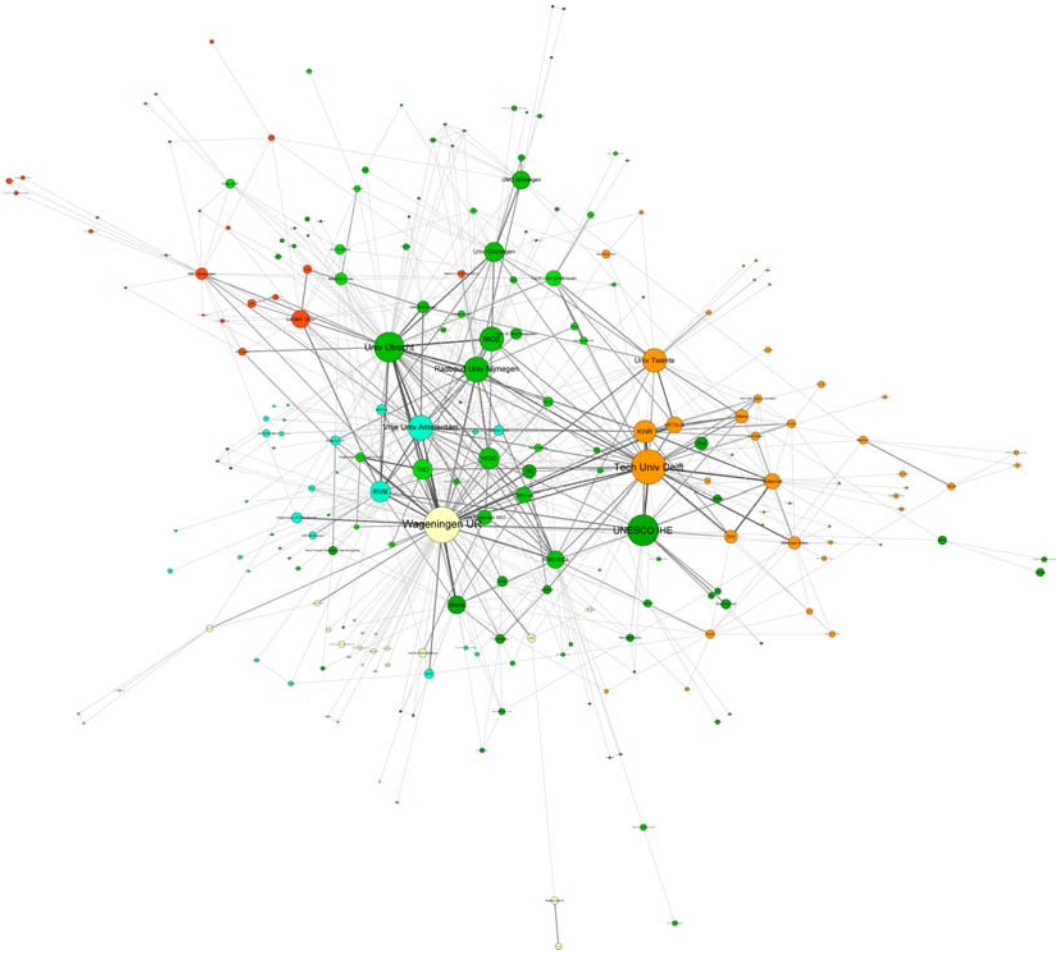
The dataset contains 97 different towns and cities (Figure 1).
Figures 2 and 3 show the co-author network in two different layouts. Figure 2 presents the network in a conventional manner, using colours to demarcate clusters in the network. In Figure 3 collaborating institutes have been mapped in their geographic location. The resulting image already indicates that most co-authorships occur among organisations in a relatively small part of the country. There is a belt of dense collaborations from west to east in the middle of the country. Even though the organisation of the water sector has a strong regional component (think, for example, of the distribution areas of drinking water companies or regional water authorities), there are no clear regional clusters in water-related research collaboration.

Figure 1 The locations of the organisations in our dataset; each dot depicts a city where one or more organisations are located.

Created by GPSVisualizer.com

Rathenau Instituut

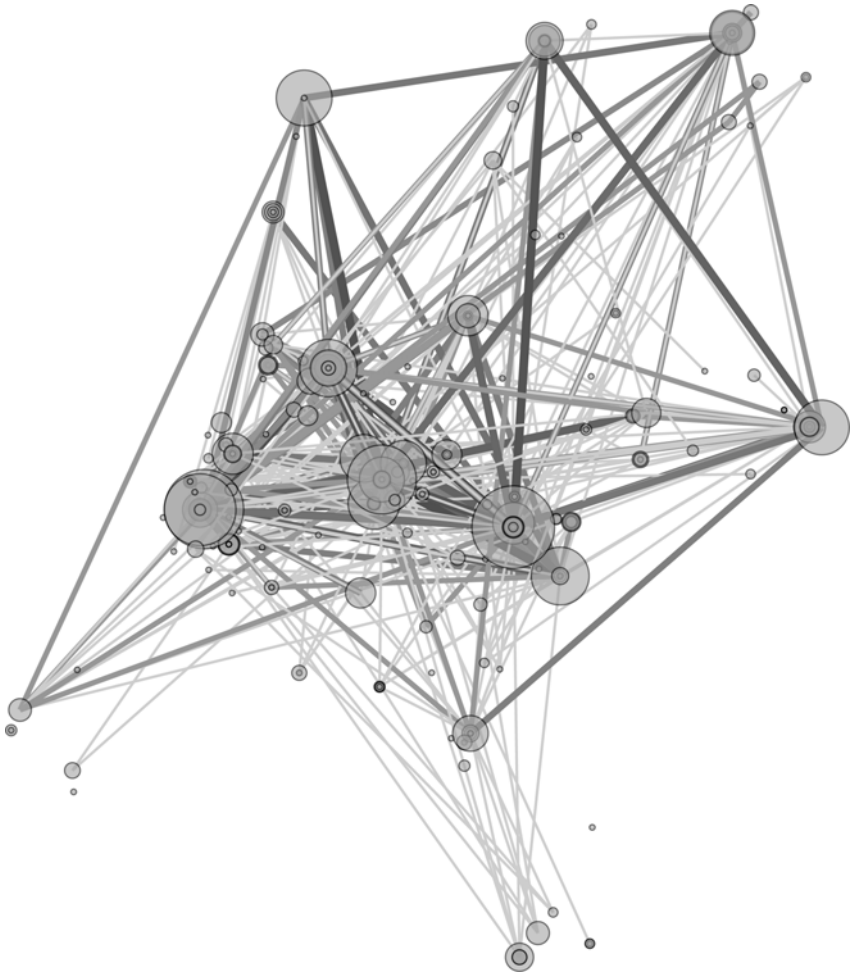
Figure 2 Co-author network in water-related scientific research in the Netherlands[2].



Rathenau Instituut

2 The size of nodes and font size of names are scaled to the number of publications produced by each organisation. The weights of edges represent the number of co-authorships. Colours indicate clusters in the network, determined using the community detection algorithm of Blondel et al. (2008).

Figure 3 Map of the co-authorship patterns in our dataset; the size of the nodes represents the number of publications of the organisation, the thickness of the edges represents number of co-authorships.



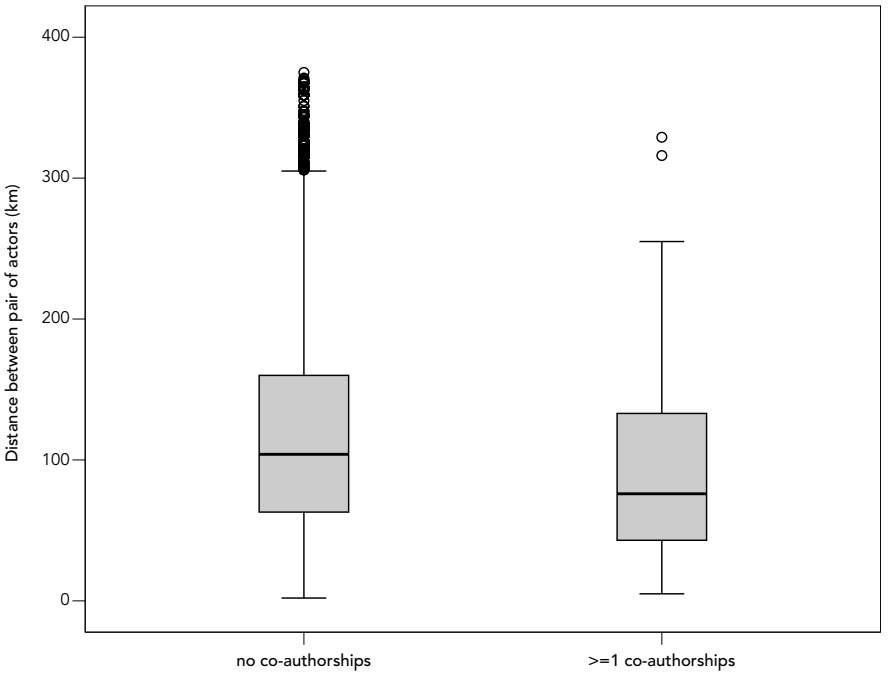
Rathenau Instituut

2.4.2 Comparing distance with and without collaboration

First, we compare the distance between organisations that have one or more co-authorships (n=646) with the distance between all pairs of organisations that have no co-authorships (n=46,035). The null hypothesis is that distance does not play a role in developing co-authorship relations, which means that average distances in the two groups should be more or less equal. Since the data are not normally distributed, we use a non-parametric test (Mann-Whitney) to test this hypothesis. The null hypothesis is rejected. There is a significant difference

between the two groups. Median distance in the group without co-authorships is 105.0 kilometers, while median distance in the group with co-authorships is 76.5 kilometers (Z= -10.052, p = .000). Figure 4 compares the distribution of distances in the two groups. This result confirms that geographic proximity has an effect: more proximate organisations are more likely to collaborate.

Figure 4 Boxplots of the distances between organisations with and without co-authorships respectively



2.4.3 Gravity model

We use a gravity model to expand our analysis, particularly to take into account the mass of collaborating organisations. Table 3 presents the results of the basic gravity model. The model has three explanatory variables: the mass of organisation A (defined as the total number of publications of that organisation in the dataset), the mass of organisation B, and the geographic distance between A and B. All parameters are highly significant (p<.001), which may be partly because the large number of observations.

Once again, the null hypothesis is rejected. There is a negative relationship between geographic distance and the number of co-authored publications (Z= -13.31; p<0.001). Mass has a positive effect: the larger an organisation, the more

co-authorships it has (Z=36.48; p<.001 for A; Z=37.25; p<.001 for B).[3] The pseudo-R^2 is 0.318. If the natural logarithm of distance increases with one unit, the model predicts a decrease of .526 in the log of co-authored publications.

Table 3 Negative binomial regression for the basic gravity model

Goodness of fit:		
Dependent variable	**Co-authored pubs**	
N	46681	
Log likelihood	-2848.0934	
Pseudo R2	0.3181	
LR chi2	2657.39 ***	
LR $\alpha = 0$	592.51 ***	
Model:		
	Coefficient	**Z-score**
Ln (distance)	-.5260984***	-13.31
Ln (mass A)	.9718572***	36.48
Ln (mass B)	.8658232***	37.25

*** Two-sided significance at 1%-level.

Rathenau Instituut

The choice for a negative binomial model stems from overdispersion in the data. A likelihood ratio (LR) test confirms that parameter α is significantly different from zero (p =<.001). This shows that a negative binomial distribution is indeed more appropriate than a 'common' Poisson distribution because of the overdispersion. The model is, however, robust for slightly different specifications of the model: specifying it as a Poisson model does not alter the conclusions.

2.4.4 Extending the model with organisational proximity

The number of co-authorships between organisations is not only determined by their size and geographical distance. Research collaboration and innovative outcomes are associated with multiple dimensions of proximity, including geographic, organisational, social, and cognitive proximity (see also chapter 3). In this chapter, we test the effects of organisational proximity. Organisational proximity is defined in terms of differences between organisations in culture, incentive systems, knowledge bases, and so on (Boschma, 2005; Knoben and Oerlemans, 2006).

3 The dataset is constructed in such a way that every possible combination of two organisations in the set occurs exactly once. It is hence arbitrary whether an organisation is mentioned as "A" or "B". The fact that the coefficient of mass A is a bit higher is merely coincidence.

We have extended our gravity model with dummies for different types of organisations to test whether organisational proximity matters. We distinguish seven different types of organisations: universities, academic hospitals, (semi) public research organisations, consultancy firms, government bodies, industrial firms, and other organisations. The dummies are designed to take a value of one if two organisations belong to the same type of organisation and a value of zero if they belong to different types. The null hypothesis is that organisational proximity has no effect on co-authorship, which is why co-author relations between dissimilar organisations serve as a baseline.

The results are presented in Table 4. The parameters from the basic model hardly change. Geographic distance has a negative and highly significant effect (Z=-12.90; p<.001). The effect of organisational mass is positive and highly significant (Z=35.35; p<.001 and Z=35.96; p<.001). Universities, (semi) public research organisations, and other organisations are not likely to co-author more with organisations of the same type than with organisations of a different type. However, we find a significantly higher likelihood for academic hospitals (Z=8.10; p<.001), consultancies (Z=5.95; p<.001), industrial firms (Z=4.84; p<.001), and governmental bodies (Z=3.32; p=.001). The LR test confirms that in this situation a negative binomial distribution is appropriate. This implies that collaboration patterns are influenced both by organisational proximity – whether or not the two organisations have a similar background – and by the organisational background in itself. "Knowledge users" (academic hospitals, consultancies, industry, government) have a tendency to collaborate with organisationally proximate counterparts, while "knowledge producers" (public research organisations, universities) do not have a preference to collaborate with organisationally proximate partners.

Table 4 Extended gravity model with dummies for pairs of organisations of the same type

Goodness of fit:		
Dependent variable	**Co-authored pubs**	
N	46681	
Log likelihood	-2800.1604	
Pseudo R2	0.3296	
LR chi2	2753.26***	
LR $\alpha = 0$	582.03***	
Model:		
	Coefficient	**Z-score**
Ln (distance)	-.5084501***	-12.90
Ln (mass A)	1.006832***	35.35

continued on the next page ➔

Model:		
	Coefficient	**Z-score**
Ln (mass B)	.9222609***	35.96
Both university	.1834895	0.90
Both (semi) public res	.1482864	0.78
Both academic hospitals	1.672508***	8.13
Both consultancy	1.532714***	5.98
Both industrial	1.269188***	4.86
Both governmental	1.474479***	3.33
Both other	.3824729	0.99

Rathenau Instituut

*** Two-sided significance at 1%-level.

The impact of all dummies together is very small: the pseudo R^2 increased only slightly (pseudo R^2 =.330). A small effect does not imply that organisational proximity does not matter. It does, however, suggest that the dummies may explain part of the variance that is already explained by the variables of the basic model. To verify this, we have constructed a correlation matrix (Table 5).

Table 5 Correlation matrix of the dummies of the extended model against the variables of the basic gravity model.

	Ln (distance)	**Ln (mass A)**	**Ln (mass B)**
Both university	.004	.071***	.122***
Both (semi) public res	-.001	.057***	.038***
Both academic hospitals	-.024***	.010**	-.012***
Both consultancy	-.009**	-.045***	-.071***
Both industrial	.039***	-.048***	-.057***
Both governmental	-.010**	-.007	-.006
Both other	-.033***	-.020***	-.033***

Rathenau Instituut

** Two-sided significance at 5%-level.
*** Two-sided significance at 1%-level.

The correlation matrix shows that although many correlations are significant, they are not very strong. Apparently, for our data geographic proximity and organisational mass have a much stronger impact on collaboration than organisational proximity.

2.5 Conclusions and discussion

The production of scientific knowledge is increasingly a collaborative effort. Geographic distance is one of the factors that may explain the intensity of collaboration between organisations. The influence of geographic distance may be different at different spatial levels. The existing literature on the effects of

distance has basically focused on two levels: the micro-level of one organisation or building, and the macro-level of a large group of countries or an entire continent.

We show that geographical proximity also has a strong effect at the intermediate level of a small country, even in the geographically delineated water sector. Applying a gravity model to organisational co-authorships in scientific papers, we show that there is a clear negative relation between geographic distance and the number of co-authorships. As the distance between two organisations increases, the number of co-authored publications to which they both contributed decreases.

In a more elaborate specification of the gravity model we have tested whether organisational proximity has a similar effect on research collaboration. Organisations that can be characterised as "knowledge users" tend to collaborate more with organisationally proximate alters, whereas organisations that can be characterised as "knowledge producers" have no such preference. The effect of geographical distance does not change when controlling for organisational proximity.

2.5.1 Discussion

Our study finds a robust effect of geographical proximity on research collaboration. This finding corroborates the results of earlier studies, for example, by Ponds et al. (2007), Katz (1994), and Bouba-Olga et al. (2012). The effects found in this study are stronger than those found by Ponds et al. (2007) for the Netherlands. A possible explanation for this difference is that where Ponds et al. examined fundamental research, the subject of our study – water research – tends to be more applied. d'Este and Iammarino (2010) suggest that geographical proximity is much more important in applied research than in basic science. Applied research requires relatively more tacit (i.e. non-codified) knowledge, providing an incentive to collaborate with geographically proximate collaborators. The knowledge that is codified in co-authored publications is not universal, but is adapted to local questions and special circumstances. This "contextualized knowledge effect" can promote proximate collaborations. The contextualized knowledge effect may apply to all research fields that adapt to local conditions, such as most environmental sciences.

The effect of geographical proximity seems at least as significant and robust in the Netherlands as comparable studies have found for much larger countries (Smith and Katz 2000, d'Este and Iammarino 2010). On the other hand, geographical proximity was found to have a relatively weak effect for Australia, another large country (Katz, 1994). Does nation size matter? The difference between large and small countries may be moderated by people's perception of distance in addition to actual geographical distances. In small countries with geographically dense networks and many partners in close proximity, geographical distances

that are considered small in very large countries can be perceived as prohibitive to collaboration. This may be true for our specific case as well. The Netherlands is a densely populated area with clustered economic activity, where the location choices of high-tech industry depend especially on the local presence of knowledge infrastructure (Van der Panne and Dolfsma 2003).

One limitation of our study is that we have included only publications of which all contributors are located in the Netherlands. It is quite possible that long-distance collaborations in small countries always involves international partners. The smaller the country, the lower the odds that a suitable partner can be found within national borders. In addition, it is widely known that international research collaborations are more prestigious and produce a higher citation impact (Narin et al. 1991, Katz and Martin 1997), creating an incentive to find international partners. The open borders in the European Union and modern means of communication can make international distances easier to overcome. On the other hand, Hoekman et al. (2010) show that while the impact of national borders on co-authorships has decreased, the impact of absolute physical distance has not.

The strength of the effect of geographical proximity may be related to the nature of the collaboration partners. For example, more prestigious research appears to be less susceptible to distance (Sutter and Kocher, 2004). This is why we also tested for organisational proximity. We find that organisational proximity has an effect on research collaboration. Similar conclusions have been drawn in a number of studies that analyse various dimensions of proximity (e.g. Balland, 2012).

Ponds et al. (2007) found that including organisational proximity weakened the effect of geographical proximity in the physical sciences and even eliminated the effect in a few specific specialisations, especially for academic collaborations. They suggest that geographical proximity helps to overcome institutional or organisational differences between academic and non-academic organisations, even more so in the physical sciences which has a more mature structure with longer established relations between actors.

This interaction between organisational and geographical proximity does not occur in our data. The Dutch water sector has a long tradition of collaboration between different types of organisations (academic, semi-public, commercial, governmental). Nevertheless, the effect of geographical proximity is much stronger than the effect of organisational similarity. One explanation may be the contextualised knowledge effect. Another explanation is historical: actors may stick with their established, geographically proximate collaborators, even when more distant partners can be more easily reached. We do find that organisational proximity is important for organisations that can be characterised as "knowledge users". This is in line with the results of Scherngell and Barber (2009) that organisational proximity is more important for collaborations in industrial research than in public research. It is also likely that organisational proximity is

more important in applied fields of research than in fundamental research (d'Este and Iammarino 2010).

2.5.2 Future research

Three questions deserve more elaboration in future empirical research. The first question concerns the underlying causes of the effect of geographical proximity. Is it a deliberate choice of researchers to search for local collaborators? Various causes can be proposed. Researchers may be convinced that their questions are so contextualised and localised that only local partners can be of use in answering them; collaboration networks may have grown gradually from local roots that continue to form the backbone of partnerships; the transaction costs of maintaining long-distance collaborations may be prohibitive; or researchers are so inward-looking that they do not search for or meet potential partners from less proximate places. The second and related question concerns the balance between factors that promote proximate collaboration and factors that promote long-distance collaboration. In this study, we have shown that the former are stronger at the spatial scale of a small country. However, there is very little insight in the interactions between these "push and pull" mechanisms. The third question concerns the interaction with other dimensions of proximity (see also Frenken et al. 2009). Our analysis reveals that the type of organisation to which collaborators belong has an effect on the intensity of collaboration. This is confirmed by other studies as well. Yet, our understanding of the interplay between the different dimensions of proximity needs to be extended.

2.6 References

Acedo, F. J., C. Barroso, C. Casanueva & J. L. Galán (2006) Co-Authorship in Management and Organizational Studies: An Empirical and Network Analysis. *Journal of Management Studies,* 43, 957-983.

Aguiléra, A., V. Lethiais & A. Rallet (2012) Spatial and non-spatial proximities in inter-firm relations: *An empirical analysis. Industry and Innovation,* 19(3). 187-202.

Aldieri, L. (2011) Technological and geographical proximity effects on knowledge spillovers: evidence from the US patent citations, *Economics of Innovation and New Technology,* 20(6), 597-607.

Allen, T. J. (1977) *Managing the flow of technology: Technology transfer and the dissemination of technological information within the research and development organization.* Boston, MA: Massachusetts Institute of Technology.

Andersson, Å. E. & O. Persson (1993) Networking scientists. *The Annals of Regional Science,* 27, 11-21.

Balland, P.A. (2012) Proximity and the evolution of collaboration networks: Evidence from research and development projects within the Global Navigation Satellite System (GNSS) industry, *Regional Studies,* 46(6). 741-756.

Beaver, D. D. (2001) Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics,* 52, 365-377.

Beckmann, M. J. (1994) On knowledge networks in science: collaboration among equals. *The Annals of Regional Science,* 28, 233-242.

Blondel, V. D., J. L. Guillaume, R. Lambiotte & E. Lefebvre (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008, 10, P10008.

Boschma, R. (2005) Proximity and innovation: a critical assessment. *Regional Studies,* 39, 61-74.

Bouba-Olga, O., M. Ferru & D. Pépin (2012) Exploring spatial features of science-industry partnerships: A study on French data. *Papers in Regional Science,* 91(2): 355-376.

Bozeman, B. & E. Corley (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy,* 33(4), 599-616.

Breschi, S. & F. Lissoni (2003) *Mobility and social networks: Localised Knowledge Spillovers revisited.* Milan, CESPRI working paper 142.

Broekel, T. & R. Boschma (2012) Knowledge networks in the Dutch aviation industry: the proximity paradox. J*ournal of Economic Geography,* 12. 409-433.

Burger, M., F. van Oort & G. J. Linders (2009) On the specification of the gravity model of trade: zeros, excess zeros and zero-inflated estimation. *Spatial Economic Analysis,* 4, 167-190.

Chessa, A, et al. (2013) Is Europe Evolving Toward an Integrated Research Area? *Science,* 339(6120), 650-51.

De Solla Price, D. J. & D. Beaver (1966) Collaboration in an invisible college. *American psychologist,* 21, 1011.

Eichengreen, B. & D. A. Irwin (1998) The role of history in bilateral trade flows. In: Frankel, J.A. (Ed.) (1998) The Regionalization of the World Economy. *Chigago: University of Chigago Press*, pp. 33-63.

EIP, European Innovation Partnership on Water (2014), 'Barriers and bottlenecks for innovation in the water sector', *European Commission, Brussels*, 9 pp.

Este, P. d' & S. Iammarino (2010) The spatial profile of university-business research partnerships. *Papers in Regional Science,* 89(2), 335-351.

Flowerdew, R. & M. Aitkin (1982) A Method of Fitting The Gravity Model Based on the Poisson Distribution. *Journal of Regional Science,* 22, 191-202.

Frenken, K., S. Hardeman & J. Hoekman (2009) Spatial scientometrics: Towards a cumulative research program. *Journal of Informetrics,* 3, 222-232.

Frenken, K., R. Ponds & F. van Oort (2010). The citation impact of research collaboration in science-based industries: A spatial-institutional analysis. *Papers in Regional Science,* 89(2), 351-271.

Gazni, A., C. R. Sugimoto & F. Didegah (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the American Society for Information Science and Technology,* 63(2), 323-335.

Hagstrom, W. (1965) *The Scientific Community.* New York: Basic Books.

Havemann, F., M. Heinz & H. Kretschmer (2006) Collaboration and distances between German immunological institutes–a trend analysis. *Journal of Biomedical Discovery and Collaboration*, 1, 6.

Hoekman, J., K. Frenken & R. J. W. Tijssen (2010) Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy,* 39, 662-673.

Hoekman, J., K. Frenken & F. van Oort (2009) The geography of collaborative knowledge production in Europe. *The Annals of Regional Science,* 43, 721-738.

Katz, J. S. (1994) Geographical proximity and scientific collaboration. *Scientometrics*, 31, 31-43.

Katz, J. S. & B. R. Martin (1997) What is research collaboration? *Research Policy,* 26, 1-18.

King, G. (1988) Statistical models for political science event counts: Bias in conventional procedures and evidence for the exponential Poisson regression model. *American Journal of Political Science*, 838-863.

Knoben, J. & L. Oerlemans (2006) Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8, 71-89.

Kraut, R., C. Egido & J. Galegher. 1988. Patterns of contact and communication in scientific research collaboration. In: *Proceedings of the Conference on Computer-Supported Cooperative Work*, Portland OR.

Laudel, G. (2002) What do we measure by co-authorships? *Research Evaluation,* 11, 3-15.

Maggioni, M. A. & T. E. Uberti (2009) Knowledge networks across Europe: which distance matters? *The Annals of Regional Science,* 43, 691-720.

Malecki, E. J. (2010) Everywhere? The Geography of Knowledge. *Journal of Regional Science,* 50, 493-513.

Melin, G. & O. Persson (1996) Studying research collaboration using co-authorships. Scientometrics, 36, 363-377.

Melin, G. (2000). Pragmatism and self-organization: research collaboration on the individual level. *Research Policy*, 29(1), 31-40.

Merino, F. & L. Rubalcaba (2012) Are knowledge-intensive services highly concentrated? Evidence from European regions. *Tijdschrift voor Economische en Sociale Geografie*, 104 (2), 215-232.

Murray, A. T. (2010) Quantitative Geography. *Journal of Regional Science*, 50, 143-163.

Nagpaul, P. (2003) Exploring a pseudo-regression model of transnational cooperation in science. *Scientometrics*, 56, 403-416.

Narin, F., K. Stevens & E. S. Whitlow (1991) Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21, 313-323.

Panne, G. van der & W. Dolfsma (2003) The odd role of proximity in knowledge relations: high-tech in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie*, 94 (4), 453-462.

Paci, R. & S. Usai (2009) Knowledge flows across European regions. The Annals of *Regional Science*, 43, 669-690.

Ponds, R., F. van Oort & K. Frenken (2007) The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86, 423-443.

Rijnsoever, F. J. van & L. K. Hessels (2011) Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy*, 40, 463-472.

Scherngell, T. & M.J. Barber (2009) Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU Framework Programme. *Annals of Regional Science*, 46, 247-266.

Sen, A. & T. E. Smith (1995) *Gravity models of spatial interaction behavior*. Heidelberg: Springer.

Silva, S. & S. Tenreyro (2006) The log of gravity. *Review of Economics and Statistics*, 88, 641-658.

Smith, D. & J. S. Katz (2000) Collaborative approaches to research. A report to the Higher Education Funding Council for England. Centre for Policy Studies in Education, University of Leeds. Smith, M. (1958) The trend toward multiple authorship in psychology. *American psychologist*, 13, 596.

Sternberg, R. (2007) Entrepreneurship, proximity and regional innovation systems. *Tijdschrift voor Economische en Sociale Geografie*, 98 (5), 652-666.

Subramanyam, K. (1983) Bibliometric studies of research collaboration: A review. *Journal of information Science*, 6, 33-38.

Sutter, M. & M. Kocher (2004) Patterns of co-authorship among economics departments in the USA. *Applied Economics*, 36, 327-333.

Thomas, A. Duncan & R. R. Ford (2005) *The Crisis of Innovation in Water and Wastewater*. Cheltenham, UK, Edward Elgar. 336 pp.

Ukrainski, K., J. Masso & H. Kanep (2014) Cooperation patterns in science in Europe: the standpoint of small countries. *Scientometrics*, 99, 845-863.

Wagner, C. S. & L. Leydesdorff (2005) Mapping the network of global science: comparing international co-authorships from 1990 to 2000. *International Journal of Technology and Globalisation*, 1, 185-208.

Weterings, A. & R. Ponds (2009) Do Regional and Non-Regional Knowledge Flows Differ? An Empirical Study on Clustered Firms in the Dutch Life Sciences and Computing Service Industries, *Industry and Innovation*, 16 (1), 11-31.

3 How do dimensions of proximity relate to the outcomes of collaboration? A survey of knowledge intensive networks in the Dutch water sector[4]

3.1 Introduction

The literature agrees on the benefits of collaboration in knowledge intensive processes (Hagedoorn, Link and Vonortas, 2000; Hoekman, Frenken and Tijssen, 2010; Katz and Martin, 1997). Much less is known about configurations that stimulate effective collaboration, leading to targeted outcomes such as knowledge production, innovation and joint publications. Research policy favours specific collaborations, such as public-private partnerships, while it is not clear what conditions are favourable for what kind of outcomes. Various studies suggest that proximity is a key concept in understanding the configurations of collaboration in knowledge production (see Boschma, 2005 for an overview). The basic premise is that proximate people have a tendency to collaborate, as it is easier to communicate with people who are close. On the other hand, the advantage of collaboration may disappear when people become "too close" (Nooteboom et al., 2007). There is a substantial body of work on the relation between geography and innovation (Autant-Bernard et al., 2007; Broekel and Boschma, 2012; Porter, 2000). Gravity models show that geographic proximity can explain coauthorship in scientific publications (Hoekman, Frenken and Tijssen, 2010; Ponds, Oort and Frenken, 2007). Ethnographic studies, for example on business development and technology acquisition around CERN, show the importance of cognitive and social proximity for successful collaboration (Autio, Hameri and Vuola, 2004).

The growing body of literature on proximity is rich and diverse, but contributions often share three limitations. First, most empirical studies focus on one dimension of proximity. The earlier work by economic geographers on co-location has led to the insight that in addition to geographical proximity other dimensions are relevant in knowledge production and innovation (Boschma, 2005). This has resulted in studies that analyse the effect of diverse dimensions of proximity in recent years (e.g. Aguiléra, Lethiais and Rallet., 2012; Broekel and Boschma,

4 This chapter has been published as: Heringa, P.W., E. Horlings, M. van der Zouwen, P. van den Besselaar and W. van Vierssen (2014) How do dimensions of proximity relate to the outcomes of collaboration? A survey of knowledge-intensive networks in the Dutch water sector. *Economics of Innovation and New Technology* 23(7), 689-716.

2012). Second, the analysis of the impact of proximity on the outcomes of collaboration mostly focuses on publications and patents. The focus on publications and patents as proxies for learning, knowledge production or innovation may give an incomplete picture of the effect of proximity. The use of datasets on patents and publications without additional data limits the possible indicators of proximity to the variables stored in the dataset, which sometimes are at best proxies for the dimensions of proximity. Third, they consider a relatively homogeneous group of people from one societal sector (for example scientists or professionals from industry). Analysing a homogeneous group of actors (all from science, or all firms, for example) may give a limited view on the effect of proximity. Proximity may work differently in a field with relatively much organisational and cognitive variance (i.e. people with strongly different expertises and from very different organisations) compared to a field that is relatively homogeneous.

In this study we contribute to the proximity literature by investigating the relation between different outcomes of collaboration (such as publications, innovations, but also more intangible outcomes like exchange of ideas) and the degree of geographical, social, cognitive and organisational proximity between collaborators. In our data, we do not distinguish between outcomes of collaboration that are expected and that are already achieved. So, throughout this article, when we refer to outcomes, this concerns both expected and achieved outcomes. We elaborate further on this point in section 3.3.5. Our empirical analysis is based on a survey among professionals in the Dutch water sector. The water sector involves a wide variety of knowledge disciplines and societal sectors, resulting in a large variety in organisational and cognitive backgrounds of collaborators. The use of a survey allows us to use a larger number of indicators for different dimensions of proximity than the analysis of patents and publications. Our study is part of a recent trend to use surveys to assess the different dimensions of proximity. Aguiléra, Lethiais and Rallet(2012) have used survey data in their study on the impact of proximity on network formation to develop a typology of eight different types of relationships, each with their own geographical scale and need for coordination. Ferru (2010) combines contract data with survey data. This allows her to show that the pattern of local partnerships tends to be reinforced over time, because people prefer to collaborate with alters they know – even if those are not the most appropriate partners in terms of available resources – over searching for new partners. Weterings and Ponds (2009) use a survey and (for geographical proximity) arrive at different conclusions than conventional studies: they show that although most collaborations occur within a region, the most valuable knowledge exchange takes place in interregional collaborations.

The remainder of this chapter is structured as follows. In section 3.2 we introduce a conceptual framework and explain how it relates to earlier research. In section 3.3 we explain how we applied the concepts to our case and how we have collected our data. In section 3.4 we discuss the results. In section 3.5 we draw conclusions and raise some issues for future research.

3.2 Conceptual framework

We apply a multidimensional model of proximity that includes a geographical, social, organisational and a cognitive dimension between two collaborators, the ego and the alter. Our aim is to find out what dimensions of proximity are conducive to the outcomes of collaboration.

3.2.1 Dimensions of proximity

The first literature on proximity focused entirely on geographical proximity (e.g. Audretsch and Feldman, 1996). Over time other dimensions, such as organisational, institutional, cultural, cognitive, technological and social proximity have been added. Authors have come up with a wide range of categories of proximity, each with their own definition and operationalisation (see for an overview Knoben and Oerlemans, 2006). The common denominator of these dimensions is that being proximate in any of them can enhance coordination, reduce uncertainty and thus contribute to knowledge production and innovation (Boschma, 2005). Review papers by Boschma (2005) and Knoben and Oerlemans (2006) show that there is much overlap between some of the concepts in the literature, either because different labels are used for the same idea or because umbrella terms are used that include several other concepts. To give one example: what is termed 'social proximity' in this thesis, is also called 'personal proximity' or 'relational proximity' by others (Schamp, Rentmeister and Lo, 2004; Coenen, Moodysson and Asheim, 2004). Knoben and Oerlemans (2006) distinguish three dimensions: organisational, cognitive (or technological) and geographical proximity. Boschma (2005) identifies two more: social and institutional.

In our analysis we distinguish four dimensions of proximity, namely social, organisational, cognitive and geographical. We disregard the institutional dimension. Institutional proximity entails humanly devised constraints that structure political, social and economic interaction (North, 1991). At the dyadic level of individual interactions, institutional differences and similarities can be considered part of organisational proximity (Knoben and Oerlemans, 2006). Ponds, Oort and Frenken (2007) for example use the difference between academic and non-academic organisations as an indicator of institutional proximity. In our framework this is part of organisational proximity. At the level of communities and systems, institutional proximity can also concern differences in values and norms, the macrolevel in North's framework. This is sometimes measured using proxies such as language or shared law systems (Boschma, 2005). In a small and culturally homogeneous country, measuring such differences with data on one sector would require questions that are difficult to implement concisely in a survey (Aguiléra, Lethiais and Rallet, 2012). We do distinguish between organisational and social proximity. Social proximity refers to personal aspects of collaboration (mutual trust, kinship), whereas organisational proximity (at the dyadic level) focuses on similarities and differences in the organisational context. The same four dimensions of proximity are selected in a recent empirical study on the Dutch aviation industry (Broekel and Boschma, 2012). In Table 6 we

give an overview of the dimensions of proximity we use, with references to recent empirical works that use the same (or a similar) concept. In section 3.3.2 we describe in more detail how these four dimensions are operationalized and measured in our study.

Table 6 Correlation matrix of the dummies of the extended model against the variables of the basic gravity model.

Proximity dimension	Description	References
Geographical	Distance "as the crow flies" between working place of ego and alter (sometimes combined with other geographical indicators such as national and regional borders)	Aguiléra, Lethiais and Rallet, 2012 Aldieri, 2011 Autant-Bernard et al., 2007 Balland, 2011 Broekel & Boschma, 2012 Cunningham & Werker, 2012 Ferru, 2010 Hoekman, Frenken and Tijssen, 2010
Social	Social embeddedness of ego and alter (involving trust, based on friendship, kinship, personal experiences)	Aguiléra, Lethiais and Rallet, 2012 Autant-Bernard et al., 2007 Balland, 2011 Broekel & Boschma, 2012 Cunningham & Werker, 2012 Fleming, King and Juda, 2007 Ter Wal, 2009
Organisational	Similarity in incentives and routines between organisations of ego and alter	Aguiléra, Lethiais and Rallet, 2012 Balland, 2011 Broekel & Boschma, 2012 Cunningham & Werker, 2012 Ponds, Oort and Frenken, 2007
Cognitive	Similarity in the professional knowledge base of ego and alter	Aguiléra, Lethiais and Rallet, 2012 Aldieri, 2011 Balland, 2011 Broekel & Boschma, 2012 Cantner & Meder, 2007 Cunningham & Werker, 2012 Nooteboom et al., 2007

3.2.2 Outcomes of collaboration

Aguiléra, Lethiais and Rallet (2012) distinguish three streams of literature on proximity. The first stream studies the links between proximity and network formation (e.g. Autant-Bernard 2007; Ferru, 2010). The second stream analyses the impact of proximity on the economic performance of firms (e.g. Broekel and Boschma, 2012). The third stream investigates the impact of the different dimensions of proximity on knowledge production and sharing (Boschma, 2005, Knoben and Oerlemans, 2006). Our study can be positioned in this last stream. An overview of findings in the literature since 2005 is provided in Table 7. It immediately stands out from this overview that earlier studies either measure the impact on innovative performance, or on one single type of outcome.

Table 7 Overview literature findings since 2005 on the effect of the dimensions of proximity.

Source	Outcome	Scale/field	Geo	Soc	Org	Cogn	Findings
Aldieri, 2011	Patents	Europe, USA, Japan.	+			+	Knowledge streams from patent citations stronger between proximate firms, both geographically and technologically. Geographical effect stronger in Europe.
Autant-Bernard et al., 2007	Projects in Framework Programme6	EU; nanotech	0/+	+			Social distance matters more than geographical distance. Geographical does not matter for firms involved in more projects.
Balland, 2012	Projects in FP6	EU; satellite navigation	+	0	+	0	Cognitive proximity does not have an effect because organisations also need access to different knowledge in the industry branch.
Bouba-Olga, Ferru and Pépin, 2012	Science-industry alliances	France	+		+		Strong differences between regions in the effects of geographical and sectoral proximity.
Broekel & Boschma, 2012	Innovative performance (interviews)	Dutch aviation industry	+	+	0	-	Strong evidence that too much cognitive proximity lowered firms' innovative performance, and organisational proximity did not have an effect.
Cantner & Meder, 2007	Patents	Germany				+	Technological overlap of two organisations increases probability that they cooperate.
Cunningham and Werker, 2012	Publications	EU; nanotech	+		+	+	Geographical proximity not only physical distance, also territorial. Cognitive proximity neither too far nor too distant. Organisational proximity only indirectly: non-academic partners are cognitively more proximate.
Hoekman, Frenken, and Tijssen, 2010	Publications	EU	+				Impact of physical distance is stable over time; impact of territorial borders has decreased over time (2000-2007).
Nooteboom et al., 2007	Innovative performance (survey)	Chemical, pharmaceuti-cal, automotive				+/-	U-shaped relationship between cognitive proximity and innovative performance. Effect is stronger in explorative collaborations than in exploitative collaborations.
Ponds, Oort and Frenken, 2007	Publications	8 Dutch scientific fields	+		+		Geographical proximity has a stronger effect if organisational proximity is lower.
Ter Wal, 2009	Patents	Germany biotech	+	+		+/-	Social proximity is the strongest predictor. Controlling for social, the effect of geographical is weak. The effect of cognitive is positive, but turns negative if controlling for the other two dimensions.
Weterings & Ponds, 2007	Knowledge exchange (survey)	Dutch computing & life sciences	+				Firms have more knowledge exchange with proximate others, but the knowledge flows exchanged with more distant partners are valued higher.

Table 7 also makes clear that, so far, the literature focused on hard, tangible outcomes of collaboration. Many studies are based on data about co-authorship of publications (e.g. Hoekman, Frenken and Tijssen, 2010) or co-ownership of patents (e.g. Wal, 2009; for a more extensive overview see Bouba-Olga, Ferru and Pépin, 2012). There are large datasets with these types of data, which allows for an analysis of many different relations. However, knowledge production and innovation entail much more than can be captured in publications and patents. Many innovations for example are not patented but protected in other ways or even shared openly. Especially non-profit organisations store and share their knowledge in other forms than patents and scientific publications, for example by personal communication between people or in non-scientific publications.

It is an important question whether proximity has the same impact on tacit knowledge (which is often shared informally and cannot be traced in patents or journal publications) as on formal codified knowledge (Aguiléra, Lethiais and Rallet, 2012; Balland, Suire and Vicente, 2013). To date, there is little empirical work on the relation between proximity and informal knowledge production and sharing between collaborators. An exception is the study of Aguiléra, Lethiais and Rallet (2012) who assume that collaborators who indicate a great need of coordination will also exchange tacit knowledge. They then show that non-spatial proximities are especially important in relations in need of coordination. Another exception is the work of Weterings and Ponds (2009) who excluded all formal R&D collaborations in their study on the difference between intra-regional and inter-regional knowledge flows. Attention for informal knowledge production and exchange is especially important in the water sector where patenting is rare even for profit organisations, and where many non-profit organisations are involved in knowledge production.

3.2.3 The relation between proximity and outcomes of collaboration

Each dimension of proximity has an impact on the outcomes of collaboration. We briefly discuss earlier findings per dimension.

Geographical proximity can stimulate and facilitate processes of learning and innovation, sometimes by complementing or substituting other dimensions of proximity (Rallet and Torre, 1999). Earlier studies on patents and publications confirm that collaboration is more intense across smaller geographical distances (see Bouba-Olga, Ferru and Pépin, 2012 for an overview). However, Weterings and Ponds (2009) use data from a telephone survey to show that, although most collaborations are geographically proximate, the ones across larger distances are considered more valuable and more often concern knowledge exchange on technological issues.

Social proximity is considered to facilitate and foster joint knowledge production and knowledge exchange (Broekel and Boschma, 2012). It has been shown empirically that social proximity (using the proxy of a collaboration history in the

past) leads to more joint patents (e.g. Wal, 2009). For collaboration in EU Framework Programme projects on micro- and nanotechnologies, the number of common acquaintances in the network and network distance have an effect on the likelihood of collaboration (Autant-Bernard et al., 2007). Then again, Balland (2012) shows for Framework Programme projects in the navigation industry that the partners of partners in the project (which he defines as social proximity) are not more likely to interact than random actors. It is also argued that too much social proximity can be detrimental for effective learning and innovation because a relationship largely based on trust and loyalty may lead to an underestimation of opportunistic behaviour (Boschma, 2005). However, to the best of our knowledge, this has yet to be proven empirically.

Organisational proximity is said to reduce the uncertainty and opportunism involved in knowledge creation. It provides control mechanisms required to protect intellectual property and ensure rewards for the knowledge produced (Boschma, 2005). Broekel and Boschma (2012) show a positive effect of organisational proximity on knowledge network formation among firms, but no effect on their innovative performance. Cunningham and Werker (2012) find that collaborations with only academic partners are better able to overcome large technical distances than mixed or non-academic collaborations. There is no empirical evidence for a negative effect of too much organisational proximity on (forms of) knowledge production and exchange.

Regarding cognitive proximity, Nooteboom (1999) argued that for novelty cognitive distance is required, small enough to be able to understand each other and efficiently process the acquired information, yet large enough to yield new knowledge. The empirical evidence is mixed. Cantner and Meder (2007) use patent data to show that technological overlap between collaborators contributes to the likelihood that they collaborate. Wal (2009), also using patent data, finds a weak negative effect of cognitive proximity in a multivariate model that controls for geographical and social proximity, but a positive effect in a univariate model. Broekel and Boschma (2012) find a negative effect on innovative performance. Cantner and Meder (2007) explicitly test for an inverted U-curve, but do not find one. However, Nooteboom et al. (2007) find an inverted U-curve for explorative patents (though not for exploitative patents).

Few studies include an interaction effect between different dimensions of proximity. They examine the effect on network formation and give mixed results. Breschi and Lissoni (2003) find with patent data that geographical proximity is only relevant if there is a social connection between patents. Ponds, Oort and Frenken, (2007) find a smaller effect for geographical proximity in collaborations between academic organisations than in collaborations between academic and non-academic organisations. However, Broekel and Boschma (2012) find that geographical, social, organisational and cognitive proximity all four have an effect on knowledge network formation, also when controlling for the other

dimensions. This is to the best of our knowledge the only study that includes interaction effects and tests four dimensions. Cunningham and Werker (2012) test a model with geographical, organisational and technical proximity. They find that geographical proximity is statistically most significant, although technical proximity has the largest effect. Organisational proximity only has an indirect effect; the different types of organisations differ in their absorption of new knowledge, with non-academic organisations being more specialised. Wal (2009) finds that the positive effect of cognitive proximity turns into a weak negative effect if he controls for geographical and social proximity.

Our hypothesis is that proximity has a different effect on different outcomes of collaboration. We expect that proximate relations yield "everyday" outcomes of knowledge exchange; the intangible outcomes like exchange of knowledge or support for ideas. Because the more distant relations have higher transaction and coordination costs, such relations probably aim for specific, tangible outcomes like innovation or publications.

3.3 Data and Methods

Our results are based on a survey among the members of the Royal Dutch Water Network. The Network is a society of 3,468 individual water professionals aiming to increase their expertise by exchanging experiences and knowledge. All members have received a personal invitation to answer a variety of questions. Respondents (egos) have been asked to:

- provide information on personal characteristics (age, educational level, etcetera);
- randomly select three persons from their external professional network (alters);
- provide their perspective on a number of personal characteristics of those alters;
- assess the proximity of the relation by answering questions on each dimension; and
- identify the benefits that were expected or had been achieved in each relationship.

A total of 618 respondents have returned the questionnaire. Since each respondent was asked to provide information on three relationships, the maximum number of relationships that can theoretically be analysed is 1,854. However, not all respondents have provided complete information on all three relations. In this study we only analyse the 1020 relationships for which all questions were answered. There is a number of limitations to the survey data. First, we have only asked the respondents about their perception of the collaboration with three of their alters; we cannot observe how that differs from the perception of the alters on the same collaboration. Second, the survey data are inherently subjective in nature; we measure the perceptions of the respondents. Third, there may be

biases by the alters in the selection of alters to report on. We elaborate in more detail on these limitations in section 3.5.3.

3.3.1 The Dutch water sector

Our data have been collected in the water sector in the Netherlands. The Netherlands is a small country in geographical terms; relatively small differences in distance can have considerable impact on people's perception. However, it is densely populated and shows high internal diversity. About 50% of the country (the western part, where about 70% of its GDP is earned) is low-lying and flood-prone, because, although safely behind dikes, it is below sea level (Kabat et al., 2005). One can understand that water safety and security are considered important. This diversity in combination with its relatively compact make the country interesting for proximity research.

Proximity mechanisms may have different effects in different sectors. Vinciguerra et al. (2011) show that the importance of geographical proximity may be technology-specific. We study the water sector; this is delineated as all activities related to the water cycle (production, collection, distribution (grid maintenance) and treatment of drinking water and wastewater; water management). The water sector is directly linked to grand societal challenges. Rockström et al. (2009) have identified nine planetary boundaries; transgressing them is potentially catastrophic because of the risk of transgressing thresholds that trigger abrupt environmental changes in continental and even planetary-scale systems. A deeper understanding of water and water management is required for several of these planetary boundaries, notably global freshwater use, climate change and the nitrogen and phosphorus cycle. This is also recognized by policymakers; it is for example directly related to several of the grand challenges mentioned in Horizon 2020 as crucial for Europe (notably climate change and depletion of natural resources and food security and sustainable agriculture).

Regarding organisational and cognitive proximity, it is important to note that the Dutch water sector itself entails a set of heterogeneous actors. A water sector typically envelopes a whole range of intertwined organisations specific, yet complementary roles. Therefore, when we refer to Dutch "water sector" we first of all mean the collaborative community of public organisations such as water utilities (10 drinking water companies), water boards (25) and municipalities (408). But also the attached industrial conglomerate of service providers, R&D departments of technology manufacturers as well as the public research infrastructure of universities and applied research institutes and research intermediaries who commission research.. Moreover, private consultants play an important role in the generation and transfer of knowledge to the operations (Muizer and Van den Berg, 2002).

In many aspects the Dutch water sector is similar to the ones in other European countries. Its utilities are public as is the case in the vast majority of countries in

Europe, with the exception of the UK and France. However, the consolidation process in a sector that is, worldwide, notorious for its fragmentation, is remarkable in the Netherlands. The scale and geographical coverage of the drinking water companies has increased substantially over the past 70 years. In 1940 there were 210 water supply companies in the Netherlands; this decreased to 14 in 2004 (Moel, Verberk and Dijk, 2006) and 10 at present. All companies have their own service area; there is hence no direct competition in drinking water supply and distribution or wastewater treatment. The consolidation in the domain of waste water treatment and waster safety (waterboards) is even more considerable. The number decreased from approximately 2600 in 1945 to 25 in 2013. Van Vierssen (2012) estimates that for Europe as a whole, the Netherlands has meanwhile scaled up operations with a factor 100 as compared to the average situation in Europe.

However, as Thomas and Ford (2005) state, there are concerns that because the sector is too orthodox and lacks an innovative culture, it will fail to deliver the breakthroughs required for high-quality water services in the coming century. This is attributed to a lack of integration and collaboration between actors of different types (e.g. firms with knowledge institutes), and myopia with regards to technology and innovation, which is reinforced by regulatory and policy frameworks (Thomas and Ford, 2005).

The Dutch water sector would like to invest in knowledge production and innovation to strengthen its (international) position (Stumpe, 2011); this will require stronger collaboration, both between different organisation types and between different subsections of the water sector (Muizer and Van den Bergh, 2002). Traditionally, the sector is strongly organised in pillars (like drinking water, wastewater, distribution, water management); there is recently attention for the need to integrate those. Governmental agencies from across the sector (from national agencies to municipalities and from drinking water related agencies to water management agencies) have expressed their willingness to collaborate with private parties and research organisations on innovative projects; water also has a clear position in the Dutch sectoral innovation policy (Stumpe, 2011). There is also more attention for integration with other sectors; water management for example is now more integrated with related policy fields such as nature preservation, spatial planning, agriculture than a few decades ago; parties in the water sector are in have interactions with other relevant actors (Brugge, 2009). It is hence a very interesting field to test how mechanisms like organisational and cognitive proximity currently shape patterns of collaboration in knowledge production.

3.3.2 Operationalising dimensions of proximity

The choice for a survey to collect the data allows for more refined indicators of the other dimensions of proximity than the ones that are common in the literature. Per dimension we will explain how it is usually measured and how our measures relate to the definition of each dimension.

Geographical proximity was measured by asking the respondents to list both the city where they (most often) work and the city where their relations work. Due to a technical error, the cities of the relations were not stored in our dataset. However, other details (such as the name of the organisation) were stored, and we have used that information to retrieve the cities of the relations where possible. We have identified the latitude and longitude of each city and calculated the distance between each pair of cities using the formula for great-circle distances (Sinnot, 1984). In other words, distances refer to the shortest possible distance between two points on a sphere, "as the crow flies".

Social proximity refers to the social embeddedness of the collaboration. Social embeddedness involves trust, based on friendship, kinship, personal experiences (Boschma, 2005; Broekel and Boschma 2012). This cannot be inferred directly from data on patents or publications. Many studies therefore measure the social connectedness based on the collaboration history of actors (such as earlier co-authorships) (Breschi and Lissoni, 2003) or the geodesic distance in a social network (Balland, 2012; Cunningham and Werker, 2012) as a proxy for social proximity. Such social connectedness can indeed be a source and indication of social proximity: the fact that an ego repeatedly collaborates with the same alter indicates a basic form of mutual trust and social proximity. However, the fact that collaborators do not have a formal track record of past publications does not imply they are not socially proximate. Moreover, the fact that people have a history of collaboration may say as much about their cognitive proximity (their ability to understand each other's knowledge so they can fruitfully collaborate) as about their social proximity. We have hence decided to measure social proximity more directly by asking about trust and the nature of the relationship. Trust is considered a central element of social proximity. For measuring trust (the items Trust, Effort, and Share), we have used questions from existing surveys on trust (Levin and Cross, 2004, McAllister, 1995). In addition, we asked for details about the nature of the relationship, for example whether ego and alter know each other as peers in former jobs or went to school together or have a contractual relationship. By asking for personal characteristics of both the respondent and his or her relations, we could also examine whether similarity in age and gender contributes to social proximity.

Organisational proximity can be defined as the degree of similarity in routines and incentive mechanisms (Metcalfe, 1994). In innovation literature a distinction is often made between profit and non-profit organisations, as they clearly have different incentive mechanisms and, hence, different routines. Profit organisations for example have incentives to hide knowledge from their competitors, whereas non-profit organisations often have a mission for open knowledge exchange (Broekel and Boschma, 2012). Given the large variety of organisations in our sample, we have extended the possible categories to four societal sectors (business, government, academia, NGO). We have added a question to ask specifically about the differences in intellectual property protection between the

organisations of alter and ego. Moreover, in the literature on organisational cultures (Ashkanasy, Wilderom and Peterson, 2000; Delobbe et al., 2002; Denison and Mishra, 1995; Hofstede 1998) many indicators are described to give some basic characterisation of an organisation, focusing on differences in incentive mechanisms and routines in organisations. They are therefore useful for measuring organisational proximity. As the range of organisations in our group of respondents is very wide, we have chosen a few universal indicators. They measure a focus on procedures versus results, the capacity to adapt to new circumstances, the strictness of planning and financial management and the freedom to engage in external contacts.[5]

Cognitive proximity concerns the similarity in the knowledge base of alter and ego (Boschma, 2005; Aguiléra, Lethiais and Rallet, 2012). It is very similar to the concept of technological proximity. However, technological proximity is often defined a bit more narrowly as differences in the technical knowledge base of collaborators (Knoben and Oerlemans, 2006). This is usually operationalised as a similarity in technical class, e.g. on the basis of industrial classisifications (such as the NACE classification) or by creating technological profiles for each organisation based on patent classifications (Aguiléra, Lethiais and Rallet, 2012; Wal, 2009). Cognitive proximity is somewhat broader; it refers to all knowledge actors hold, and their ability to interpret or absorb the knowledge exchanged (Mattes, 2012). We have measured the cognitive distance using items that indicate whether ego and alter share specific expertise. Using the same concepts and terms (speaking the same "language") is an indicator of a similar knowledge base. We have therefore included a question on the extent to which alter and ego use the same jargon when they interact. The same goes for expertise on specific instruments and machinery, the second indicator we have included. Furthermore, they indicate to which part of the water cycle their own work and that of their relations belongs. This is an additional measure for overlap in technical expertise.

3.3.3 The explanatory variables

The four dimensions of proximity have been measured using different questions, thus producing the explanatory variables in our model. Table 8 describes the explanatory variables in detail. All ordinal variables in this table were measured as a 5 or 6 point Likert scale.

Geographic proximity is defined as the inverse of geodesic distance between the cities where ego and alter work. The more proximate cities are, the shorter the distance between them. By using the inverse, more proximate relations have a

5 To keep the required response time for the survey within limits, we have not asked the respondents to score both their own organisation and that of their collaborators on these items. Instead, we have asked about the difference between the two organisations.

higher score, in line with the other variables. Many scholars employ further alterations to avoid the problem that the inverse of zero distance is not defined (see e.g. Aldieri and Cincera, 2009). However, the smallest distance in our case is 3.5 km between cities. For collaboration within the same city we have assumed a fixed distance. Sensitivity analysis shows that different standard values for this fixed distance do not alter the effect size or significance level of any of the results. We have tested several values in a range from 1 to 5 km and we use a distance of 5 km as standard value in the reported figures.

Social proximity has been operationalized using two groups of variables. SP-Effort, SP-Trust and SP-Share provide a direct indication of social proximity, while variables such as age and gender, that have a primary function as control variable, are also informative with respect to social proximity. SP-Effort, SP-Trust and SP-Share were measured on a 6-point Likert scale. However, in each case, few respondents indicate low proximity. For statistical purposes we have aggregated the scores 1 and 2 into one group.

Age difference, Frequency and Time are categorical variables. Age difference had five categories (from much younger to much older). As proximity is about distance, the answer categories have been recoded to "more or less the same age", "some difference in age", and "large difference in age". The question on Frequency had six response categories, but the frequencies at both extremes (scores 1 and 6, meeting daily and meeting less than once a year respectively) were so low that they have been aggregated with their adjacent categories. Time has five categories (from less than one year to over 10 years).

The variables that measure organisational proximity were measured on a scale from 'organisation A much more so than organisation B' to 'organisation B much more so than organisation A'. However, from a proximity point of view it does not matter which organisation has a higher score, but rather how large the difference between the two organisations is. Therefore, the answers to these variables have been recoded to a scale ranging from 'there is a large difference between the organisations' to 'the organisations are about the same'.

To identify a common domain in the water sector, respondents were asked to indicate whether or not they considered themselves experts in nine areas within the water sector (collection of drinking water, production of drinking water, distribution of drinking water, sewerage collection, sewerage transport, wastewater treatment, water management, another water area or no expertise related to the water cycle). The respondents were also asked to indicate whether or not they considered their relations as experts in these fields. Many professionals appear to have expertise in more than one of these areas. Factor analysis revealed five strong clusters: drinking water, sewerage, wastewater treatment, water management and non-water cycle. These five are used to measure if respondents and their relations have at least one common area of expertise.

Table 8 Description and descriptive statistics of the explanatory variables.

Variable	Description	N	Median	Min.	Max.	Measurement type
Geographical proximity						
Inverse geodesic distance	The inverse of the geodesic distance between the work locations of ego and alter (measured at city level).	541	.0200 (50.0 km)	.004 (261.4 km)	.29 (3.5 km)	Continous
Social proximity						
SP-Effort:	Willingness of respondent to put effort into something the alter asks him to do.	1010	5	2	6	Ordinal
SP-Trust:	Trust ego has in contributions of alter.	1007	5	2	6	Ordinal
SP-Share:	Willingness of ego to share information with alter.	1013	5	2	6	Ordinal
Same Gender:.	Whether ego and alter are of same gender	1020	1	0	1	Dichotomous
Age difference:	Age difference between ego and alter.	1015	2	1	3	Ordinal
Time:	Time the ego and alter know each other.	1015	4	1	5	Ordinal
Frequency:	Frequency at which ego and alter meet each other.	1020	3	2	5	Ordinal
Private:	Dummy indicating whether or not ego and alter have a non-business relation (one or more of: members of the same association, friend, had the same education, former colleague/employee).	1020	0	0	1	Dichotomous of underlying questions
Organisational proximity						
OP-Adapt:	Degree of difference between organisation of ego and alter in (easily) adapting to new circumstances.	925	2	1	3	Ordinal
OP-Management:	Degree of difference between organisation of ego and alter in strictness of planning and financial management.	900	2	1	3	Ordinal
OP-External:	Degree of difference between organisation of ego and alter in freedom to initiate relations outside own organisation.	908	3	1	3	Ordinal
OP-Procedures:	Degree of difference between organisation of ego and alter in preferring sticking to procedures over achieving results	881	2	1	3	Ordinal

continued on the next page ➔

Variable	Description	N	Median	Min.	Max.	Measurement type
OP-IP:	Degree of difference between organisation of ego and alter in importance attached to protecting intellectual property.	897	3	1	3	Ordinal
Same Soc. Sector:	whether or not ego and alter work in the same societal sector (business, government, academia, NGO)	1020	0	0	1	Dichotomous
Cognitive proximity						
CP-Jargon:	Degree of difference in technical terms and jargon used by ego and alter.	1019	4	1	5	Ordinal
CP-Machines:	Degree of difference in specialised instruments, software, machines that ego and alter use.	935	3	1	5	Ordinal
Common Domain:	Dummy to indicate whether or not have a specific field in the water sector as a common expertise.	1020	1	0	1	Dichotomous
Common Activity	Dummy to indicate whether or not respondent and alter have a daily activity in common. The daily activities listed included: management, policymaking, research, maintenance, operations.	1020	1	0	1	Dichotomous

3.3.4 Methodology for constructing a variable per dimension

Most dimensions of proximity are operationalised using a set of items that together measure the score on that dimension. We have used exploratory factor analysis to test whether different items measure a common variable. The results are shown in Table 9.

Table 9 Results of exploratory factor analysis.

Variables	Components				
	1	2	3	4	5
OP-Adapt:	.746				
OP-Management:	.687				
OP-External:	.647				
OP-Procedures:	.609				
OP-IP:	.594				
Same Soc. Sector:	.440				.401
SP-Effort:		.776			
SP-Trust:		.768			
SP-Share:		.672			
CP-Jargon:			.721		
CP-Machines:			.717		
Common Domain:			.689		
Common Activity			.480		
Time:				.671	
Frequency:		.451		-.565	
Private:				.558	
Age difference:				.444	.405
Same Gender:.					.749

Rathenau Instituut

Note: Principal components analysis with an orthogonal rotation (Varimax with Kaiser normalisation), only showing factor loadings over 0.4.

Five factors are distinguished. Factors 1 and 3 contain all items that measure organisational and cognitive proximity respectively, and no other variables have a substantial loading on them. Factor 2 contains the variables that ask about social aspects of the interactions in the collaboration. We have termed this interaction-based social proximity. Factor 4 contains the variables that were constructed as potential sources of social proximity (age differences, having a private relationship, and the time the collaborators know each other). We have termed this identity-based social proximity, because it is based on comparing aspects of personal identity of ego and alter. Factor 5 captures gender differences, but also loads on age differences and being in the same sector or not. This may be related to the distribution of the data. On average the women in the dataset are much younger than the men, which explains the correlation between gender and age differences. Apparently this also relates to having relations in the same

sector or not. All in all there are four strong and clear factors: organisational proximity, cognitive proximity, interaction-based social proximity and identity-based social proximity.

The scores per item were aggregated to produce a single score per variable, thus allowing us to analyse the outcomes per dimension of proximity. This is done by averaging the scores of the different questions in each factor. Some items were measured on a different scale (e.g. dichotomous rather than a 5-point Likert scale). Diverging items were rescaled in order to combine items with different scales in an aggregate variable. For dichotomous variables (for example yes or no, male or female) we assigned the two options a value of 1 and 5 respectively and then included them in the calculation of averages. Sensitivity analysis showed that assigning different values (2 and 4) had no significant influence on the results.

3.3.5 Outcomes

Six different outcomes will be examined (Table 10). They are measured as variables that can take a value of 0 or 1. They include tangible outcomes of collaboration (such as publications), but also for intangible outcomes (such as exchange of knowledge). To enable comparison with outcomes that are not knowledge-related we also included financial turnover as an outcome. To get some more understanding of how collaborations at personal level are brought to collaboration at organisational level, we also included joint programmes as an outcome. Patents, copyrights, and trademarks will be excluded from the analysis, since this item was hardly selected as an outcome. Our survey does not distinguish between achieved and expected outcomes. This implies that some respondents may have indicated results they expect to be realized in the future, while others describe actually achieved results from the past. Of course, the fact that collaborators expect a specific outcome does not imply that this outcome will indeed be realised as expected (see for example Ariño and Doz, 2000). However, by far most relationships in the dataset are well established (almost all alters and egos have known each other for at least a few years). Most outcomes will hence have been realised already, or there is a realistic expectation that they will occur in the (near) future. Moreover, we have no reason to assume that more proximate collaborators have a tendency to report on achieved outcomes while less proximate people would report expected outcomes or vice versa.

Table 10 Description of the outputs analysed.

Outcome	Description	Times selected (out of 1020)
Innovation	Product, process or organisational innovations. No strict definition in survey; interpretation of respondent whether e.g. incremental innovations are included.	361
Joint publications	Scientific papers as well as policy documents and other publications.	299
Shared knowledge	Any form of knowledge exchange.	654
Patents, copyrights, trademarks	Ideas that are protected with a patent, copyright, trademark.	26
Support for ideas	A bit more specific than knowledge exchange: the relation supports ideas of the respondent.	472
Joint programmes	Collaborations at organisational level (joint programmes, projects, collaboration agreements).	632
More financial turnover	Money inflow for the organisation of the respondent.	209

Rathenau Instituut

3.4 Results and Analysis

Table 11 presents the degree of association between outcomes and the indicators of proximity. We use two statistical approaches that match the skewed distribution of values. For each combination of outcome and proximity variables, we first measure the degree of correlation (Kendall's τ). Then, we compare the group of respondents who do report a specific outcome with the group of respondents who do not report the outcome using a Mann-Whitney test. In the table we report r for effect size – Mann-Whitney's Z-score divided by the square-root of N – to overcome the Mann-Whitney test's sensitivity to sample size.

3.4.1 Results per dimension

Geographical proximity has a negative effect on three of the six outcomes: the longer the geographical distance between the two collaborators (ego and alter), the more likely it is that the respondent reports the outcomes innovation, joint publications, or financial turnover. This is in line with Weterings and Ponds (2009), who, also for empirical data on the Netherlands, find that knowledge obtained through non-regional knowledge flows (i.e. flows across larger distances) is valued higher than the knowledge obtained in regional flows. Long-distance collaboration is scarcer than short-distance collaboration, but people are willing to afford higher (transaction) costs and uncertainty if the collaboration will yield valuable outcomes.

Our finding seems to contradict earlier studies on the impact of geographical proximity on publications in other fields than water. (Hoekman et al., 2010 find a positive effect: co-authors tend to be geographically proximate (their study is at European rather than national scale, but they also find that a large share of the scientific collaborations takes place within countries). Also studies on patents (a "hard" outcome too) find a positive effect, such as Wal (2009), who analysed the

Table 11 Associations between the outcomes of collaboration and the indicators of proximity.

	Innovation		Joint publications		Financial turnover		Support for ideas		Collaboration programmes		More shared knowledge	
Indicator	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**
Geographical												
Geodesic distance	-.072**	.088	-.076**	.092	-.068**	.082						
Social												
SP-Effort			.145***	.156	.107***	.115	.125***	.145				
SP-Trust	.069**	.074	.092***	.104			.099***	.106			.103***	.110
SP-Share	.063**	.068	.160***	.173	.113***	.122	.163***	.176	.082***	.088	.080***	.087
Frequency	.069**	.072	.100***	.105	.105***	.111	.092***	.097	.089***	.093		
SP_Average	.071***	.082	.165***	.191	.109***	.126	.176***	.204	.060**	.069	.082***	.095
Age difference							-.056*	.058				
Time									-.046	.051		
Private			.084***	.084			.106***	.106			.153***	.153
SP2_Average							.088***	.102				
Same gender	.108***	.108			.073**	.073						
Organisational												
OP-Adapt					-.123***	.128					.060*	.062
OP-Management	-.107***	.111	-.055*	.057	-.110***	.114	-.065**	.068				
OP-External												
OP-Procedures	-.070**	.072	-.058*	.060	-.123***	.128						
OP-IP	-.059*	.062	-.080**	.084								

continued on the next page ➔

	Innovation		Joint publications		Financial turnover		Support for ideas		Collaboration programmes		More shared knowledge	
Indicator	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**	**Kendall's τ**	**r**
Same soc. sector	-.088***	.088	-.169***	.169	-.103***	.103			.147***	.147	.058*	.058
OP_Average	-.088***	.103	-.087***	.101	-.148***	.172						
Cognitive												
CP-Jargon							.073**	.078	.069**	.074	.150***	.162
CP-Machines	.089***	.098	.067**	.074			.055*	.061	.066**	.073	.170***	.188
Common Domain	.077**	.077	.136***	.136	.123***	.123	.100***	.100	.100***	.100	.109***	.109
Common activity							.111***	.111				
CP_Average	.053*	.063	.091***	.108	.049*	.058	.121***	.142	.103***	.121	.153***	.180

Note: Only significant variables are reported.
*p <= .10.
**p <= .05.
***p <= .01.

biotechnology industry in Germany. However, it is important to keep in mind that such studies measure a different thing. Our analysis finds that out of all sorts of collaborations that people have, the (geographically) more distant ones produce joint publications, innovations and higher turnover. The analyses of patents and publications only observe relations that have actually achieved patents or publications and cannot compare with collaborations in which these outcomes were not realized. Instead, they show that even collaborations with publications and patents occur across smaller distances than we would observe in a world where collaborations are distributed randomly across space. This suggests that geographic proximity has a positive effect on network formation, and a negative effect on specific outcomes.

To confirm this, we have compared the distance between actual collaborators with the distance between any random ego-alter pair in the dataset.[6] A Mann-Whitney test proves that collaborators work across much smaller distances than any random combination of egos and alters in the dataset (Z=-16.069; p=.000; median of actual collaborations is 50.0 km; median of potential collaborations is 75.4 km).

There is a clear difference between "hard" and "soft" outcomes. Geographical proximity only has an effect on hard (i.e. tangible) outcomes: innovations, joint publications, and financial turnover. It has no effect on soft (i.e. intangible) outcomes: support for ideas, collaboration programmes, and more shared knowledge.

For identity-based social proximity we only find a (positive) effect on support for ideas. Gender correlates with the indicators of identity-based social proximity, but does not belong to the same factor and is hence treated separately. It only has a (positive) effect on innovations and turnover. Interaction-based social proximity has a positive effect on all outcomes. Although the operationalisation of social proximity is different, for hard outcomes this is in line with the findings of Broekel and Boschma (2012) and Wal (2009).

Organisational proximity has a significant negative effect on the hard outcomes: innovations, publications, and financial turnover. The aggregated variable has no effect on the soft outcomes and even among the specific items only a few results were found. We are not aware of any earlier literature that finds an effect for organisational proximity on knowledge-related outcomes. Probably the explanation is similar to geographical proximity: most collaborations are with proximate alters (Broekel and Boschma (2012) find a positive effect on network formation in the Dutch aviation industry), but collaborations across larger organisational distances are selected for the likelihood of producing valuable, hard outcomes.

6 The few foreign addresses in the dataset were excluded to avoid biases.

Interestingly, differences in protecting intellectual property are negatively associated with publications and innovations. Apparently, differences in intellectual property regimes do not hinder such outcomes, and may even be necessary for collaboration.

Cognitive proximity has a positive effect on all outcomes. When ego and alter have a common knowledge base, all outcomes are reported more often. As we have seen in the literature overview, the empirical evidence on this point is inconclusive so far. However, our findings corroborate the results of e.g. Cantner and Meder (2007) who, based on German patents, find that cognitive similarity is associated with higher odds on outcomes. The strongest effects are found among the soft outcomes. Using the same jargon is only associated with soft outcomes.

We have also tested to what extent the different outcomes are correlated. The results are shown in Table 12. This confirms the existence of hard and soft outcomes; all hard outcomes are correlated at .10 level, all soft outcomes are correlated at .01 level. The weaker correlations among the hard outcomes seem to be because financial turnover is much less knowledge-intensive than the other outcomes. The strongest associations are between innovations and joint publications and between shared knowledge and support for ideas.

Table 12 Correlations (phi coefficients) between the different outputs of collaboration.

	Innovation	Joint publications	Financial turnover	Shared knowledge	Support for ideas	Joint programmes
Innovation	X	.235***	.056*	.092***	.185***	.035
Joint publications	.235***	X	.068**	.154***	.180***	.150***
Financial turnover	.056*	.068**	X	-.071**	.065**	.013
Shared knowledge	.092***	.154***	-.071**	X	.206***	.167***
Support for ideas	.185***	.180***	.065**	.206***	X	.132***
Joint programmes	.035	.150***	.013	.167***	.132***	X

Rathenau Instituut

*p <= .10.
**p <= .05.
***p <= .01.

Table 13 Odds ratios for the association between dimensions of proximity and hard outputs.

Items per dimensions of proximity	Innovation					Joint publications					Financial turnover				
	Least proximate			Most proximate		Least proximate			Most proximate		Least proximate			Most proximate	
Social (direct)															
SP-Effort						0.19***	0.41**	1	1.30	1.99***	1.41	0.15***	1	0.93	2.17***
SP-Trust	0.54	0.72	1	1.29	1.39	2.08*	0.28**	1	1.39*	1.65**					
SP-Share	1.08	1.28	1	1.61***	1.55***	0.80	0.48	1	1.54**	2.53***	1.22	0.59	1	1.36	2.25***
Same gender															
Social (sources)															
Age difference												0.39*	1	1.43**	1.85**
Time															
Frequency		0.54	1	1.28*	1.17		0.84	1	1.76***	1.29					
Private							0.67***	1							
Organisational															
OP-Adapt												2.85***	1	0.92	
OP-Management		1.01	1	0.61***			0.94	1	0.75*			1.66**	1	0.69**	
OP-External															
OP-Procedures		1.12	1	0.76*			0.84	1	0.72**			2.85***	1	0.92	
OP-IP		1.18	1	0.82			1.86***	1	0.91						
Same soc. sector		1.46***	1				2.23***	1				1.72***	1		
Cognitive															
CP-Jargon															
CP-Machines	0.74	0.75	1	1.27	1.10	0.68	1.10	1	1.19	1.33					
Common Domain		0.69***	1				0.49***	1				0.47***	1		
Common activity															

*p <= .10.
**p <= .05.
***p <= .01.

Table 14 Odds ratios for the association between dimensions of proximity and soft outputs.

Items per dimensions of proximity	Innovation					Joint publications					Financial turnover				
	Least proximate			Most proximate		Least proximate			Most proximate		Least proximate			Most proximate	
Social (direct)															
SP-Effort	0.59	0.81	1	1.46**	1.87**										
SP-Trust	0.52	0.59	1	1.32	1.56**						0.32**	0.70	1	1.21	1.65**
SP-Share	0.97	0.64	1	1.86***	2.44***	0.50**	0.50**	1	1.01	1.24	0.58	1.28	1	1.35*	1.52**
Same gender															
Social (sources)															
Age difference		1.20	1	0.82											
Time						1.07	1.17	1	0.99	0.83					
Frequency		0.68	1	1.56***	1.14		0.61	1	1.35**	1.39					
Private		0.63***	1									0.48***	1		
Organisational															
OP-Adapt												0.94	1	1.29*	
OP-Management		0.77	1	0.67***											
OP-External															
OP-Procedures															
OP-IP															
Same soc. sector							0.52***	1				0.77*	1		
Cognitive															
CP-Jargon	0.54	1.01	1	1.32	1.49**	0.23**	0.79	1	0.93	1.37	2.15	0.89	1	1.42**	2.76***
CP-Machines	0.83	0.85	1	1.17	1.01	0.73	0.97	1	1.15	1.21	0.51***	0.72	1	1.44*	1.89*
Common Domain		0.64***	1				0.64***	1				0.61***	1		
Common activity		0.63***	1												

*p <= .10.
**p <= .05.
***p <= .01.

3.4.2 The inverted U-shaped curve of proximity

Earlier literature suggests that the relation between proximity and outcomes is not linear but has the shape of an inverted U-curve: it is better to be closer (or more similar) than very far apart (or very different) but being too close or too similar also has a negative effect on outcomes (Boschma, 2005). This is assessed by calculating odds ratios.[7]

The odds ratios are shown in Table 13 and Table 14. The odds ratios are computed only for the items for which significant results were reported in Table 11. We have not computed ratios for the aggregated variable for each dimension, as it proved to be very complicated to construct an aggregate variable in such a way that the shape of the curves can be analysed.

No odds ratios have been calculated for geographical proximity, as it is a continuous variable. Instead we have computed the values of the median and quartiles of the groups that do or do not report an outcome. This confirms the negative effect of geographical proximity, yet does not suggest an inverted U-shape.

All odds ratios suggest a linear pattern. This is line with most of the empirical literature, that does not report inverted U-curves.[8] There are three possible explanations for the absence of inverted U-curves: (1) respondents report outcomes achieved in a time when they were less proximate to their alters, but they have since become more proximate; (2) the collaborators are all relatively proximate, especially in a geographical (all within the Netherlands) and cognitive (all within the water sector) sense; maybe they are all relatively in such close proximity that we cannot find an inverted U-curve or (3) the optimal level of proximity is far more proximate than the literature suggests and the downward sloping part of the curve is beyond our measurement scale.

3.4.3 Interaction effects between the different dimensions

The literature suggests that the different dimensions of proximity may complement or substitute each other (e.g. Breschi and Lissoni 2003; Broekel and Boschma 2012). We have applied multivariate logistic regression to quantify the interactions among the dimensions of proximity. The results are shown in Table 15.

---

7 We use Pearson's Chi Square test to determine whether an odds ratio is significantly different from its neutral value and, hence, whether there is an actual effect. For variables that can only take two values, we have corrected with Yates' Continuity Correction. Pearson's Chi Square may overestimate the effect, because it (incorrectly) assumes that the discrete probability of observed binomial frequencies in the table can be approximated by the continuous chi-squared distribution. The correction subtracts 0.5 from each difference between observed and expected value, leading to higher p-values.

8 An exception is the work by Nooteboom, et al. (2007), who show an inverted U-curve for cognitive proximity in explorative patents.

Table 15 Multivariate logistic regression of the dimensions of proximity on the outcomes of collaboration.

N=402	Innovation			Joint publications			Financial turnover		
	B	SE	Exp (B) (p)	B	SE	Exp (B) (p)	B	SE	Exp (B) (p)
GP	-3.257	1.970	0.039*	-1.613	1.984	.199	-4.249	2.717	.014
SP_Average	.347	.185	1.415*	.627	1.98	1.872***	.803	.249	2.232***
SP2_Average	.101	.111	1.107	.004	.116	1.004	-.114	.147	.892
OP_Average	-.587	.235	.556**	-.854	.249	.426***	-1.513	.321	.220***
CP_Average	.238	.118	1.268**	.277	.124	1.319**	.087	.151	1.091
Constant	-1.747	1.045	.174*	-2.569	1.105	.077**	-1.643	1.358	.193
Goodness of fit									
-2LL	518.679			484.543			345.651		
X^2	18.29***			27.21***			38.28***		
Cox-Snell R^2	.044			.065			.091		
Nagelkerke R^2	.060			.091			.148		
	Shared knowledge			Support for ideas			Collaboration programmes		
	B	SE	Exp (B) (p)	B	SE	Exp (B) (p)			
GP	-.559	1.909	.572	1.873	1.858	6.507	1.551	1.966	4.718
SP_Average	.039	.194	1.039	.686	.187	1.986***	.308	.187	1.361*
SP2_Average	.181	.119	1.199	.100	.109	1.105	-.176	.113	.838
OP_Average	.094	.247	1.099	-.608	.233	.544***	-.145	.237	.865
CP_Average	.431	.121	1.539***	.307	.116	1.360***	.277	.117	1.319**
Constant	-1.620	1.087	.198	-3.161	1.047	.042***	-.859	1.046	.424
Goodness of fit									
-2LL	475.042			527.202			503.527		
X^2	18.44***			29.73***			10.96**		
Cox-Snell R^2	.045			.071			.027		
Nagelkerke R^2	.063			.095			.037		

*p <= .10.
**p <= .05.
***p <= .01.

The multivariate regression shows that the effect of geographical proximity on the hard outcomes is much smaller (and indeed in two of the three cases insignificant) when controlling for the other three dimensions. This seems to be in line with Ponds, Oort and Frenken (2007) who find with Dutch publication data that the effect of geographical proximity is smaller if controlling for organisational differences. For the soft outcomes, the effect of social proximity becomes less significant, and in the case of shared knowledge even insignificant. Support for ideas shows a significant effect for organisational proximity, which it did not in the bivariate analysis.

Multivariate analysis hence proves that there are indeed interaction effects among dimensions of proximity, where a lack of proximity in one dimension can be bridged by proximity in other dimensions.

3.4.4 Soft versus hard outcomes

There is a remarkable difference between what we have termed the "soft" or intangible outcome (shared knowledge, support for ideas, collaboration programmes) and the "hard" or tangible outcomes (innovations, publications, financial turnover) of a relation. The dimensions of proximity have a different effect on hard outcomes and soft outcomes. Geographical proximity has a negative association with hard outcomes, but no association with soft outcomes. The same goes for organisational proximity. Jargon, an indicator of cognitive proximity, has a positive effect on all soft outcomes and no effect on hard outcomes.

Our expectation was that would be relatively few distant relationships that are only established if the collaborators expect clear pay-offs in the form of hard outcomes, and that proximate relations are more common and involve more informal knowledge sharing with soft outcomes. This clearly holds for geographical and organisational proximity: most relationships are relatively proximate, but the odds of producing hard outcomes are higher for distant relations than for more proximate relations. This finding matches the result of Arundel and Geuna (2004) who found that European firms that stress the importance of informal contacts to learn about public research results attach lower value to the geographical proximity of the provider of these results. Our result also seems in line with what Ibert (2010) terms relational distance in a case study of one innovation at the intersection of science and business. The (socio)cultural tensions that can come with geographical and organisational distance may be conducive to hard outcomes like innovation. The statistical relationship is different for social proximity. This might be explained by the fact that relations that involve hard, tangible outcomes probably require social proximity and mutual trust to assure the collaborators that collaboration will prove useful and is worth the investment. More common, closer relationships may involve more face-to-face contact and build up social proximity through daily interactions. More or less the same seems to hold for cognitive proximity: both soft and hard outcomes appear to require a relatively high level of cognitive proximity.

A mix of both proximate and distant relationships appears to be optimal for the production and exchange of knowledge-related outcomes in collaboration. This is in line with the work of Uzzi (1997) on overembeddedness, who (for social proximity) also recommends a mix of relationships. The results in this way corroborate the suggestion of a proximity paradox, where being proximate is considered conducive to network formation, yet has a negative impact on innovative performance (Broekel and Boschma, 2012).

In addition to that, Cantwell and Santangelo (2002) suggest that actors who are very proximate in one dimension should avoid being proximate in others. They find that cognitively very proximate firms are very reluctant to co-locate. In the literature on related variety it has also been suggested that the negative impact of very high proximity in one dimension could be counterbalanced by a lower proximity in other dimensions (Boschma and Frenken, 2010).

3.4.5 Most relations are proximate

Our dataset appears to contain more proximate relationships than distant relationships. This may be partly explained by self-selection. Although we asked respondents explicitly to randomly select three of their professional relationships, it is not unlikely that many respondents focused on socially proximate relationships. This may be deliberate (for example because of privacy issues) or accidental (because socially proximate collaborators simply came to mind earlier when filling in the survey).

An alternative possibility is that respondents only report about proximate relations because their entire network consists of relatively proximate people. This would suggest that the entire Dutch water sector consists of cliques of people who are proximate in all four dimensions. Potential other collaborators (even within the Dutch water sector) may remain out of sight. Such a situation can be very risky in the longer term. Drejer and Vinding (2007) for example show that firms with a limited absorptive capacity in sparsely populated regions also tend to collaborate with domestic partners rather than looking abroad. Such behaviour may lead to group-think and can hamper the creation of new knowledge, because the existing knowledge of all people in a clique is already very similar.

3.5 Conclusions and discussion

3.5.1 Conclusions

Our analysis clearly shows that proximity matters for the outcomes that people report from collaborations with other professionals in their sector. We have also found that the effects of proximity vary by dimension of proximity and by outcomes. There is a difference between "hard" outcomes (innovations, publications, financial turnover) and "soft" outcomes (shared knowledge, collaboration programmes, and support for ideas). Both geographical proximity and organisational proximity have a negative association with the hard outcomes, and no

association with the soft outcomes. Social and cognitive proximity have a positive effect on all six outcomes.

We have also shown that there are interaction effects between the different dimensions of proximity. In particular, the effect of geographical proximity becomes much weaker when controlling for the other dimensions for the hard outcomes. Also, the effect of social proximity becomes weaker for the soft outcomes.

Our empirical analysis does not reveal the inverted U-curves suggested by the literature. The patterns are generally linear, either in a positive or negative direction.

It is important to note that we do not assess the effectiveness of collaboration. Some people in the dataset may have only one outcome of collaboration (say joint publications), and yet consider their collaboration highly effective, because they only look for this specific outcome. The results should hence not be interpreted in terms of effective collaboration. Our model reflects how proximity relates to different outcomes of collaboration.

3.5.2 Discussion

Our empirical case is the Dutch water sector. The effect of proximity may be specific to a country and to the specific configuration and infrastructure of a sector. Caution is needed if our findings are generalized to more generic situations. Proximity may, for example, work differently in geographically larger areas or in regions with more institutional diversity.

As we have explained in section 3.3.1, one of the peculiarities of the Dutch water sector is that the service-providers (drinking water suppliers and wastewater treatment plants) all have their own geographically discrete service areas and hence do not face any direct competition pressure. This may affect the role of the proximity dimensions for the employees of these organisations. It will probably be easier to build up social proximity with people from other service providers as the levels of trust will be higher than if they were actual competitors.

On the other hand, geographical proximity will always be lower between, for example, two water suppliers, as there is per definition just one supplier in each region. However, we do not expect that this phenomenon had a large effect on our findings. The sector consists of many organisations, in a wide range of environments, from regional authorities (non-competitive) to consultants (highly competitive). Moreover, as the results on organisational proximity show, many collaborations exist across different organisation types. The high share of people from environments with a low level of competition (authorities, NGO's, etc) may contribute to the high scores on social proximity.

In addition, the existence of a dense network with many heterogeneous players in a relatively small country may lead to economies of scale and network exter-

nalities: the more players there are in a network, the more valuable it is for entrants to become well embedded in the network. In addition to this, the sector is strongly organised with for example network organisations and structures like the regional division of water suppliers. Although this may influence the fact that many collaborators are proximate to each other, it is not very likely that it also influences the relation between proximity and the outcomes of collaboration.

More empirical work is needed to compare different sectors and different countries or regions. An interesting question is whether or not the different dimensions of proximity can complement or substitute each other. Of course, collaboration is driven by a far more complicated interplay of factors than we have tested in this article. For example, personal characteristics also determine the outcomes of collaboration. Further research should examine how the entire complex of factors (including the various dimensions of proximity) create patterns of collaboration.

3.5.3 Methodology

The use of survey data has clear benefits: it allows for more refined indicators of proximity and for the analysis of a broader range of outcomes. However, it also introduces potential measurement problems. First, all questions on the relation between alter and ego have only been answered by the egos (the respondents). It is hence their perception of the relation that we measure. Some indicators of proximity (such as the city of work of ego and alter) are not very susceptible to differences in perception, but others, such as the indicators for social and cognitive proximity may be perceived differently by alter and ego. Because we do not use a closed network (egos are free to select alters outside the network of invited respondents), and because it was not required to fill in the names of the alters, we cannot check if there are "mirroring" responses or how diverging they are. The effect on our findings is probably very small, as we have no reason to assume that the alters systematically have different perceptions on the collaborations than the egos. Moreover, the perceived proximity to a (potential) collaborator will have more impact on the collaboration decisions of an ego than the "actual", objectified proximity (insofar as that can be measured at all). Second, proximity is dynamic and accumulates over time. This holds especially for social and cognitive proximity. For example, the very fact that an alter and ego publish a report together may increase their cognitive and social proximity. This implies that the direction of the causality between proximity and outcomes is not straightforward. The realisation of the outcomes may have caused collaborators to become more proximate. In fact, this is exactly the assumption of most studies that use patent or publication datasets: earlier co-patents or co-publications are assumed to indicate proximity. Future research should address this dynamic character of proximity. In that respect, it would also be good to not only make an explicit distinction between achieved and expected outcomes in the future, but also to monitor whether expectations regarding outcomes that have not come true in turn also have a reverse impact on the perceived proximity between actors.

3.5.4 Policy recommendations

Our analysis provides fruitful insights for future policy design. We elaborate on two of them. First, research policy should take the difference between 'hard' and 'soft' outcomes into account. Many research policy instruments steer specifically at some proximity dimensions. EU policy, for example, promotes the emergence of a European Research Area, where knowledge can flow without hindrance of geographical borders, and many national research programmes have specific incentives for collaborations between research organisations and firms. However, as our analysis shows, the dimensions of proximity work differently for different outcomes of collaboration. For fruitful policy design it is hence useful to first determine what kind of outcomes are to be stimulated exactly, and then per dimension of proximity develop incentives to promote collaborations with high or low proximity.

Second, despite popular belief that geographical proximity will promote fruitful collaboration (which is often the basic premise behind policy to create for example science parks), our analysis shows that although indeed many people tend to have geographically proximate collaborations, the more distant collaborations result more often in publications and innovations. That effect becomes smaller if one controls for the other dimensions of proximity. This suggests that initiatives like science parks are probably only effective (in producing more publications and innovations) if they bring together people that would collaborate anyway but would otherwise have to travel long distances to meet.

3.6 References

Aguiléra, A., V. Lethiais & A. Rallet. 2012. Spatial and Non-spatial Proximities in Inter- firm Relations: An Empirical Analysis. *Industry and Innovation* no. 19(3): 187-202.

Aldieri, L. 2011. Technological and geographical proximity effects on knowledge spillovers: evidence from the US patent citations. *Economics of Innovation and New Technology* no. 20(6): 597-607.

Aldieri, L. & M. Cincera. 2009. Geographic and technological R&D spillovers within the triad: Micro evidence from US patents. *The Journal of Technology Transfer* no. 34(2): 196-211.

Ariño, A. & Y. Doz. 2000. Rescuing troubled alliances… before it's too late. *European Management Journal* no. 18(2): 173-182.

Arundel, A. & A. Geuna. 2004. Proximity and the use of public science by innovative European firms. *Economics of Innovation and new Technology* no. 13 (6): 559-580.

Ashkanasy, N.M., C.P.M. Wilderom, and M. F. Peterson. 2000. *Handbook of organizational culture and climate*: Sage.

Audretsch, D. B. & M. P. Feldman. 1996. R&D spillovers and the geography of innovation and production. *The American Economic Review* no. 86(3): 630-640.

Autant-Bernard, C., P. Billand, D. Frachisse & N. Massard. 2007. Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies*. *Papers in Regional Science* no. 86(3): 495-519.

Autio, E., A.P. Hameri & O. Vuola. 2004. A framework of industrial knowledge spillovers in big-science centers. *Research Policy* no. 33(1): 107-126.

Balland, P.A. 2012. Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies* no. 46(6): 741-756.

Balland, P.A., R. Suire & J. Vicente. 2013. Structural and geographical patterns of knowledge networks in emerging technological standards: evidence from the European GNSS industry. Economics of Innovation and New Technology no. 22(1): 47-72.

Boschma, R. 2005. Proximity and innovation: a critical assessment. Regional Studies no. 39(1): 61-74. Boschma, R. & K. Frenken. 2010. The spatial evolution of innovation networks. A proximity perspective. *Handbook of evolutionary economic geography*: 120-135.

Bouba-Olga, O.r, M. Ferru & D. Pépin. 2012. Exploring spatial features of science-industry partnerships: A study on French data. *Papers in Regional Science* no. 91(2): 355-375.

Breschi, S. & F. Lissoni. 2003. Mobility and social networks: Localised knowledge-spillovers revisited. Milan: *University Bocconi, CESPRI Working Paper* (142).

Broekel, T. & R. Boschma. 2012. Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography* no. 12(2): 409-433.

Brugge, R. 2009. *Transition dynamics in social-ecological systems: the case of Dutch water management*: Erasmus University Rotterdam.

Cantner, U. & A. Meder. 2007. Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination* no. 2(1): 45-65.

Cantwell, J. & G. D. Santangelo. 2002. The new geography of corporate research in Information and Communications Technology (ICT). *Journal of Evolutionary Economics* no. 12(1): 163-197.

Coenen, L., J. Moodysson & B. T. Asheim. 2004. Nodes, networks and proximities: on the knowledge dynamics of the Medicon Valley biotech cluster. *European Planning Studies* no. 12 (7): 1003-1018.

Cunningham, S. W. & C. Werker. 2012. Proximity and collaboration in European nanotechnology. *Papers in Regional Science* no. 91(4): 723-742.

Delobbe, N., R. R. Haccoun & Chr. Vandenberghe. 2002. Measuring core dimensions of organizational culture: A review of research and development of a new instrument. *Unpublished manuscript, Universite catholique de Louvain, Belgium*.

Denison, D.R. & A.K. Mishra. 1995. Toward a theory of organizational culture and effectiveness. *Organization Science*: 204-223.

Drejer, I. & A. L. Vinding. 2007. Searching near and far: determinants of innovative firms' propensity to collaborate across geographical distance. *Industry and Innovation* no. 14(3): 259-275.

Ferru, M. 2010. Formation process and geography of science-industry partnerships: the case of the University of Poitiers. *Industry and Innovation* no. 17(6): 531-549.

Fleming, L., Ch. King & A.I. Juda. 2007. Small worlds and regional innovation. *Organization Science* no. 18(6): 938-954.

Hagedoorn, J., A. N. Link & N. S. Vonortas. 2000. Research partnerships. *Research Policy* no. 29(4): 567-586.

Hoekman, J., K. Frenken & R.J.W. Tijssen. 2010. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy* no. 39(5): 662-673.

Hofstede, G. 1998. Identifying organizational subcultures: An empirical approach. *Journal of Management Studies* no. 35(1): 1-12.

Ibert, O. 2010. Relational distance: sociocultural and time–spatial tensions in innovation practices. *Environment and planning. A* no. 42(1): 187.

Kabat, P.l, W. van Vierssen, J. Veraart, P. Vellinga & J. Aerts. 2005. Climate proofing the Netherlands. *Nature London* no. 438 (7066): 283.

Katz, J.S. & B.R. Martin. 1997. What is research collaboration? *Research Policy* no. 26(1):1-18.

Knoben, J. & L.A.G. Oerlemans. 2006. Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews* no. 8(2): 71-89.

Levin, D.Z. & R. Cross. 2004. The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*: 1477-1490.

Mattes, J. 2012. Dimensions of proximity and knowledge bases: Innovation between spatial and non-spatial factors. *Regional Studies* no. 46(8): 1085-1099.

McAllister, D.J. 1995. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of management journal*: 24-59.

Metcalfe, S. 1994. The economic foundations of technology policy: equilibrium and evolutionary perspectives. In *The Handbook of Industrial Innovation*, edited by Mark Dodgson and Roy Rothwell, 409-512. Cheltenham: Edward Elgar.

Moel, P. J. de, J. Q.J.C. Verberk & J.C. van Dijk. 2006. *Drinking water: principles and practices*: World Scientific.

Muizer, A. & M. van den Bergh. 2002. Strategic watercards: International opportunities for the Dutch water sector. Delft: Netherlands Water Partnership.

Nooteboom, B. 1999. Innovation and inter-firm linkages: new implications for policy. *Research Policy* no. 28(8): 793-805.

Nooteboom, B., W. van Haverbeke, G. Duysters, V. Gilsing & A. van den Oord. 2007. Optimal cognitive distance and absorptive capacity. *Research Policy* no. 36(7): 1016-1034.

North, D. C. 1991. Institutions. *The Journal of Economic Perspectives* no. 5(1): 97-112.

Ponds, R., F. van Oort & K. Frenken. 2007. The geographical and institutional proximity of research collaboration*. *Papers in Regional Science* no. 86(3): 423-443.

Porter, M. E. 2000. Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* no. 14(1): 15-34.

Rallet, A. & A. Torre. 1999. Is geographical proximity necessary in the innovation networks in the era of global economy? *GeoJournal* no. 49(4): 373-380.

Rockström, J., W. Steffen, K. Noone et al. 2009. Planetary boundaries: exploring the safe operating space for humanity. *Ecology and society* no. 14(2).

Schamp, E. W, B. Rentmeister & V. Lo. 2004. Dimensions of proximity in knowledge- based networks: the cases of investment banking and automobile design. *European Planning Studies* no. 12(5): 607-624.

Sinnott, R.W. 1984. Virtues of the Haversine. *Sky and telescope* no. 68: 158.

Stumpe, J. 2011. Bestuursakkoord Water. The Hague: Ministry of Infrastructure and Environment. Thomas, Duncan A, and Roger R Ford. 2005. *The crisis of innovation in water and wastewater*: Edward Elgar Publishing.

Uzzi, B. 1997. Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*: 35-67.

Vierssen, W. van. 2012. *Water@ work!* Inaugural address. January 18, 2012.

Vinciguerra, S., K. Frenken, J.Hoekman & F. van Oort. 2011. European infrastructure networks and regional innovation in science-based technologies. *Economics of Innovation and New Technology* no. 20(5): 517-537.

Wal, A. L. J. ter. 2009. *The Structure and Dynamics of Knowledge Networks: a Proximity Approach*. PhD Thesis, Utrecht University, Utrecht.

Weterings, A. & R. Ponds. 2009. Do Regional and Non-regional Knowledge Flows Differ? An Empirical Study on Clustered Firms in the Dutch Life Sciences and Computing Services Industry. *Industry and Innovation* no. 16 (1): 11-31.

4 The dynamics of the European water research network: A diversifying community with a stable centre[9]

4.1 Introduction

Water research has great potential to contribute to societal challenges. The new framework programme for research in the European Union, Horizon 2020, addresses several grand challenges for Europe that are essentially intertwined with water-related research: 'Climate action, resource efficiency and raw materials', and 'Food security, sustainable agriculture, marine and maritime research and the bio-economy'.

To tackle these societal challenges, both researchers in water management and in research policy have called for an integral approach. Specifically in water management, the paradigm has shifted from the government as an exclusive authority in managing resources to a multi-stakeholder approach where many stakeholders, with different institutional backgrounds, all contribute to the management of a resource. For knowledge production and use, this implies that participatory approaches are needed, where different actors together develop new knowledge (Pahl-Wostl, 2007; Frijns et al., 2013). More in general, literature on knowledge production suggests that collaboration among different institutional partners drives innovation (Etzkowitz and Leydesdorff, 2000). Inspired by notions such as Triple Helix and Mode 2 knowledge production, both national and international research policies promote collaboration among organisations of different institutional backgrounds (Nieminen and Kaukonen, 2001; Potì and Reale, 2007). Collaboration in networks is hence essential to solve the water-related societal challenges.

Although knowledge about water often is specific to local (environmental) conditions, the challenges in water typically cut across administrative and cultural borders. Many countries struggle with similar issues, while the actors are often still organised in national systems (EIP, 2014; Thomas and Ford, 2005). Mutual learning requires international collaboration in knowledge production, which could benefit from insight in the functioning of international research networks. However, our understanding of research networks in the water sector is still limited. Existing literature (Breschi and Cusmano, 2004; Heller-Schuh et al., 2011) has provided insight into the features of the European research network at a

9 This chapter has been submitted - in slightly different form - to *Water Policy*.

generic level. Given the large differences between network dynamics across scientific fields and economic sectors, however, these generic analyses have limited value when it comes to understanding the composition of the network in a specific research field such as water research. The need for a better understanding of such collaboration networks is twofold. First, (research) policy makers need deeper insights in the functioning and configuration of the networks to design effective funding instruments. Second, organisations in the water field can benefit from network analysis which provides strategic insights in their positioning in the networks. This helps them in fulfilling the shift to an adaptive and integral approach of water management.

The aim of this chapter is to contribute to the understanding of the European water research network by an analysis of the participation of different actors in European Framework Programmes (FPs). What types of organisations are most central in the network, and what is their geographical distribution? We will enrich our analysis by a comparison between the water research network and the generic network constructed from all FP projects.

To enhance our understanding of the role of different organisations, such as businesses, universities and governments, this chapter specifically addresses the centrality of different types of organisations in research networks. Organisations become part of a research network by engaging in collaborative activities with others. Centrality refers to the position of individual organisations in a network, in terms of the number of partners organisations have, and the extent to which they are a crucial link in the network to connect other organisations (e.g. Newman, 2004). It hence indicates which organisations in the network have the best access to other organisations in the network.
Our two research questions are as follows:

- How has the composition of the European water research network developed, in terms of participation of different types of organisation and geographical distribution?
- How can the high centrality of certain actors in the water research network be explained?

Access to different sets of data on the level of countries allows us to make more in-depth analyses of the geographical distribution of the network. This will hence also be our focal point in addressing the first question.

The rest of this article is organised as follows: in section 4.2 we give an overview of existing literature and present our conceptual framework. In section 4.3 we explain how we collected and processed our data and what methodology we used to analyse it. In section 4.4 we present our findings and results. In section 4.5 we present our conclusions and discuss implications for future research and policy.

4.2 Theoretical framework

In addressing the questions set out above, our study builds on several bodies of existing literature. We elaborate below on the three core elements of our research questions. First, we briefly survey existing literature on the analysis of international research networks. Second, we describe earlier findings regarding the participation of different types of organisations in research networks. Third, we give a brief summary of the literature on the development of research networks in geographical terms.

International research networks

Many studies have analysed characteristics of research collaboration networks, including the networks that emerge through FP funding. Some studies focus on changes in the knowledge flows between regions (Scherngell and Lata, 2012; Foddi and Usai, 2013).

Other studies analyse the network at the participant level. They suggest that collaboration is facilitated by prior acquaintance, thematic proximity and geographical proximity (Paier and Scherngell, 2011). Social distance seems to be more decisive than geographical distance (Autant-Bernard et al., 2007). However, a study on project proposals did not observe any significant tendency by organisations to preserve the same consortium across projects or to form a consortium with organisations of the same type (Hazir and Autant-Bernard, 2012). In the R&D literature it is assumed that the benefits of bilateral collaboration also stem from individual characteristics of the actors involved. Studies of the individual characteristics of nodes in a research network have shown that a company's research potential and absorptive capacity promote the probability of collaboration, while small size has a negative impact (Autant-Bernard et al., 2007).

An important question at the participant level is how well different types of actors are embedded in the network. A common notion to assess the position of an actor in the network is centrality. The concept can be operationalised in such a way that it accounts both for the access an actor has to other players, and for how often an actor forms an essential link between to other players. Earlier research has identified some relationships between the centrality of actors and their institutional characteristics. In general organisations in higher education and public research tend to take the most central positions in a research network, while SMEs have weaker positions (Protogerou et al., 2010). However, the specific roles and positions of different organisation types vary strongly across research themes. For example, in aerospace companies have most central positions, while in the life sciences universities have very central positions (Heller-Schuh et al., 2011). It has also been shown that the geographical position of an actor can be of large influence to the centrality in the research system (Foddi and Usai, 2013).

Participation of different types of organisations

The literature suggests that research networks have an increasingly heterogeneous composition in terms of the types of organisation involved. Under the heading 'Mode 2 knowledge production' it has been claimed that the organisational diversity of research tends to increase (Gibbons et al., 1994), which means that a larger variety of organisations cooperate in the production of knowledge. Studies using the 'Triple Helix' framework have also reported that the interactions among governments, universities and industry are intensifying (Etzkowitz and Leydesdorff, 2000). Both bodies of literature suggest that the intensifying interactions are important drivers of innovation and deserve to be stimulated and supported by public policies. Indeed, over the past few decades policies for research and innovation have provided incentives for universities to strengthen their collaboration with industry, the government and other organisations (Nieminen and Kaukonen, 2001; Potì and Reale, 2007).

Empirical evidence suggests that the interactions between different types of organisations collectively embarking on research activities are indeed intensifying (Geiger and Sa, 2008; Hicks and Katz, 1996). However, little is known about the position businesses and governmental organisations adopt in heterogeneous research networks. Institutional characteristics, such as the routines, aims and incentives of an organisation, are expected to influence centrality of an organisation in a network (Owen-Smith and Powell, 2004). Because knowledge production is not a primary task of non-academic organisations, one may wonder whether they will ever become central in research networks, as equal partners with universities or public research organisations. The water sector involves actors from a broad range of institutional backgrounds (EIP, 2014), but it is unclear what that implies for the representation and centrality of actors in the research and knowledge production network.

Geographical distribution

The development of international research networks has been studied not only in terms of organisation types but – particularly in Europe – also in geographical terms. A common notion used to refer to the increasing integration of national research activities is Europeanisation (Barré et al., 2013). In this chapter we focus on the behaviour of organisations, more specifically organisations that play a role in collaborative knowledge production and research. Concerns about the fragmentation and compartmentalisation of national research efforts led in 2000 to a political desire for a European Research Area (ERA). The main instruments used to accomplish this are the funds of the Framework Programmes (Breschi and Cusmano, 2004). These programmes promote the development of research networks across Europe. Since FP5 especially, they have included explicit incentives to integrate research in Europe and to transcend geographical borders (DeLanghe et al., 2009).

The empirical evidence about the effects of these policies on international

research networks is ambivalent. Some studies have found evidence for the development of a European Research Area. Geographical factors (physical distance, territorial borders, language areas) have less impact on collaboration patterns in FP projects than on the collaboration patterns that emerge from co-patenting (Lata et al., 2012). Once the collaborative links are established through FP projects, they tend to continue after FP funding has ceased (Defazio et al., 2009; Hoekman et al., 2013). Recent decades have witnessed a slight decrease in the importance of territorial borders within Europe (Scherngell and Lata, 2012). Still, there is no evidence of large structural changes in the collaboration patterns between member states which would indicate further Europeanisation of research networks (Hoekman et al., 2010; Chessa et al., 2013). Also for water research in particular, it has been claimed that the research is still largely organised in national systems (EIP, 2014; Thomas and Ford, 2005). The creation of a European Research Area is also still incomplete in the sense that there are inequalities in the participation of different member states of the European Union in the European research network. The participation rate among the new member states in Framework Programme projects, for example, lags behind that of the older member states (Annerberg et al., 2010).

4.3 Data and Methodology

4.3.1 Data selection

This study is based on an analysis of data on participation in projects under the European Framework Programmes (FPs). Research networks are often analysed on the basis of research outputs, such as scientific papers, patents (Lata et al., 2012; Chessa et al., 2013) and/or survey data (Weterings and Ponds, 2009; Ferru, 2010). Data on the collaborative links in projects funded by the FPs are a promising alternative source of data for studying patterns in research networks (Heller-Schuh et al., 2011). The research networks emerging from the FPs are well suited to answer questions about the relative impact of the country of origin and organisation type of an actor on its centrality in a research network in relation to the funding it receives, because the FPs explicitly aim both to promote Europeanisation and to involve actors with different organisational backgrounds. For our analysis we used the EUPRO database , which contains a cleaned and harmonised version of the data about FP projects that is publicly available through the information service CORDIS (Barber et al., 2008). EUPRO is produced by the Austrian Institute of Technology (AIT). Based on a set of selection keywords we used all water-related projects in the EUPRO database from FP1 to FP7 whose latest updates were in March 2010. Since FP7 continued for several more years, our analysis did not allow us to draw conclusions about trends that might have occurred in the last few years of FP7, such as an increasing emphasis on innovation and a weaker emphasis on the inclusion of new member states.

In order to construct a database that contains all FP projects on 'the water sector', we developed a set of keywords that filtered out the relevant projects.

In this study we define the water sector as consisting of all human activities associated with the water cycle: production/purification and transport of drinking water, collection, transport and treatment of wastewater, water storage, water use and water management, including flood protection. We excluded water-borne transport, maritime and off-shore activities, oceanography, coastal research and fisheries. Although the boundaries between these fields are sometimes blurred, they are generally separate communities with their own actors and knowledge disciplines. Starting from this definition, we have selected relevant keywords. In doing so, it is important to find a good balance between precision and recall, making sure most selected projects are relevant but also taking care not to exclude too many relevant projects. We took a three-stage-approach: we first used a very broad set of keywords to extract all potentially relevant projects, see appendix A. Second, we employed a more refined set of keywords – building on Wen et al. (2011) – to filter out all false positives. To find additional terms related to trends and policy contexts, we consulted several policy documents, such as the Strategic Research Agenda of the Water Supply and Sanitation Technology Platform, the Strategic Implementation Plan of the European Innovation Partnership on Water, and EU Framework directives related to water. In the third step, we consulted several water experts from across Europe with experience of FP projects to add more keywords, with special attention to keywords typically used in the 1980s and 1990s. We have tested all potential keywords with random samples of projects to explore what kind of projects were extracted from the database. The final keyword set is in appendix A.

Data on scientific publications and R&D budgets

Besides the data on projects in the FPs, we used other datasets to explore explanations on the distribution of projects over countries. We extracted data from the Web of Science (WoS) about publications in scientific journals in the period 2006-2008. We selected publications on water research with the help of keywords, similar to (though less extensive than) the set of keywords used for the main data in this article. We also extracted data on the R&D budget per country in 2011 from the European statistics office, Eurostat.

4.3.2 Methodology

The network is based on participation in FP projects. Joint participation in a project is the basis for having a link in the network. The network is hence an affiliation network, with information of subsets of actors (the organisations) that participate in the same event (an FP project). We have constructed the network as a unipartite structure of organisations, linked by undirected edges. This is done because we are essentially interested in the role of organisations, not of the projects (cf. Protogerou et al., 2010). All partners in a project are assumed to collaborate with each other. If an organisation participates in more than one project, the collaborators in one project become (indirectly) linked to the collaborators in the other project. To get more insight into the network and its constituents, we first generated a few basic statistics of the network, which are

listed in Table 16 (see for example Newman, 2004 for a broader discussion on these statistics). We then extended the results with both more elaborate statistical counts and comparisons with data on scientific publications and R&D budgets and more in-depth social network analysis using the concept of centrality. To enrich our interpretation of the results, we have consulted experts from state funding agency AgentschapNL (now RVO.nl), research council NWO, KWR Watercycle Research Institute, consultancy Evers+Manders, and European technology platform WSSTP.

Table 16 Basic statistics for network description.

Statistic	Definition
Size of the network	Number of unique organisations (nodes) in the network
Average degree	Average number of unique organisations an organisation collaborated with
Average weighted degree	Average number of unique organisations an organisation collaborated with, weighted for collaboration in more than one project.
Largest component	Largest connected group of organisations in the network
Average distance	Mean node-node distance between connected individuals in the network
Largest distance	Maximum node-node distance between connected individuals in the network
Average clustering coefficient	Mean probability that if A has a collaboration with B and with C, then B and C also have a collaboration.
Density	Ratio of actual number of edges to the number of possible edges.
Modularity	Measure of the density of links inside a community compared to the links between communities (where a community is a set of highly interconnected nodes).

Rathenau Instituut

Operationalising centrality

To assess the centrality of organisations in the network, we used two indicators from social network analysis. First, weighted degree centrality counts all the ties (edges) a participant has, taking into account how often two actors have collaborated (Newman, 2004). Second, we also measured eigenvector centrality, which accounts for the possibility that an organisation has only two direct collaborators but still functions as the only link between two otherwise separated parts of the network. This indicator assigns relative scores to all nodes in the network depending not only on their own edges but also the edges of collaborators; the scores are normalised values between zero and one. This measure can help assess what kind of nodes have most influence on the network (Bonacich, 2007).

Operationalising geographical and institutional diversity

In operationalising institutional diversity, we assume that a number of organisation types can be distinguished that each share a number of institutional characteristics, such as their use and dissemination of scientific findings and resilience. Following the classification in the EUPRO dataset (Barber et al. 2008), we distinguished seven different organisation types (see Table 17).

Table 17 The organisation types and their definitions.

Organisation type	Definition
Consultancy	Consultancy businesses
Education	Faculties of universities and other educational institutions
Government	Governmental institutions and organisations
Industry	Businesses, including for-profit industrial research centres
Non-profit	Non-commercial or non-profit institutions and organisations
Research	Publicly-funded research organisations and their constituent research areas
Other	Organisations that do not fit one of the other types

Rathenau Instituut

To investigate the geographical diversity of the network and the degree of Europeanisation, we explored the differences between the centrality of organisations from traditional European member states, the newer entrants and non-European partner states.

4.3.3 The organisation as aggregation level of analysis for network studies

This study analyses a research network at the level of organisations and sub-organisations. Innovation is more and more dependent on the ability of organisations to access the newest technological insights and to establish connections to bring products and services to new and existing markets. To acquire new knowledge, organisations depend on their knowledge networks (Augier and Vendelo, 1999). The organisational level is also important in the networks that result from the EU FPs. Most of the projects are conducted by a consortium of organisations. Some instruments used in the FPs explicitly encouraged specific organisation types to participate, in particular firms (EC, 2007).

However, even within a sectoral network, large organisations sometimes operate in a variety of unrelated activity areas. For the analysis of research networks it is more informative if such organisations are split into sub-entities that represent coherent activity areas. In our analysis, therefore, we broke down universities into faculties or schools, and research institutes into research areas. Insufficient information is available on organisational structures in industry, consultancy and non-profit organisations, so only multinationals were broken down into national branches (see Barber et al., 2008). The issue is also less urgent when it comes to these organisation types because they generally participated in only a few projects in our database. The term ‘organisation’ as used below therefore also includes sub-organisations.

4.4 Analysis and results

4.4.1 Geographical and institutional development of the network

We first show how the European research network on water has developed over time. We also explore how its dynamics compare to the development of the network emerging from all FP projects.

Table 18 presents the descriptive statistics of each FP as a separate network, and the network that is constituted by the grand total of all FPs. The network has grown enormously, in terms of projects per FP and in terms of participants. The number of participants per project and the average degree of organisations have strongly increased over time. These developments are similar to the developments in the generic network constituted by all FP projects (Heller-Schuh et al., 2011). They are mainly consequences of the strong increase of both the familiarity and the available funds in the FPs over time. Probably as a co-effect, the interconnections in the water network have become stronger over time; the giant component is also much larger in relative terms in the latest FPs than it was in the first FPs, this is a common feature of networks with a growing number of nodes. The large difference in size between the generic network and the water network also implies that it is not very informative to compare them on this aspect. The average distance has not changed much over time; the figures are comparable to the generic FP network (cf. Heller-Schuh et al., 2011). The density decreased over time until FP6, which was only to be expected, as the number of nodes and hence the number of potential links has grown so strongly over time. The incidental increase of density in FP6 is associated with the strong increase in the number of partners per project at that time, especially in Integrated Project and Networks of Excellence. In FP7 there have been very few calls for Networks of Excellence; the number of partners per project dropped again (Arnold et al., 2009). The largest distance is smaller in the water network (6-8) than in the general network (7-11). Combined with the relatively high clustering coefficients this implies that even more than the general network, the water network is a smallworld type network, in which knowledge can flow relatively easily and quickly through the network (Heller-Schuh et al., 2011; Cowan, 2006). To some extent this may be attributed to the fact that the generic network is simply much larger, but still the relatively high clustering is remarkable given the fact that this sector is also characterized in the literature as fragmented and bound in national systems (EIP, 2014, Thomas and Ford, 2005).

Table 18 Basic statistics of the European research network on water in seven Framework Programmes; participant counts based on subentities.

Statistic	FP1	FP2	FP3	FP4	FP5	FP6	FP7*	Total
Starting year	**1984**	**1987**	**1990**	**1994**	**1998**	**2002**	**2007**	
Size of network (nodes)	198	547	807	1854	3007	3311	1318	7767
Number of projects	167	139	256	616	942	647	295	3062
Average partners per project	2.1	6.1	5.4	5.1	5.5	8.7	6.9	6.1
Average countries per project	1.7	3.6	3.7	3.5	3.8	5.0	4.3	4.2
Average degree	4.3	15.3	10.8	10.2	15.9	34.5	20.5	18.9
Average weighted degree	5.7	16.5	11.7	11.2	16.7	37.3	21.4	21.0
Largest component (%)	47.0	89.6	87.4	83.0	85.0	96.6	96.1	93.7
Average distance	3.207	2.976	3.412	3.738	3.462	3.048	3.220	3.181
Largest distance	8	6	8	8	8	7	8	7
Average clustering coefficient	0.48	0.84	0.83	0.79	0.79	0.86	0.87	0.79
Density	0.086	0.028	0.013	0.005	0.005	0.010	0.016	0.004
Modularity	0.62	0.65	0.67	0.70	0.67	0.58	0.67	0.50

*Last updated until March 2010

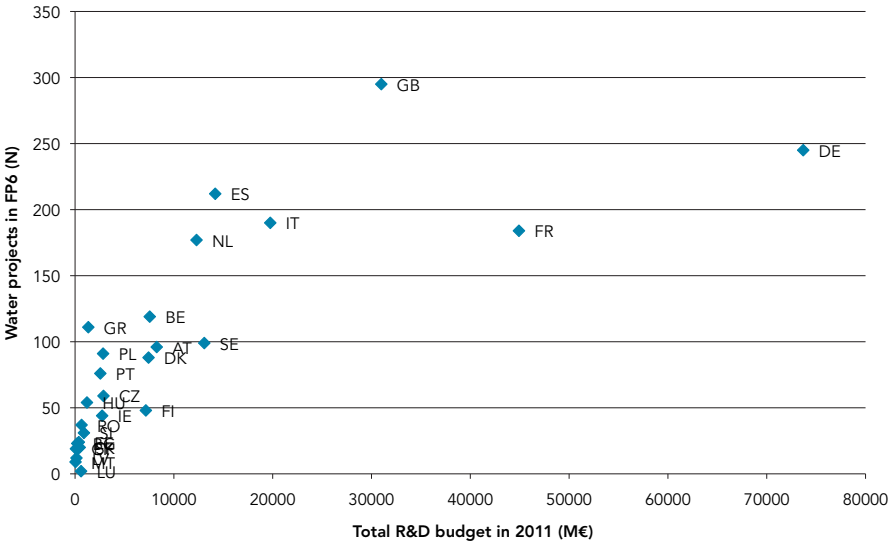
Rathenau Instituut

Explaining the geographical dynamics

We have examined the geographical patterns that we found in more depth. Using FP6 data we explored three different factors that may play a role the geographical dynamics observed: the R&D budget, projects in the entire FP and the scientific output on water of each country.

The large disparities in absolute counts of projects among countries can be partially explained by a difference in overall R&D capacity (see Figure 5). The positive relationship which is visible between water projects and R&D budget corresponds with an earlier analysis of the generic research network, based on projects in FP2 and FP3 (Sharp, 1998). However, the ratio between the two is still quite skewed across countries, indicating that new member states and countries with a relatively small research budget such as Portugal have a relatively large number of projects on water.

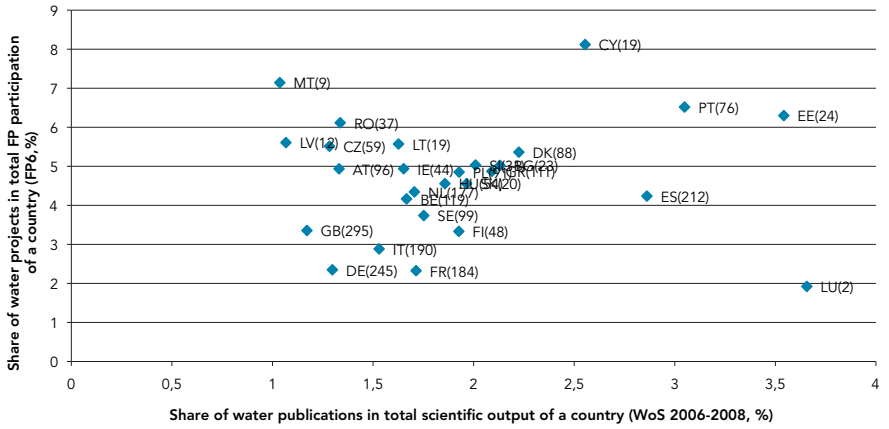
Figure 5 Number of water projects in which countries from EU27 participated in FP6 versus the amount of money (M€) the respective country spent on R&D in 2011[10].



To account for the possibility that some countries specialise more prominently in water-related knowledge production than others, we compared the relative share of water in FP6 projects with the relative share of water in scientific publications, see Figure 6. Two groups of countries stand out, with a surprisingly large share of water projects in FP6. The first group of countries, comprising Malta, Romania, Latvia, Lithuania and the Czech Republic, have a small share in the scientific output compared to other countries, while they have a large share in the FP. A second group of countries (Cyprus, Estonia, Portugal) seem more genuinely specialised in water research: they are relatively prevalent both in the FPs and in terms of scientific output.

10 R&D data are from Eurostat, latest updates 2012. R&D data on Greece from 2007 due to data availability.

Figure 6 Share of water (%) in FP6 versus Web of Science (WoS) 2006-2008[11].



Rathenau Instituut

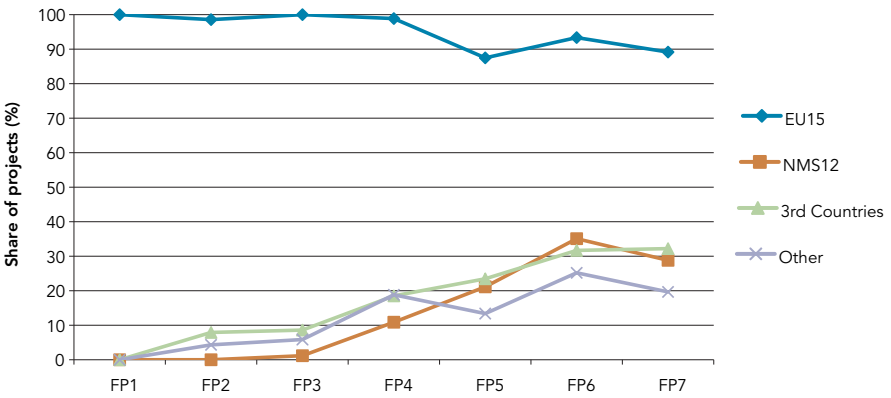
Altogether, these three indicators show that small countries have a relatively large share: they have more projects per euro invested in R&D. This effect is stronger in the field of water than in the Framework Programme in general. For some small countries this may relate to a specialisation in water research, for other countries it seems that water is a relatively accessible field to start participating in the FPs.

Distribution of projects in groups of countries

The differences in participation also relate to the member status of a country: EU15, new member state, associated country, or other, see Figure 7. This shows the continuing large share of the EU15: in the first FPs there was at least one EU15 country in literally every project, and there is still at least one partner from the EU15 in about 90% of the most recent projects. The share of projects with at least one NMS partner increased until FP6, but dropped slightly in FP7, to below 30%. The rise until FP6 can be partially attributed to explicit incentives in the calls to include organisations from the NMS. The share of associated countries has increased strongly over time, and is now also around 30%. This is even more remarkable if one notes that the lion's share comes from only three countries: Norway, Switzerland and Israel. This is not only the case in water projects; these three countries have a strong participation in FP6 in general.

11 Data on total projects per country in FP6 based on the final review of FP6 (EC, 2008).

Figure 7 Share of water projects in which at least one country in a group participated, per FP[12].



Participation of different types of organisations

Universities and public research organisations have the largest representation in the network, see Figure 8. They have increased over time in absolute numbers, but their relative share has decreased. This is mainly due to the rise of industrial partners (up to about 30% in FP7). Governmental organisations have a relatively small but stable share in the network. The 'other' group was marginal in the first FPs, yet nowadays this group is larger than the group of governmental organisations. Interestingly, the shares of Industry and Other organisations have been larger in the network across all FPs than in any of the individual FPs. This suggests that participation by organisations in these categories often remains limited to one FP. Organisations in Education and Research tend to participate in more subsequent FPs than organisations in Industry and in Other organisations.

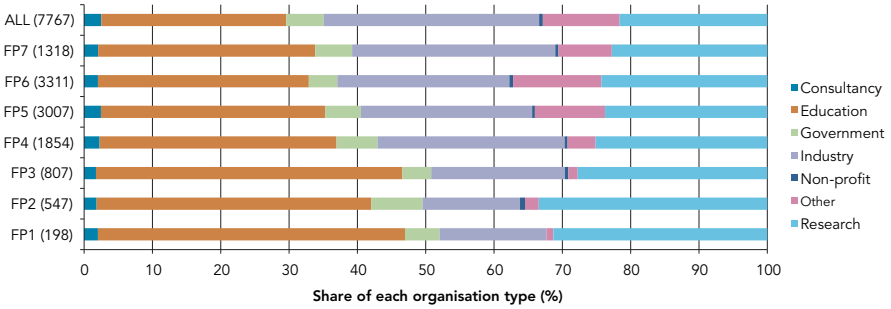
The composition of the water network differs from the composition and dynamics in the generic network. In the overall network, the share of higher education organisations increased from 32% in FP3 to 37% in FP6 (EC, 2004). In the water network, it decreased from 45% in FP3 to 31% in FP6. The share of industry (including consultancies) decreased in the general network, from 35% in FP3 to 30% in FP6 (EC, 2004), whereas it increased in the water network, from 22% in FP3 to 27% in FP6. For public research organisations it decreased both in the general network and in the water network: from 30% to 26% and from 28% to 24% respectively. All other categories together increased in the general network

12 Note that the sum of the curves exceeds 100% because many projects include countries in several groups.

to 7% in FP6, but had a higher share in the water network: 20% in FP6. The differences in share may partly relate to differences in classifications, but these cannot completely account for the trends observed. Two clear developments can be identified: first, despite all the attention for the relevance of the Triple Helix and participation of actors outside the traditional research, the generic network is hardly diversifying, industry participation is even decreasing over time.

Second, the water network is in terms of composition much more diverse than the generic network, and diversifies over time. This corroborates the characterisation of the water sector as consisting of many different organisation types, from utilities to water authorities, from consultancies to university departments (EIP, 2014; Thomas and Ford, 2005) and also shows that all these actors have found their way to EU funding.

Figure 8 Share of each organisation type in participation by unique entities, per FP (count of total unique participating entities per FP in brackets).



To summarise, the FPs have witnessed strong growth over time in terms of projects, and even stronger in terms of participating organisations. Compared to the generic network, the water network can be characterised as a small world network, where information flows relatively easily. Although the 'old' countries still dominate the network in absolute terms, the water network stands out by the fact that small countries (in terms of R&D budget and scientific output) have a relatively strong participation. The water network has also diversified over time in terms of institutional backgrounds. It is now institutionally more diverse than the generic network.

4.4.2 Centrality of organisations in the network

In this section we analyse how the centrality of actors in the network relates to institutional and geographical characteristics and the criteria of the available funding instruments. There are strong differences between the average centrality across organisation types, see Table 19. The distribution of the centrality mea-

sures shows that all organisation types contain a large group with a low centrality, having few collaborations and positioned far from the core of the network. In some organisation types, in particular Education and Research, there also is a large group with a very high centrality; this is lacking in other types such as Industry. The differences between the organisation types have increased greatly over time. For example, in FP1, the median of weighted degree ranged between 1 and 3 (with the exception of the 'Other' category, yet there were only two entities in this group in FP1); in FP7, the medians ranged between 11 and 20.5, and the full ranges were even more diverse.

Organisations in Education have the highest centrality in the network. They are immediately followed by the organisations in Research. Governmental organisations have a much lower centrality than the first two. They lack a subgroup of entities with an extraordinarily high centrality. However, compared with all other categories they are relatively central. Industry has a relatively low centrality: the average organisation in Industry has few links, and the ones it has are not important for the network. This seems to deviate from the generic network: there Education and Research have the highest centralities as well, but 15 out of the 100 most central organisations in FP6 are from industry (Heller-Schuh et al., 2011). This probably has to do with the fact that some multinational firms with a variety of activities have a strong position in the generic network; their activities in water alone (if any) are insufficient to make them a key player in this specific network.

Table 19 Median and interquartile range of weighted degree and eigenvector centrality per organisation type per FP.

Orgtype	Measure	FP1	FP2	FP3	FP4	FP5	FP6	FP7	ALL
Consultancy	Median weighted degree	1	6	4	7.5	8	15	16	10
Education		3	10.5	7	8	12	31.5	18	21
Government		2.5	24	9	8	12	24.5	16	15
Industry		1	6	6	6	7	12	11	9
Non-profit		*	16	4	8	8	29	15.5	13
Other		9.5	8	3	7	8	15	11	12
Research		2	13	7	8	11	29	20.5	17
Consultancy	Interquartile range weighted degree	2	9	4	7	8	27	14	16
Education		8	18	13	10	21	51	19	47
Government		9	20	7	7	19	41	13	24
Industry		3	9	6	7	6	17	14	9
Non-profit		*	20	6	5	11	46	20	24
Other		*	22	5	7	7	18	16	15
Research		8	22	11	9	18	48	26	40

continued on the next page ➔

Orgtype	Measure	FP1	FP2	FP3	FP4	FP5	FP6	FP7	ALL
Consultancy	Median eigen-vector centrality	0.001	0.015	0.007	0.009	0.006	0.008	0.051	0.005
Education		0.007	0.060	0.040	0.025	0.036	0.029	0.061	0.017
Government		0.005	0.191	0.029	0.024	0.032	0.023	0.042	0.011
Industry		0.001	0.009	0.012	0.007	0.005	0.005	0.024	0.003
Non-profit		*	0.101	0.021	0.007	0.032	0.024	0.038	0.012
Other		0.030	0.039	0.004	0.013	0.009	0.009	0.032	0.005
Research		0.009	0.074	0.029	0.021	0.027	0.025	0.076	0.012
Consultancy	Interquartile range eigenvector centrality	0.004	0.061	0.022	0.028	0.035	0.057	0.057	0.014
Education		0.104	0.140	0.117	0.057	0.101	0.095	0.105	0.049
Government		0.105	0.293	0.109	0.030	0.098	0.069	0.078	0.019
Industry		0.011	0.058	0.050	0.017	0.014	0.010	0.061	0.006
Non-profit		*	0.144	0.024	0.008	0.069	0.080	0.055	0.025
Other		*	0.302	0.019	0.026	0.022	0.020	0.068	0.011
Research		0.069	0.239	0.095	0.044	0.074	0.095	0.138	0.035

* Measure cannot be computed because there are not enough organisations in the respective category.

Rathenau Instituut

Characteristics of organisations with a high centrality in the network

Given our finding that the largest differences between the organisation types in terms of centrality are among the most central participants per type, let us now focus on the 50 organisations per organisation type with the highest scores on weighted degree and eigenvector centrality, to explore these organisations in terms of country of origin, time since first participation in FPs, and number of projects. It turns out that the participants with the strongest positions participated in much more projects than the average participant. The difference amounted to a factor between two and eight, see Table 20. Their relevance to the network as measured by their eigenvector centrality differed even more from the average participant. The central organisations in higher education and public research have relatively long experience of FP participation, but organisations in the other categories do not. However, once they participate, they are less likely to leave: the actors with a high centrality have participated in many more FPs than the average. By far the most organisations with a high centrality are from the EU15.

All in all, we found that the skewed distribution of centrality across different organisation types has become even more skewed over time. The population of organisations with the highest centrality scores is dominated by research organisations and universities. Many actors from industry are involved, but almost all of them have a very low centrality.

Table 20 Comparison of characteristics of all entities versus the top 50 of each organisation type.

	CON	EDU	GOV	IND	NFP	OTH	RES
Average number of projects per entity, total	1.5	3.4	2.0	1.3	1.8	1.2	3.4
Average number of projects per entity, top50	2.4	21.0	6.0	4.5	1.8	2.4	28.8
Median weighted degree, total	10	21	15	9	13	12	17
Median weighted degree, top50	33	294	90	76	13	64	330.5
Median eigenvector centrality, total	0.005	0.017	0.011	0.003	0.012	0.005	0.012
Median eigenvector centrality, top50	0.022	0.288	0.084	0.081	0.012	0.071	0.332
Average start year, total	2001	2000	2001	2002	2000	2003	2000
Average start year, top50	2001	1992	2000	2000	2000	2004	1993
Average number of FPs per entity, total	1.22	1.78	1.32	1.12	1.24	1.08	1.64
Average number of FPs per entity, top50	1.66	4.08	2.19	2.22	1.24	1.47	4.57

Note: CON = Consultancy, EDU = Education, GOV = Government, IND = Industry, NFP = Not-for-profit, OTH = Other, RES = Research.

Rathenau Instituut

4.5 Conclusions and discussion

In the concluding section of this chapter, we present the answers to our research question, set out policy recommendations and present suggestions for future research.

4.5.1 Conclusions

Our analysis has shown that the network of research collaborators in water has grown strongly since the first FP in 1984 in terms of projects and in terms of participating organisations. The smaller largest distances in the water network, combined with the relatively high clustering coefficients imply that the water network, even more than the generic network is of the small world type, where information can flow relatively quickly. The dispersion over countries follows the enlargement of the European Union over time. Remarkably, however, small countries in terms of R&D budget, including many new member states, have a relatively large number of projects. This effect is stronger in the water sector than in the generic FP network; the water network is geographically more diverse. It may be that the scarcity of funding in their own country prompts these organisations to look for funding from European programmes. However, in absolute terms, the larger and older members still dominate the projects. There have hardly been any projects so far with no participant from the EU15. Also in institutional terms the water network has diversified over time. Especially in the later FPs it is institutionally more diverse than the generic network. We can conclude that the water network is geographically still dominated by older and larger member states of the EU, yet more diverse than the generic research network, and is also institutionally more diverse than the generic network.

The organisations with the highest centrality are primarily organisations oriented towards knowledge production (higher education and public research organisations), mostly from 'traditional' countries. All organisation types include a large number of organisations[13] with a low centrality in the network, but only some organisation types include a substantial group of organisations with a high centrality in the network as well. Interviewed experts state that universities and public research institutes often benefit from dedicated resources such as project management offices to write strong applications and carry out complex projects. Governmental organisations have a relatively small share of project participation, yet their centrality is relatively high. Businesses, on the other hand, have a relatively large numerical share, but their centrality is low. On this point the water network differs from the generic network, which contains a group of important key players from industry. It has proved increasingly difficult to involve businesses. Our observation that the composition of the knowledge network is diversifying confirms claims made in the literature on the Triple Helix and Mode 2 knowledge production (Gibbons et al., 1994, Etzkowitz and Leydesdorff, 2000). Still, our analysis indicates that universities and public research institutes remain the central hubs in the network.

Organisations with a high centrality differ in general from the others first in that they started participating slightly earlier, and second in that they have participated over a longer period of time and in a larger number of projects. In contrast to findings on the overall network (Heller-Schuh et al., 2011), the strong organisations in the water network are not only large, broad-based research organisations, but also smaller organisations with a thematic focus. Overall, our analysis shows that the centrality of an organisation in the research network is strongly associated with its institutional characteristics and national background: organisations from higher education and research and organisations from the oldest member states generally have higher centrality.

4.5.2 Policy implications

Before we close, let us elaborate on three policy implications.
First, the stable network centre can be helpful for the governance of the network. The network diversifies: "New" organisation types, especially not-for-profits and firms appear on the edges of the research network. This implies a form of network governance in water research that resembles the principle of integrated water resources management (Pahl-Wostl, 2007). The downside of such governance modes is that a clear division of tasks, responsibilities and accountabilities for addressing grand challenges is lacking (Muro and Jeffrey, 2012; Biswas, 2004). In the water research network under study, however, the fact that the core of the

13 As explained in section 4.3.3, we broke down many organisations into more coherent sub-organisations for our analysis. Whenever we use the term 'organisations', we are therefore referring to the entities we analysed.

network is stable, enables the further, directed development of the network. It must be noted at the same time that the presence of a dominant perspective from these core actors can limit the introduction of new insights and approaches. There are some promising developments in Horizon 2020 to ensure long term commitment of industry. European Innovation Platforms (EIP) have been launched to speed up the development of innovations and ensure their uptake by the market. Moreover, the European Council has announced a 'Fast Track to Innovation' pilot scheme that specifically aims to speed up the process from idea to market in projects.[14] This track will probably follow a bottom-up logic, allowing participants to submit research and innovation projects at any time, making businesses less dependent on the specific themes of calls.

Second, funding instruments should be tailored to the needs and characteristics of specific research fields. Our analysis shows that the composition and dynamics of research networks vary across fields. In the field of water there are more small, specialised research organisations in the backbone of the network than in other fields. The network also includes many more organisations in the categories 'non-profit' and 'other' than the generic network. Both findings may relate to the relatively applied and challenge-oriented nature of the field. These findings indicate that research policy should be tailored to the needs and characteristics of a field. The Responsible Research and Innovation approach which features prominently in Horizon 2020 demands active participation by NGOs and other stakeholders in the research process. This interactive and integrated approach to innovation will probably develop relatively easily in the water domain, but it will require more efforts to involve these organisations in other domains.

Third, the European Commission could consider strengthening the European Research Area by inducing collective knowledge needs. The distribution of projects over countries in general is very skewed. However, our analysis also shows that water projects are relatively accessible for new countries. This may be because most research is relatively applied and hence does not require large, well-established infrastructures for basic research, and requires less expensive equipment than other research fields. Moreover, EU legislation on water, such as the Drinking Water Directive and the Water Framework Directive, have given rise to a need to develop knowledge in this field to guarantee compliance. Apparently, regulation can act as a catalyst for participation by new countries and the development of the ERA.

4.5.3 Suggestions for future research

We close by discussing some recommendations for future research. First, it is an inherent feature of the type of data we use that we are only able to analyse

14 Council of the European Union, 2013, press release 'Agreement on "Horizon 2020": the EU's research and innovation programme for the years 2014 to 2020', 17 July 2013.

structural and quantitative properties of the network and its constituents. To gain further insights into the mechanisms behind network formation, such as the role of proximity and other factors in the selection of research partners, an analysis of data on both project proposals that received funding and those that were rejected would be promising. A second important research avenue deals with the performance of the research network, and in particular the synergy benefits for the research network that emerge from an ERA. Our study has aimed to add to our understanding of the dynamics required for such synergies.

4.6 References

Annerberg, R., et al. (2010), 'Interim Evaluation of the Seventh Framework Programme: Report of the Expert Group', (European Commission).

Arnold, E., et al. (2009), 'Evaluation of the sixth Framework Programmes for Research and Technological Development', Report of the Expert Group (Brussels).

Augier, M. & M.T. Vendelø (1999), 'Networks, cognition and management of tacit knowledge', *Journal of knowledge management*, 3(4), 252-61.

Autant-Bernard, C., et al. (2007), 'Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies*', *Papers in Regional Science*, 86(3), 495-519.

Barber, M., et al. (2008), 'NEMO - Network Models, Governance and R&D Collaboration Networks. Deliverable D4.1: The sysres EUPRO database manual.'

Barré, R., et al. (2013), 'Measuring the integration and coordination dynamics of the European Research Area', *Science and Public Policy*, 40(2), 187-205.

Biswas, A.K. 2004. Integrated Water Resources Management: A Reassessment. *Water International*, 29(2): 248-256

Bonacich, Ph. (2007), 'Some unique properties of eigenvector centrality', *Social Networks*, 29(4), 555-64.

Breschi, S. & L. Cusmano (2004), 'Unveiling the texture of a European Research Area: emergence of oligarchic networks under EU Framework Programmes', *International Journal of Technology Management*, 27(8), 747-72.

Chessa, A, et al. (2013), 'Is Europe Evolving Toward an Integrated Research Area?', *Science*, 339(6120), 650-51.

Cowan, R. (2006), 'Network models of innovation and knowledge diffusion', in Stefano Breschi and Franco Malerba (eds.), Clusters, networks and innovation. Oxford: *Oxford University Press*.

Defazio, D., A. Lockett & M. Wright (2009), 'Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program', *Research Policy*, 38(2), 293-305.

Delanghe, H., U. Muldur & L. Soete (2009), European science and technology policy: Towards integration or fragmentation? *Edward Elgar Publishing*.

EC (2004), 'Five-Year Assessment of the European Union Research Framework Programmes 1999-2003' *DG Research*, Brussels.

EC (2007), 'SMEs and Research: An Impact Assessment of R&D Funding Schemes'. *DG Research*, Brussels.

EC (2008), 'FP6 final review: Subscription, implementation, participation', *DG Research*, Brussels.

EIP, European Innovation Partnership on Water (2014), 'Barriers and bottlenecks for innovation in the water sector', *European Commission*, Brussels, 9 pp.

Etzkowitz, H. & L. Leydesdorff (2000), 'The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations', *Research Policy*, 29(2), 109-23.

Ferru, M. (2010), 'Formation process and geography of science–industry partnerships: the case of the University of Poitiers', *Industry and Innovation*, 17(6), 531-49.

Foddi, M. & S. Usai (2013), 'Regional knowledge performance in Europe', *Growth and Change*, 44(2), 258-86.

Frijns, J., C. Büscher, A. Segrave & M. van der Zouwen (2013), 'Dealing with future challenges: a social learning alliance in the Dutch water sector', *Water Policy*, 15, 212-222.

Geiger, R. & C.M. Sá (2008), Tapping the riches of science: Universities and the promise of economic growth, *Harvard University Press*.

Gibbons, Michael, et al. (1994), The new production of knowledge: The dynamics of science and research in contemporary societies, *Sage*.

Hazir, C. S. & C. Autant-Bernard (2012), *Using Affiliation Networks to Study the Determinants of Multilateral Research Cooperation Some empirical evidence from EU Framework Programs in biotechnology*, Working Papers. Lyon: Groupe d'Analyse et de Théorie Économique Lyon - St Étienne.

Heller-Schuh, B., et al. (2011), JRC scientific and technical reports: 'Analysis of Networks in European Framework Programmes' (*1984-2006*).

Hicks, D. M. & J.S. Katz (1996), 'Where is science going?', *Science, Technology & Human Values*, 21(4), 379-406.

Hoekman, J., K. Frenken & R.J.W. Tijssen (2010), 'Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe', *Research Policy*, 39(5), 662-73.

Hoekman, J., et al. (2013), 'Acquisition of European research funds and its effect on international scientific collaboration', *Journal of Economic Geography*, 13(1), 23-52.

Lata, R., T. Scherngell & T. Brenner (2012), *Integration Processes in European R&D: A comparative spatial interaction approach using project based R&D networks and co- patent networks*, Paper presented at the 52th European Congress of the Regional Science Association International.

Muro, M. & P. Jeffrey 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.

Newman, M.E.J. (2004), *Coauthorship networks and patterns of scientific collaboration*, Proceedings of the National Academy of Sciences of the United States of America, 101 (Suppl 1), 5200-05.

Nieminen, M. & E. Kaukonen (2001), 'Universities and R&D networking in a knowledge-based economy', A glance at Finnish development. *Sitra Reports series*, 11.

Owen-Smith, J. & W.W. Powell (2004), 'Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community', *Organization Science*, 15(1), 5-21.

Pahl-Wostl, C., M. Craps, A. Dewulf, E. Mostert, D. Tabara & T. Taillieu (2007), 'Social learning and water resources management'. *Ecology and Society* 12(2): 5. [online] URL: http://www.ecologyandsociety.org/vol12/iss2/art5/

Paier, M. & T. Scherngell (2011), 'Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model', *Industry and Innovation*, 18(1), 89-104.

Potì, B. & E. Reale (2007), 'Changing allocation models for public research funding: an empirical exploration based on project funding data', *Science and Public Policy*, 34(6), 417-30.

Protogerou, A., Y. Caloghirou & E. Siokas (2010), 'Policy-driven collaborative research networks in Europe', *Economics of Innovation and New Technology*, 19(4), 349-72.

Scherngell, T. & R. Lata (2012), 'Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data*', *Papers in Regional Science*.

Sharp, M. (1998), 'Competitiveness and cohesion – are the two compatible?', *Research Policy*, 27, 569-588.

Thomas, D. A. & R. R. Ford (2005), 'The Crisis of Innovation in Water and Wastewater', Cheltenham, UK, *Edward Elgar*. 336 pp.

Weterings, A. & R. Ponds (2009), 'Do Regional and Non-regional Knowledge Flows Differ? An Empirical Study on Clustered Firms in the Dutch Life Sciences and Computing Services Industry', *Industry and Innovation*, 16(1), 11-31.

Wen, B., E. Horlings, M. van der Zouwen, P. van den Besselaar & W. van Vierssen (2011). 'Mapping science through bibliometric triangulation: an experimental approach applied to water research', *ASIS&T SIG/MET Paper Contest*, May 2011

5 How do dimensions of proximity influence international research collaboration?[15]

5.1 Introduction

Over the past few decades the interest in collaborative knowledge production has grown, both in scientific literature and in policy discourse. Two trends are of special importance: an increase in international collaboration, and a growing emphasis on collaboration across different societal sectors, often described in the literature under the label of 'Triple Helix' (Etzkowitz and Leydesdorff, 2000).

Especially in Europe there is a great interest in collaborative knowledge production across geographical distances. This is obvious in the literature (Hoekman et al., 2010; Chessa et al., 2013), but also in policy initiatives, which show a special focus on crossing geographical borders. In addition to national research policies to stimulate international collaboration and exchange, the European Union (EU) has formulated explicit goals to build a European Research Area (ERA), where knowledge can circulate without hindrance of national borders (Delanghe et al., 2009). The EU also plays an important role in accomplishing these goals by means of the Framework Programmes and the recently launched successor Horizon 2020. The FPs contain several funding instruments to stimulate research and knowledge creation and to build and strengthen research networks (Heller-Schuh et al., 2011; Delanghe et al., 2009).

Intertwined with this process, the emphasis on Triple Helix collaboration has grown. The conceptual idea that collaboration among partners with different institutional backgrounds drives innovation (Etzkowitz and Leydesdorff, 2000) has been implemented in both national and international research policies (Nieminen and Kaukonen, 2001; Potì and Reale, 2007). The FPs stimulate research organisations to involve their stakeholders in the knowledge production process, and they stimulate firms to involve public research organisations in their precompetitive innovation process.

The effects of these two trends on the dynamics of research collaboration are still poorly understood. With regards to internationalization, several studies have shown that despite all attention and stimulation of this process, research is still mainly organised in national systems (Chessa et al., 2013; Frenken et al., 2009). Moreover, the effect of geographical distance does not seem to decrease over time, despite all modern communication technologies (Hoekman et al., 2010).

15 This chapter has been submitted – in slightly different form – to *Industry and Innovation*.

With regards to collaboration across institutional backgrounds, earlier studies have found a diversification of the actors involved in research networks (Heller-Schuh et al., 2011), but it has also been shown that actors still have a preference to collaborate with organisations that have a similar background (Balland, 2012; Bouba-Olga et al., 2012).

Our aim for this chapter is to contribute to the understanding of the effect of proximity on international collaboration by statistically analysing European water research projects.

Our data have been drawn from collaborative projects in the water sector. This sector is a relevant case for three reasons. First, water-related research has been funded almost since the inception of European research policy. This funding was not very visible in the first decades, as water was spread over various instruments and programmes. In recent years, its visibility has increased with the launch of several dedicated policy instruments for water, such as a European Technology Platform, a Joint Programming Initiative and a European Innovation Partnership. Second, water research involves an interesting combination of local and international scales. Like most environmental sciences water research typically deals with knowledge which is specific to local circumstances such as soil conditions and the relevant actors are mostly organised in national systems, while the challenges typically cut across organisational and national borders (EIP, 2014; Thomas and Ford, 2005). Third, knowledge production in a field like water inherently involves actors from different organisational backgrounds – from (semi)-public partners like utilities to consultancies and from regional authorities to universities. It has been suggested that this – combined with the organisation in national systems – leads to a fragmented knowledge landscape (EIP, 2014), but there seems to be no evidence on the actual research dynamics and collaborative behaviour in the sector.

To analyse the research dynamics, and more specifically the aspects of international collaboration and collaboration across different organisational types, we use the concept of proximity. The central idea is that for proximate partners (not only in geographical, but also in social or organisational sense) it is easier to collaborate: their proximity eases coordination and reduces uncertainty and transaction costs. However, too much proximity might lead to lock-in effects (Boschma, 2005). The empirical literature about the actual effects on collaborative behaviour is diverse and growing, but it has some limitations. Existing studies mostly use patents or publications as data source, which has the disadvantage that they are oriented towards scientific and commercial knowledge production respectively, which do not represent all relevant actors in the knowledge collaboration networks (see Bouba-Olga et al., 2012 for an overview of proximity studies on science-industry collaboration). Moreover, most studies analyse only one dimension of proximity. Studies that do analyse more dimensions and use broader data than patents or publications (Broekel and Boschma, 2012; Balland,

2012) use relatively small sets of very specific, relatively homogeneous groups (in the aviation industry and satellite navigation industry respectively). Proximity may have different effects in more heterogeneous and broader fields.

This brings us to our main research question: what dimensions of proximity (most strongly) influence international research collaborations in the water sector? This question is broken down into two sub-questions:

- What is the influence of the geographical, organisational and social dimension of proximity respectively on the propensity of collaboration?
- Are there relations among the different dimensions of proximity?

The rest of this article is organised as follows: in section 5.2 we present an overview of earlier literature and show our research model. In section 5.3 we describe the collection and processing of the data and the research methodology. In section 5.4 we present our findings. In section 5.5 we give our conclusions and discuss policy implications and future research questions.

5.2 Conceptual framework

In addressing the questions raised above, our study builds on several existing bodies of literature. Below we discuss earlier findings about research collaboration, the three dimensions of proximity analysed in this chapter (geographical, organisational and social) and the interactions between these dimensions.

5.2.1 Research collaboration

Research is increasingly seen as an almost inherently collaborative process. As the complexity of research problems has increased and research differentiates rapidly into specialized research fields, there also is an increasing tendency for specialization and, hence, for collaboration (Melin, 2000). Collaboration can have different manifestations, from a division of labour to access to research equipment or transmission of know-how (Laudel, 2001; Katz and Martin, 1997).
In relation to the concept of proximity, three aspects of research collaboration have been analysed in the literature (Aguiléra et al., 2012):

- How proximity affects the choice of collaboration partners and network formation (such as Autant-Bernard et al., 2007; Balland 2012).
- How proximity to collaborators affects the innovative performance of the collaborating organisations (Nooteboom et al., 2007; Broekel and Boschma, 2012).
- How proximity affects the processes of knowledge production and knowledge sharing in research collaborations (Boschma 2005; Weterings and Ponds, 2007).

This study is positioned in the first stream; we analyse how dimensions of proximity affect the propensity of collaboration between actors. Network

formation or the choice of collaborators can be studied at several levels of aggregation. Most proximity studies address one of the following levels of aggregation: regions (Scherngell and Barber, 2009; Scherngell and Lata, 2012), organisations (Balland, 2012; Broekel and Boschma, 2012) or individuals (Bozeman and Corley, 2004; see also chapter 4).

Our study is at the level of sub-organisations. Organisations are relevant entities in knowledge production. Innovation is more and more dependent on the ability of organisations to access the newest technological insights and to establish connections to bring products and services to new and existing markets. To acquire new knowledge, organisations depend on their abilities to connect with other organisations (Augier and Vendelo, 1999).

However, large organisations sometimes operate in a variety of unrelated activity areas, where it cannot be assumed that individuals in different areas are aware of the collaborations of each other. For the analysis of research collaborations it is more informative if such organisations are split into sub-entities that represent coherent activity areas. In our analysis, therefore, we broke down universities into faculties or schools, and research institutes into research areas. Insufficient information is available on organisational structures in industry, consultancy and non-profit organisations, so only multinationals were broken down into national branches (see Barber et al., 2008). Many organisations in those categories occur only once in the dataset anyway, making the issue of disaggregation much less urgent. The term 'organisation' as used below therefore also includes sub-organisations.

An central discussion in the literature on research collaboration deals with the increase of international collaborations (Leydesdorff and Wagner, 2008; Hoekman et al., 2010). This is essentially a discussion on the relevance of geographical proximity: an increase in international collaborations implies that the effect of geographical proximity has decreased. We hence elaborate further on this in section 5.2.2 under the heading of geographical proximity.

5.2.2 Findings on the three dimensions of proximity

The concept of proximity is increasingly used to analyse collaborative behaviour. In earlier literature the focus was on geographical proximity, but recently several other dimensions of proximity have been described (Knoben and Oerlemans, 2006). Building on review papers by Boschma, (2005) and Knoben and Oerlemans (2006), we distinguish three dimensions: geographical, organisational, and social proximity . Cognitive and institutional are not discussed in this study because the available data do not allow for meaningful indicators of these dimensions at the level of organisations. Below, for each of the three dimensions analysed, we elaborate on its definition, the relationship with collaboration, and earlier findings in the literature regarding this dimension.

Geographical proximity

Geographical proximity is defined as the shortest possible physical distance ("as the crow flies") between the location of two organisations. Theory suggests it promotes collaboration as it eases informal communication and facilitates processes of learning and innovation, possibly as substitute or complement to other dimensions of proximity (Rallet and Torre, 1999; Katz and Martin, 1997; Boschma, 2005). Bouba-Olga et al. (2012) provide an overview of empirical studies on the impact of the geographical dimension on research collaboration between science and industry, at several levels of aggregation, from the level of individual firms to the level of regions. Regardless of the spatial scale and the scientific field that is analysed, most studies find a positive effect for geographical proximity. In other words, more proximate actors tend to collaborate more (e.g. Balland, 2012; Broekel and Boschma, 2012). Weterings and Ponds (2009) refine this view by showing that although most collaborations are indeed geographically proximate, the ones across larger distances are considered more valuable by the collaborators. Moreover, some suggest that the effect of geographical distance does decrease over time (Scherngell and Lata, 2012). There is indeed a rise in the share of internationally co-authored papers (Leydesdorff and Wagner, 2008). International collaboration may be attractive because international publications have more citations (Narin et al., 1991) or because there is an incentive from funding instruments, like in the European Framework Programmes. Hoekman et al. (2010) have shown that in Europe, the tendency to work with physically proximate co-authors has not decreased over time, while the tendency to collaborate with collaborate with co-authors within the same territorial borders did decrease slightly over time. However, if one counts the share of international collaborations instead of papers, and considers each pair of addresses as one collaboration, the tendency to collaborate domestically remains over time (Frenken et al., 2009). Barber and Scherngell (2011) find that spatial configurations differ among thematic communities.

Organisational proximity

We define organisational proximity as a similarity in incentives and routines of two collaborating organisations (Boschma, 2005; Aguiléra et al., 2012). Sometimes the label of institutional proximity is used for this kind of proximity (Ponds et al, 2007). It can be argued that such similarities facilitate and promote collaboration as they reduce uncertainty and opportunism. A similarity in the incentives and routines is often related to the goals organisations aim for (because the organisations strive for the same types of output for example). This also provides control mechanisms to protect intellectual property, which can reduce uncertainties regarding the potential rewards for the produced knowledge (Boschma, 2005). The concept of organisational proximity relates to the concept of Triple Helix (Etzkowitz and Leydesdorff, 2000; Hessels and Van Lente, 2008). The basic assumption in this strand of literature is that universities, industry and governmental agencies are increasingly interdependent and tend to collaborate in knowledge production. Compared to geographical proximity, empirical work

on organisational proximity is much scarcer. Moreover, the way it is defined and made operational is more diverse. Most studies use a classification of organisation types, for example profit versus non-profit (Broekel and Boschma, 2012), academic versus non-academic (Cunningham and Werker, 2012), or academic, commercial and governmental (Ponds et al., 2007). Others define organisational proximity as a degree of strategic interdependence, and make it operational as whether or not two actors belong to the same corporate group (Balland, 2012). Findings differ with regard to the effect of organisational proximity. Broekel and Boschma (2012) employ a multivariate model with four dimensions of proximity and find that organisational proximity has no effect. Others who assess more dimensions find an indirect effect: non-academic partners are cognitively more proximate (Cunningham and Werker, 2012); the effect of geographical proximity is stronger if organisational proximity is lower (Ponds et al., 2007). Balland (2012) finds a direct and positive effect.

Social proximity

We define social proximity as the social embeddedness of the collaboration between actors, following from for example prior collaboration experience or other social connections (Aguiléra et al., 2012). This is beneficial for collaboration as social proximity facilitates and fosters joint knowledge production and knowledge exchange (Broekel and Boschma, 2012). Common prior acquaintances for example facilitate the circulation of information and thus reduce transaction costs (Autant-Bernard et al., 2007). Both Autant-Bernard et al. (2007) and Ter Wal (2009) find that social proximity is the strongest predictor of the dimensions studied in their respective analyses; they find a positive effect. Also Broekel and Boschma (2012), while controlling for geographical and cognitive proximity, find a positive effect for social proximity. On a similar note, Fleming et al. (2007) show with network analysis that shorter path lengths between actors correlate with increased innovation. However, Balland (2012) tests four dimensions simultaneously and finds no effect for social proximity. Comparisons between different dimensions of proximity on size of effect are inconclusive. Hardeman (2012) finds that geographical proximity has a stronger effect on collaboration than organisational proximity. However, Autant-Bernard et al. (2007), based on data about collaboration in Framework Programme projects, finds that social proximity matters more than geographical proximity, and that geographical proximity has no effect for organisations that participate in more projects. Then again, Cunningham and Werker (2012) test for organisational, geographical and technological proximity and find that geographical proximity is statistically most significant, while technological proximity has the largest effect.

5.2.3 Interactions between the dimensions of proximity

The literature suggests that the dimensions of proximity can also influence each other. One could expect that proximity in one dimension can help to overcome distance in another dimension (Boschma, 2005). It is hence important to assess the impact of each dimension in relation to the other dimensions, to get full

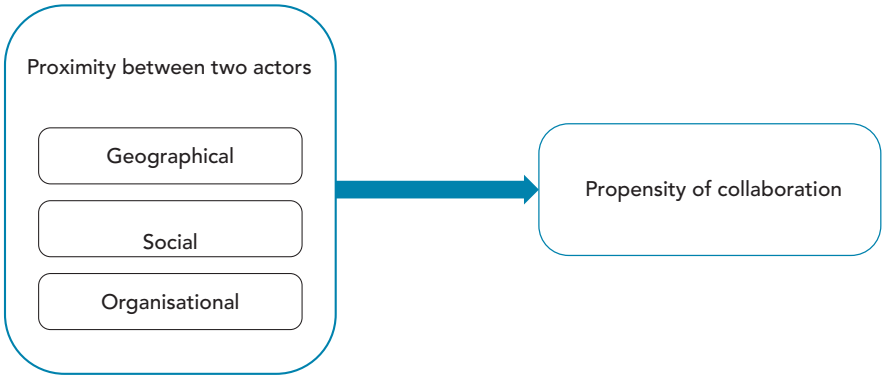
insight in the effect of each of the dimensions (Bouba-Olga et al., 2012). Ter Wal (2009) for example shows that the cognitive dimension in his analysis has a positive effect, which turns negative if controlling for other dimensions. Also, some find an indirect effect of an dimension that only becomes apparent while testing for the effect of another dimension (Cunningham and Werker, 2012).

Some literature specifically suggests that geographical proximity rather has an auxiliary effect to the other dimensions than an effect in itself (Boschma, 2005). Indeed, Breschi and Lissoni (2003) show that the effect of geographical proximity disappears if controlling for social proximity; also Ter Wal (2009) finds that the effect of geographical proximity becomes much weaker once controlling for social proximity. On the other hand, Ferru (2010) finds that social proximity conserves geographical proximity: because partners renew existing collaborations rather than initiating new ones, they also reinforce the existing geographical patterns. Others do not find any interaction between dimensions. Broekel and Boschma (2012) test a model with four dimensions (geographical, social, organisational and technical) and find that all four have a direct effect on network formation, even while controlling for the others. This is corroborated by Balland (2012) and Cunningham and Werker (2012) who also find a direct effect for geographical proximity in a multivariate model. For a proper assessment of the effect of each dimension of proximity, the dimensions should hence be tested both in bivariate and in multivariate models, to test for potential interactions between the dimensions.

5.2.4 Research model

The components discussed above together constitute the research model for our analysis, depicted in Figure 9. The three dimensions of proximity that we have defined – geographical, social and organisational – are all three expected to influence the propensity of organisations to collaborate. The dimensions are depicted as distinguished elements within proximity, as we have analysed them both one by one in separate bivariate models, and simultaneously in multivariate models.

Figure 9 Conceptual research model.



5.3 Data and methods

This study is part of a larger research project on the dynamics of collaboration and network formation as result of EU Framework Programme projects. One previous study analyses a dataset that is extracted and delineated in the same way as the data studied in this chapter. The sections on data extraction and delineation of the water sector hence largely build on chapter 4 of this thesis.

5.3.1 Data extraction

This chapter is based on an analysis of data on participation in projects under the European Framework Programmes (FPs). Research networks consist of collaborative links between actors that produce or share knowledge. Data on the collaborative links in projects funded by the FPs are a promising source of data for studying patterns in research networks (e.g. Scherngell and Lata, 2012; Paier and Scherngell, 2011). The emerging research networks are well suited to answer questions about the geographical and organisational proximity, because the FPs explicitly aim both to promote Europeanisation and to involve actors from different organisation types.

We have used the EUPRO database for data on the FP projects, just like in chapter 4. This database contains a cleaned and harmonised version of the data that is publicly available through the information service CORDIS (Barber et al., 2008). EUPRO is produced by the Austrian Institute of Technology (AIT). Based on a set of selection keywords we used all water-related projects in the EUPRO database from FP1 to FP7 whose latest updates were in March 2010. To deepen our understanding of the results, we have talked to experts from state funding agency AgentschapNL (now RVO.nl), KWR Watercycle Research Institute, consultancy Evers+Manders, and European technology platform WSSTP.

5.3.2 Delineation of the water sector

In order to construct a database that contains all FP projects on 'the water sector', we developed a set of keywords that filtered out the relevant projects (see also chapter 4). In this study we define the water sector as consisting of all human activities associated with the water cycle: production/purification and transport of drinking water, collection, transport and treatment of wastewater, water storage, water use and water management, including flood protection. Starting from this definition, we have selected relevant keywords. We took a three stage-approach: we first used a very broad set of keywords (building on Wen et al., 2011) to extract all potentially relevant projects, see appendix A. After this we employed a more refined set of keywords to filter out all false positives. To find additional terms related to trends and policy contexts, we consulted several policy documents. In the final stage, we consulted several water experts from across Europe with experience of FP projects to add more keywords. We have tested all potential keywords with random samples of projects to explore what kind of projects were extracted from the database. The final keyword set is in appendix A.

Some projects in the Framework Programmes also involve partners from outside Europe. Such participation is allowed, but not funded through the programmes. In our analysis we have only included all collaborations with partners inside the EU and the associated countries; both because collaborations outside these countries do not reflect the accomplishments of the FPs as a funding instrument, and because the inclusion of partners outside Europe would introduce large biases – particularly for geographical proximity – in the construction of a control group (see also under section 5.3.3).

The extraction and delineation described above result in a dataset that contains 2963 projects, with a starting date between 1985 and 2010 (the latest updates of the version of EUPRO that we use were in March 2010). In total 7634 organisations participate in these projects.

5.3.3 Measurement of the variables

Collaboration is defined as the joint participation of two organisations in an FP project. In many projects more than two organisations are involved, in that case each possible pair of participants in a project is considered as collaborators. The propensity to collaborate does not follow immediately from the dataset, as the pairs of actors in the data are collaborators by definition. Because we are interested in factors that influence the propensity of actors to collaborate, we compare pairs of collaborators with pairs of organisations that could have collaborated but did not. For this comparison we have constructed a control group with pairs of randomly selected actors that did not collaborate in our actual dataset. Because the eligibility criteria for partners have changed over the course of the Framework Programmes, the control group is constructed in such a way that two collaborators have never participated in the same Framework

Programme in the actual dataset. To ensure full comparability, the number of collaborations within a Framework Programme in the control group is always equal to the number of actual collaborations in that Programme. The collaboration variable is set as dichotomous, two actors either collaborate or they do not. If the same pair of organisations has collaborated more than once, each project in which they collaborated is considered as a separate observation. There are hence no weights in the collaboration variable.

Geographical proximity is measured as the shortest possible distance between the locations of two collaborating organisations ("as the crow flies"). The locations of the organisations is identified at city level in the EUPRO database. The city names are then harmonized: city names with different spelling variants get one unique name, and all cities are assigned a unique identification code. The cities are georeferenced: for each city the longitude and latitude are identified using georeferencing software (*http://www.gpsvisualizer.com/geocoder/*). The shortest possible distance is then calculated using the Haversine formula (Sinnot, 1984).

Organisational proximity is based on a categorization of organisation types. This categorization is based on differences in the routines, incentives and goals of organisations. It distinguishes seven different types: Universities and higher education organisations, Public research organisations, Governmental organisations, Industry, Consultancy, Not-for-profit organisations, and Other. This categorization is readily available in the dataset (Barber et al., 2008). Two organisations are considered to be organisationally proximate if they belong to the same category. This is hence a dichotomous variable.

Social proximity is based on a shared collaboration history of the collaborators. The social embeddedness of a collaboration between two actors is built up by, among other things, the mutual experiences in prior collaborations (Aguiléra et al., 2012; Autant-Bernard et al., 2007). Two collaborators are considered socially proximate if they have collaborated in another project in the five years before the starting date of the project under consideration. The period of five years is chosen as prior collaborations which are longer ago are considered less likely to still influence the social proximity of the prior collaborators.

5.3.4 Statistical model

We have analysed the influence of the various proximity dimensions both in bivariate and in multivariate models. We have used two categories of bivariate models. The first is used for social and organisational proximity. Because these are dichotomous variables, we have used Pearson's chi-square test. To analyse not only the statistical significance, but also the size of the effect, we have also calculated the odds ratios and Cramer's V. A second bivariate model is used for geographical proximity, because it is operationalised as a continuous variable. The effect is tested with point-biserial correlations and a binary logistic regression model with only one predictor.

For the multivariate analyses we have employed a logistic regression model. This allows for the inclusion of both dichotomous and continuous predictor variables, and for the use of a dichotomous dependent variable. The full multivariate analysis is specified as a standard binary logistic regression model.

The model is estimated using standard maximum likelihood procedures. For the optimal statistical fit of a binary logistic regression model, it is sometimes necessary to transform the explanatory variables. The optimal transformation can be found by assessing the functional form of the explanatory variable (Kay and Little, 1987). In our case a logarithmic transformation of the continuous explanatory variable – geographical proximity – gives the best results.

5.4 Results and analysis

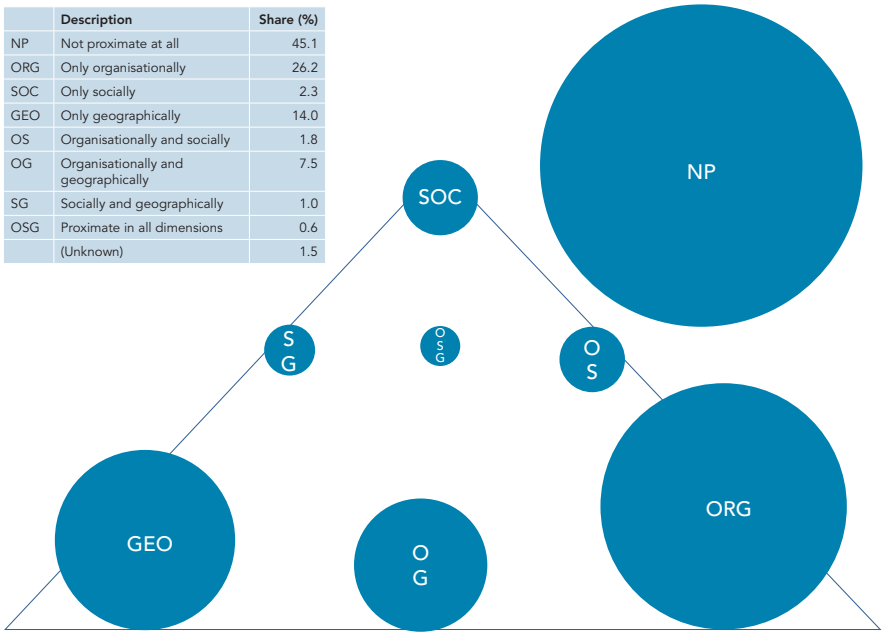
5.4.1 The effect of proximity on the propensity to collaborate

Proximity of the collaborators

Figure 10 shows that many collaborators are not proximate in any of the three dimensions. Social and organisational proximity are measured as dichotomous variables; for this analysis we have dichotomized geographical proximity. Collaborations across distances of 500 kilometre or less are considered proximate, all further away are considered as not proximate. This point is chosen as cut-off point because it is hard to make a one-day visit to a collaborator at a distance of more than 500 kilometre.

Figure 10 Shares of collaborators that are proximate in the specified dimensions. Circles are proportional to the share of collaborations with the specified proximities.

	Description	Share (%)
NP	Not proximate at all	45.1
ORG	Only organisationally	26.2
SOC	Only socially	2.3
GEO	Only geographically	14.0
OS	Organisationally and socially	1.8
OG	Organisationally and geographically	7.5
SG	Socially and geographically	1.0
OSG	Proximate in all dimensions	0.6
	(Unknown)	1.5



Rathenau Instituut

If collaborators are proximate, they are often only proximate in one dimension; only about 10% of the collaborators is proximate in more than in one dimension, and only about 1% is proximate in all three dimensions. This also shows that the FPs enable organisations to collaborate across distances. About 50% of the collaborator pairs is not proximate in any dimension. A sensitivity analysis shows that a different cut off point for geographical proximity does not alter these findings: if 200 km is set as cut off point, 54% of the collaborator pairs is not proximate at all, and if it is set at 600 km this slightly decreases to 42% (Table 21).

Table 21 Sensitivity analysis of geographical proximity. Collaborations with the specified dimensions of proximity as share of all collaborations (%) for several cut-off points for geographical proximity.

Cut-off point geographical proximity (km)	200	300	400	500	600
Not proximate at all	54.2	52.0	48.9	45.1	41.8
Only organisationally	31.2	29.9	28.2	26.2	24.4
Only socially	2.9	2.8	2.6	2.3	2.0
Only geographically	4.9	7.2	10.2	14.0	17.4
Organisationally and socially, not geographically	2.2	2.1	2.0	1.8	1.6
Organisationally and geographically, not socially	2.5	3.8	5.5	7.5	9.3
Socially and geographically, not organisationally	0.3	0.5	0.7	1.0	1.3
Proximate in all dimensions	0.2	0.3	0.4	0.6	0.8
Unknown	1.5	1.5	1.5	1.5	1.5

Rathenau Instituut

Relationships between proximity and propensity to collaborate

The bivariate analyses show that all three dimensions of proximity have a positive effect on collaboration: proximate actors are more likely to collaborate, see Table 22 and Table 23. The indicators for effect size for organisational and social proximity show that social proximity has a stronger effect. The comparison with geographical proximity is more complicated, since this is a continuous variable. However, using the dichotomized variant that considers all collaborators within 500 km of each other as geographically proximate, the odds ratio for geographical proximity is 1.65; this suggests the effect is more or less comparable in size with the effect of organisational proximity. The value for e^B - raising the base of the natural log to the power of B - implies that a change of one integer in the logarithm of the inverse of the geographical distance changes the odds that a pair of organisations is in the group of collaborators rather than the control group with a factor of 1.95. In other words: if the geographical distance between two actors decreases from 1000 to 100 km, the odds that they are collaborators almost doubles. We have also tested geographical proximity with a point-biserial correlation. This corroborates the results of the logistic regression: Spearman's r= .118 with a significance level of .000.

Table 22 Geographical proximity. Binary logistic regression with geographical proximity as only predictor.

Goodness of fit:					
N	186584				
-2Loglikelihood	254649				
Cox Snell pseudo R2	.018				
Nagelkerke pseudo R2	.024				
Overall percentage correct	54.9				
Model:					
Variable	B	eB	SE	Wald	Significance
Geographical proximity	.667	1.949	.012	3072	.000
Constant	2.082	8.018	.036	3356	.000

Rathenau Instituut

Table 23 The effect of social and organisatonial proximity on propensity to collaborate. Chi-square tests and odds ratios.

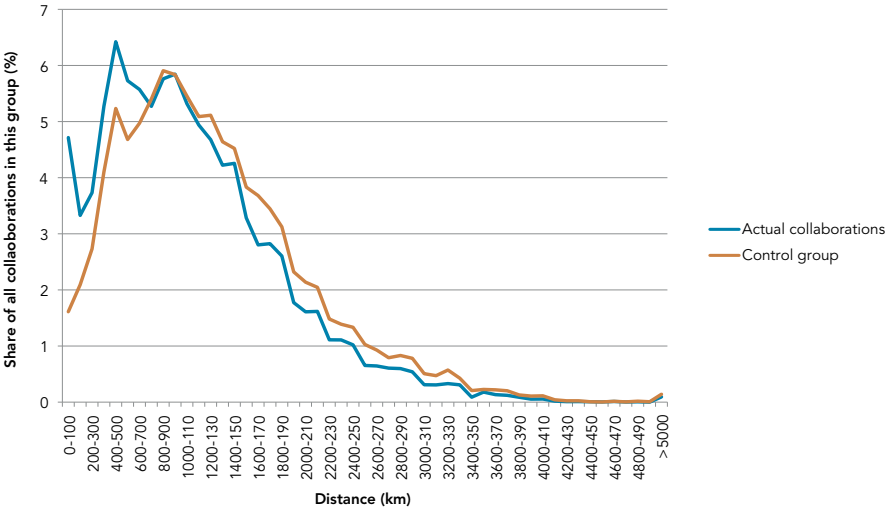
	Organisational proximity	**Social proximity**
N	189947	189947
X^2 (p-level)	2859 (.000)	4139 (.000)
Odds ratio	1.71	11.86
Cramer's V	.123	.148

Rathenau Instituut

We have shown in Figure 10 that social proximity is relatively rare, while the bivariate analysis shows that social proximity has a relatively strong effect. In other words: although it is not very likely that two actors collaborate in two in more projects, it still happens a lot more than it would if collaborations were chosen completely at random.

Figure 11 shows that there are more collaborators across small spatial distances than expected in a random distribution, with a clear peak for the smallest distances, less than 100km. However, it also shows that after about 800 km, the groups have more or less the same distribution. This implies that although the FPs enable more proximate collaborations than would emerge in a completely random network, it still results in a mix of both local and long-distance relationships.

Figure 11 The distribution of actual collaborations and the control group across geographical distance.



5.4.2 Relations among the different dimensions of proximity

The multivariate binary logistic regression model shows that all three dimensions have a positive effect, see Table 24. Here again, e^B can be interpreted as odd ratios: if the value of the predictor variable increases one unit, the odds that a pair of actors are actual collaborators increases by this factor. Because organisational proximity and social proximity are dichotomous, while geographical proximity is a continuous variable, the effect size is still difficult to compare. However, just like in the bivariate logistic regression, the value of e^B of 1.929 for geographical proximity means that a pair of organisations at 100 km distance of each other have twice the odds of collaborating compared to a pair of organisations at 1000 km distance.

Table 24 Multivariate binary logistic regression testing the effect of geographical, organisational and social proximity on the propensity to collaborate.

Goodness of fit:					
N	189947				
-2Loglikelihood	247391				
Cox Snell pseudo R2	.056				
Nagelkerke pseudo R2	.074				
Overall percentage correct	58.4				
Model:					
Variable	B	Exp (B)	SE	Wald	Significance
Logarithm of inverse of geographical distance	.657	1.929	.012	2916	.000
Organisational proximity	.532	1.703	.010	2629	.000
Social proximity	2.396	10.975	.049	2380	.000
Constant	1.836	6.272	.036	2546	.000

Rathenau Instituut

The results corroborate the findings of the bivariate analysis presented in section 5.4.1. Significance and direction of the variables does not change, and even the values of the estimated parameters change only slightly. This also implies that there is no direct interaction between the different dimensions of proximity in the sense that the effect of one dimension alters if controlling for other dimensions.

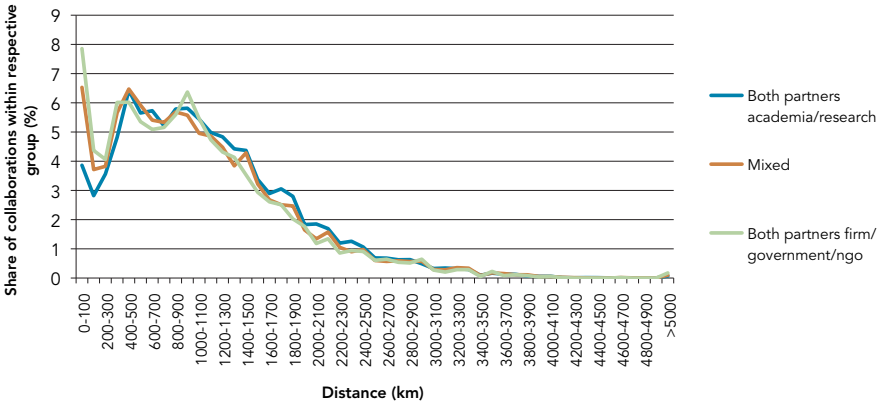
This does not imply that the dimensions do not influence each other at all. To illustrate this, we have regrouped the pairs of actual collaborators into three groups: pairs where both organisations are from academia and/or public research organisations, pairs where both partners are a firm, governmental organisation or NGO, and pairs with a mix of those two. We have then run a multinomial logistic regression model to with these groups as dependent variable and geographical and social proximity as explanatory variables. Both are significant, see Table 25. Figure 12 illustrates this effect for geographical proximity. It shows the distribution of collaborations across distance for the three groups of collaborators. The "research group" has relatively the smallest shares in geographically proximate relations, while the "knowledge user group" has most, with almost 8% of all collaborations in this group at a distance of less than 100 km. The "mixed group" is in between the two others. Apparently, it is not only organisational proximity that matters to geographical proximity, but also the organisational background in itself.

Table 25 Multinomial logistic regression testing the effect of geographical and social proximity on three groups of collaborators: (1)both in research, (2)both a firm, governmental organisation or NGO, and (3)mixed pairs. Group 1 is used as a base line in the model.

Goodness of fit:					
N	100109				
-2Loglikelihood intercept only	35570				
-2Loglikelihood Final	33843				
Cox Snell pseudo R2	.017				
Nagelkerke pseudo R2	.020				
Likelihood Ratio test (LR)	1727 (.000)				
Model:					
Variable	B	Exp (B)	SE	Wald	Significance
Geographical proximity group 2	0.411	1.508	.018	508	.000
Social proximity group 2	1.371	3.938	.062	489	.000
Intercept group 2	-1.554		.079	388	.000
Geographical proximity group 3	0.300	1.350	.014	438	.000
Social proximity group 3	.779	2.178	.034	537	.000
Intercept group 3	-0.424		.052	67	.000

Rathenau Instituut

Figure 12 The share of collaborations across distance for the groups of collaborators that are both in academia and/or public research organisations, both in firms, government and/or NGOs, and mixed pairs.



Rathenau Instituut

5.5 Conclusions and discussion

5.5.1 Conclusions

Our main conclusion is that all three dimensions have a positive effect on collaboration without any interactions among them. First, we have shown that although many collaborators are not proximate, all three dimensions of proximity do promote collaboration. About half of the collaborators are not proximate in any of the dimensions tested. All dimensions of proximity have a positive effect: proximate actors in any dimension are more likely to collaborate. The effect of social proximity is strongest, but also rarest (in our dataset). In other words: if the actors would collaborate completely at random, it would happen even much less that two organisations collaborate twice (or more) than is the case in the FP projects. The size of effect of organisational and geographical proximity is comparable.

Second, multivariate binomial logistic regression shows that there are no interactions in the sense that one dimension becomes insignificant if controlled for the other, even the effect size does not change much. This suggests that the effects of the different dimensions are more or less orthogonal to each other. The fact that there are no interactions between the dimensions does not imply that there are no correlations between organisational background, geographical distance and social embeddedness of collaborators. Our comparison of three groups of collaborators with different organisational compositions shows different distributions across geographical and social proximity. The group of knowledge users is geographically and socially most proximate; the research group is in both dimensions least proximate, while the mixed group – which is organisationally least proximate - is in between the two other groups. In conclusion: the dimensions of proximity do not directly influence each other, but the effect of social and geographical proximity does vary across groups with different organisational backgrounds.

5.5.2 Discussion

Geographical proximity

Geographical proximity is the most studied dimension of proximity. Many studies corroborate our finding that more proximate partners have a higher propensity to collaborate (e.g. Balland 2012; Broekel and Boschma, 2012). Unlike Breschi and Lissoni (2003) and Ter Wal (2009) we have not found that the effect of proximity weakens if one controls for other dimensions. Our finding that although the dimensions do not directly interact with each other, collaborations among organisations from a "research" background are socially and geographically less proximate than collaborations among "knowledge user" organisations is in line with the findings of Scherngell and Barber (2011) that geographical proximity is a much more decisive factor for collaboration in industrial R&D than for public R&D.

Organisational proximity

The finding that organisational proximity has an effect on collaboration is in line with other proximity studies (Balland, 2012; Bouba-Olga et al., 2012). The effect of organisational proximity on collaboration seems surprising in the light of the growing body of literature on the so called triple helix. The central premise of this literature that triadic collaborations among university, government and industry will grow (Etzkowitz and Leydesdorff, 2000) implies that the effect of organisational proximity weakens. Replacing the 7-fold variable for organisational proximity in our original multivariate analysis with a 3-fold according to the triple helix scheme[16] does not alter direction or significance of the effect; the effect size increases slightly[17]. Obviously, some collaborators that were not considered proximate in our original analysis are considered proximate in this scheme.

Social proximity

Our analysis shows that social proximity is a relatively strong predictor of collaboration, and not many organisations are socially proximate. This observation strongly depends on the indicator chosen, which was rather narrow in our analysis: in order to be considered socially proximate collaborators should have a formal collaboration in the sense of shared participation in an earlier FP project (within five years before the collaboration under consideration). This indicator is similar to the operationalization of social proximity in other studies (Autant-Bernard et al., 2007; Broekel and Boschma, 2012; Ter Wal, 2009). A broader indicator of social proximity would show different results: organisations may have collaborated in earlier projects through other national or international funding instruments, or they may have collaborations that are not reflected in joint participation in projects.

Interaction effects

We show that there is no direct interaction between the dimensions of proximity in the sense that the significance and size of effect of the dimensions do not change if controlling for the other dimensions. This is remarkable in the light of earlier literature which does suggest that dimensions of proximity do influence each other. Some studies empirically find that the effect of geographical proximity weakens or even becomes insignificant if controlling for other dimensions (Breschi and Lissoni, 2003; Ter Wal, 2009). However, several other recent proximity studies with multivariate models do not find interaction effects (Broekel and

16 To enable comparisons with the triple helix literature, we have aggregated the seven categories of organisational types used in our analysis into the three groups of the triple helix and tested if that alters our findings. For this analysis, we have used the following classification scheme: University (universities, organisations for higher education, and public research organisations), Government (governmental organisations), Industry (industry, not-for-profit, consultancies, others).

17 The new model gives for organisational proximity: B = .733 (p = .000); e^B = 2.081; Wald = 5678; Nagelkerke pseudo R^2 = .095

Boschma, 2012; Balland, 2012). The difference in findings might be caused by the fact that the studies by Breschi and Lissoni (2003) and Ter Wal (2009) are based on patent information, while the studies by Broekel and Boschma (2012) and Balland (2012) are based on interviews and FP project data respectively. The organisational composition of the studied networks may hence well be different. As we have shown in this study, geographical and social proximity do vary across collaborators of different organisational backgrounds.

Generalizing results

Since the effects of proximity on collaboration vary across research fields (Weterings and Ponds, 2009; Aldieri, 2011), the results of our analysis cannot be generalized to other research fields without complications. Still we believe that some of the relationships we found have a significance that goes beyond the water sector. We elaborate on three: First, the abundance of EU policy and regulation regarding water (such as framework directives on water quality) may stimulate organisations in countries where new knowledge is required to meet the rules and regulations to try and build consortia with organisations in countries that already meet the targets. This may especially have an effect on the geographical proximity of collaborators. The same probably plays a role in other sectors where regulations induce new knowledge, like energy production. Second, in the water sector, there may be other (non-proximity) conditions that play an important role in searching collaborators. It is likely for example that finding collaborators who study or operate in similar physical and geological conditions, as such conditions have a strong impact on the knowledge required. It is hence not only the physical distance that plays a role in collaboration, but also the similarity of the geographical areas where the actors operate. Such an effect is also likely in other environmental sciences where local conditions influence the knowledge agenda. Third, the water sector has a relatively long tradition in collaborating across societal sectors. As many issues interfere directly with public safety, there have long been interactions between governmental actors, research organisations and industry. This may cause a comparative advantage in overcoming low organisational proximity. In other words, other sectors where such developments are relatively new, may face a stronger effect of organisational proximity.

5.5.3 Questions for future research

We close with a few remarks on the agenda for future research in this direction. Our analysis of a cross-sectional set of collaborations in knowledge production gives a large set of observations across a relatively long timeframe, but it does not allow measuring developments over time. Against the background of the (further) development of the European Research Area, it seems promising to also investigate if and how the role of the proximity dimensions has changed over time. Second, it has been suggested that the nature of a research project (explorative, integrative, exploitative) makes a difference for the role of proximity (Balland et al., 2013). It seems worthwhile to analyze whether the growing

emphasis in the FPs on the integrative and exploitative phase of innovation processes has an influence on the proximity patterns. Third, to obtain a broader overview of the collaboration choices that actors make, it would be necessary to combine more sources of data, such as national research funding instruments and more importantly also private collaboration consortia without funding. This would give more insight in the potential interactions between funding instruments, and also give more of an overall picture of collaborative behaviour in knowledge production.

5.6 References

Aguiléra, A., V. Lethiais & A. Rallet (2012). Spatial and Non-spatial Proximities in Inter-firm Relations: An Empirical Analysis. *Industry and Innovation*, 19(3), 187-202.

Aldieri, L. (2011). Technological and geographical proximity effects on knowledge spillovers: evidence from the US patent citations. *Economics of Innovation and New Technology*, 20(6), 597-607.

Augier, M. & M. T. Vendelø,. (1999). Networks, cognition and management of tacit knowledge. *Journal of knowledge management*, 3(4), 252-261.

Autant-Bernard, C., P. Billand, D. Frachisse & N. Massard (2007). Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies*. *Papers in Regional Science*, 86(3), 495-519.

Balland, P.A. (2012). Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies*, 46(6), 741-756.

Balland, P.A., R. Suire & J. Vicente (2013). Structural and geographical patterns of knowledge networks in emerging technological standards: evidence from the European GNSS industry. *Economics of Innovation and New Technology*, 22(1), 47-72.

Barber, M., B. Heller-Schuh, T. Roediger-Schluga & T. Scherngell (2008). NEMO - Network Models, Governance and R&D Collaboration Networks. Deliverable D4.1: The sysres EUPRO database manual.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional Studies*, 39(1), 61-74.

Bouba-Olga, O., M. Ferru & D. Pépin (2012). Exploring spatial features of science-industry partnerships: A study on French data. *Papers in Regional Science*, 91(2), 355-375.

Bozeman, B. & E. Corley (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*, 33(4), 599-616.

Breschi, S. & F. Lissoni (2003). Mobility and social networks: Localised knowledge spillovers revisited. *Milan: University Bocconi, CESPRI Working Paper*(142).

Broekel, T. & R. Boschma (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, 12(2), 409-433.

Chessa, A., A. Morescalchi, F. Pammolli, O. Penner, A. Petersen & M. Riccaboni (2013). Is Europe Evolving Toward an Integrated Research Area? *Science*, 339(6120), 650-651.

Cunningham, S. W. & C. Werker (2012). Proximity and collaboration in European nanotechnology. *Papers in Regional Science*, 91(4), 723-742.

Delanghe, H., U. Muldur & L. Soete (2009). *European science and technology policy: Towards integration or fragmentation?* : Edward Elgar Publishing.

EIP. (2014). *Barriers and bottlenecks for innovation in the water sector*. Brussels: European Innovation Partnership on Water.

Etzkowitz, H. & L. Leydesdorff (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy*, 29(2), 109-123.

Ferru, M. (2010). Formation process and geography of science–industry partnerships: the case of the University of Poitiers. *Industry and Innovation*, 17(6), 531-549.

Fleming, L., C. King & A. I. Juda (2007). Small worlds and regional innovation. *Organization Science*, 18(6), 938-954.

Frenken, K., S. Hardeman & J. Hoekman (2009). Spatial scientometrics: Towards a cumulative research program. *Journal of Informetrics*, 3(3), 222-232.

Hardeman, S. (2012). *The distributed organization of science : with empirical illustrations from the field of diabetes medicine*. Technische Universiteit Eindhoven, Eindhoven.

Heller-Schuh, B., M. Barber, L. Henriques, et al. (2011). JRC scientific and technical reports: Analysis of Networks in European Framework Programmes (1984-2006).

Hessels, L. K. & H. van Lente (2008). Re-thinking new knowledge production: A literature review and a research agenda. *Research Policy*, 37(4), 740-760.

Hoekman, J., K . Frenken & R. J. W. Tijssen (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662-673.

Katz, J. S. & B. R. Martin (1997). What is research collaboration? *Research Policy*, 26(1), 1-18.

Kay, R. & S. Little (1987). Transformations of the explanatory variables in the logistic regression model for binary data. *Biometrika*, 74(3), 495-501.

Knoben, J. & L. Oerlemans (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8(2), 71-89.

Laudel, G. (2001). Collaboration, creativity and rewards: why and how scientists collaborate. *International Journal of Technology Management*, 22(7), 762-781.

Leydesdorff, L. & C.S. Wagner (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2(4), 317-325.

Melin, G. (2000). Pragmatism and self-organization: research collaboration on the individual level. *Research Policy*, 29(1), 31-40.

Narin, F., K. Stevens & E.S. Whitlow (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21(3), 313-323.

Nieminen, M. & E. Kaukonen (2001). Universities and R&D networking in a knowledge-based economy. *A glance at Finnish development. Sitra Reports series, 11.*

Nooteboom, B., W. van Haverbeke, G. Duysters, V. Gilsing & A. van den Oord (2007). Optimal cognitive distance and absorptive capacity. *Research Policy,* 36(7), 1016-1034.

Paier, M. & T. Scherngell (2011). Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model. *Industry and Innovation*, 18(1), 89-104.

Ponds, R., F. van Oort & K. Frenken (2007). The geographical and institutional proximity of research collaboration*. *Papers in Regional Science*, 86(3), 423-443.

Potì, B. & E. Reale (2007). Changing allocation models for public research funding: an empirical exploration based on project funding data. *Science and Public Policy*, 34(6), 417-430.

Rallet, A. & A. Torre (1999). Is geographical proximity necessary in the innovation networks in the era of global economy? *GeoJournal*, 49(4), 373-380.

Scherngell, T. & M. J. Barber (2011). Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU framework programme. *The Annals of Regional Science*, 46(2), 247-266.

Scherngell, T. & R. Lata (2012). Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data*. *Papers in Regional Science*.

Sinnott, R. W. (1984). Virtues of the Haversine. *Sky and telescope*, 68, 158.

Thomas, D. A. & R.R. Ford (2005). *The crisis of innovation in water and waste-water*: Edward Elgar Publishing.

Wal, A. L. J. ter (2009). *The Structure and Dynamics of Knowledge Networks: a Proximity Approach*. (PhD Thesis), Utrecht University, Utrecht.

Wen, B., E. Horlings, M. van der Zouwen, P.van den Besselaar & W. van Vierssen (2011). Mapping science through bibliometric triangulation: an experimental approach applied to water research, *ASIS&T SIG/MET Paper Contest*, May 2011

Weterings, A. & R. Ponds (2009). Do Regional and Non-regional Knowledge Flows Differ? An Empirical Study on Clustered Firms in the Dutch Life Sciences and Computing Services Industry. *Industry and Innovation*, 16(1), 11-31.

6 Conclusions and discussion

6.1 Brief recapitulation

In this thesis I have examined the role of proximity in processes of collaborative knowledge production in the field of water research. The primary research question was: How is collaborative knowledge production in the water field influenced by the geographical, organisational, social and cognitive proximity of the actors involved?

As outlined in the introduction, there were four sub-questions:

1, How has the configuration of the European water knowledge production network developed over time?
2. What explains the variation in the centrality of the actors in the network?
3. What is the influence of proximity on the propensity of actors to collaborate in knowledge production in an applied field of research, water?
4. What is the influence of proximity on the reported outcomes of collaborative knowledge production?

In this final chapter, I first present the conclusions regarding each of these sub-questions and relevant findings that overarch these sub-questions. I then synthesise the results to answer the main question, before presenting recommendations for policy and further elaborating on my contribution to the scientific literature. I will close with a set of questions for future research.

6.2 Conclusions

6.2.1 How has the configuration of the European water knowledge production network developed over time?

The European water knowledge production network has grown strongly over time. Organisations from the first fifteen member states of the European Union are strongly represented in the network. Small countries are better represented in the water research network than in the overall European research network. In contrast to the overall network, the water research network has become organisationally much more diverse over time.

This thesis shows that the European water knowledge production network has grown strongly over time. The network has high clustering coefficients: if an actor collaborates with several other actors, it is likely that these actors will also collaborate with each other. It is a small-world network with many abundant ties, enabling smooth knowledge flows. This is remarkable, given that the water sector in Europe has been described as fragmented and bound within national systems (EIP, 2014; Thomas and Ford, 2005). Organisations from the EU15 (the first fifteen member states of the EU, joined in 1995 or earlier) are strongly

represented in the network. In Framework Programmes (FP) 1-4 (1987-1998), more than 98% of all projects included an EU15 country. While the EU expanded in 2004, welcoming ten new member states, about 90% of FP7 funded projects (starting in 2007) included at least one organisation from the EU15. Meanwhile, the participation of countries outside the EU15 has increased over time: in FP6 (since 2002) and FP7 about 30% of the projects funded had a participant from a new member state (joined the EU after 1995). In FP6 and FP7, small countries – in terms of R&D budget – have a large share in the water network, relative to their participation in the overall network. For large R&D countries such as Germany, the UK and France, participation in FP6 funded projects in the field of water accounts for only 2-3% of their overall FP6 participation, while for most smaller countries this is 5-7%.

Within this group of small countries, two sub-groups can be distinguished. There is a group of countries which have a much larger share in water-related FP projects than in water-related scientific publications, while another group has a relatively large share in both. The latter are probably genuinely specialised in water research, while, for the former, water research is possibly a relatively easily accessible field to join FP projects because it requires less investment in equipment than other fields, or because the knowledge has to be developed anyway to meet the requirements of European legislation. The network has become institutionally more diverse over time. The share of universities and public research organisations has decreased, while there has been a strong increase in the participation of other organisations, in particular from industry. These developments in the water research sector contrast with the overall research network, where the share of industry has decreased over time – although it is still larger than in the water network.

6.2.2 What explains the variation in the centrality of the actors in the network?

At least three factors explain the variation in the centrality of the actors in the network: the type of organisation, the country of origin and the accumulated experience with FP projects.

While the composition of the European research network diversified between 1984 and 2010, the core of the network has remained homogeneous and stable. Public research institutes and universities generally have a higher centrality – they are more embedded in the network and more often form a crucial link between two other organisations. Governmental organisations have a much lower centrality, but still generally higher than organisations from industry. In the overall EU research network, there are also organisations from industry with high centralities, often multinationals that operate across many different research areas (Heller-Schuh et al., 2011); however, in the water network, this is not the case. The organisations with a high centrality are in many cases from an EU15 country, have relatively long experience in FP projects and, more importantly, since their first

participation, they have engaged in numerous new projects in later FPs, in contrast to many organisations with a lower centrality, many of which have participated in only one FP project.

6.2.3 What is the influence of proximity on the propensity of actors to collaborate in knowledge production in an applied field of research, water?

Geographical, organisational and social proximity all have a positive effect on the propensity of actors to collaborate. For geographical and organisational proximity, this effect is very robust across different spatial levels and across collaboration types. The dimensions of proximity do not influence each other directly. Knowledge users are more susceptible to the effect of proximity than knowledge producers.

Geographical, organisational and social proximity have a positive effect on the propensity of actors to collaborate. The positive effect of geographical and organisational proximity on the choice of collaboration partners can be seen across different spatial scales and across different forms of collaborative knowledge production. I found a positive effect in analyses both of collaborations based on co-authored papers as outcome and collaborations with a wider variety of possible outcomes (Chapters 2 and 5) The first analysis was based on the analysis of collaborations, using joint scientific papers as an outcome, and hence it had a focus on academic knowledge producers (Chapter 2), while the second – based on FP project data – considered collaborations with both a wider variety of possible outcomes and a wider variety of organisational backgrounds of participants (Chapter 5). Moreover, the analysis of co-authored papers was at the national level (Chapter 2), while the analysis of FP project data concerned the transnational level (Chapter 5). The findings corroborate existing studies at different spatial scales: from the micro-level of collaboration within a building (Hagstrom, 1965; Kraut and Egido, 1988) to the level of continents (Nagpaul, 2003; Hoekman et al., 2010). In chapter 2 I confirm that this also holds at a spatial scope for which little empirical evidence is available, that of a small country, in this case the Netherlands. This suggests that the difference between large and small countries may be moderated by people's perception of distance, in addition to actual geographical distances. Visualisations of the collaborations within the Netherlands suggest that the tendency to collaborate with proximate partners is not due to more intensive collaborations within demarcated regions. In Chapter 5 I not only analysed geographical and organisational proximity, but also the social dimension of proximity. In our dataset, social proximity occurred the least of all three dimensions, but it had the strongest effect. The odds that two actors actually collaborate increased by a factor of 10 when they were socially proximate.

The effects of the dimensions do not interact with each other directly. Controlling for the other dimensions does not alter the direction, significance or

size of an effect for any of the proximity dimensions on the propensity to collaborate. This corroborates several recent studies (Broekel and Boschma, 2012; Balland, 2012; Cunningham and Werker, 2012). Nevertheless, other studies have found that the effect of geographical proximity becomes weaker when controlling for other dimensions (Breschi and Lissoni, 2003; Ter Wal, 2009).

The fact that there are no direct interactions between the dimensions does not imply that they do not influence each other at all. The combination of organisational background and organisational proximity does matter to geographical and social proximity. In both of the relevant studies (Chapters 2 and 5) I found a remarkable difference in the susceptibility to proximity between knowledge users and knowledge producers. Knowledge users are defined as firms (including consultancies), not-for-profit and governmental organisations: actors whose main role is to apply knowledge. Knowledge producers are defined as universities and research institutes: organisations whose main role is to generate new knowledge. The studies show that knowledge users are more susceptible to the influence of proximity. In the study using co-publications data I found that organisational proximity matters to knowledge users but not to knowledge producers (Chapter 2).

In the analysis of joint project participations I demonstrated that pairs of knowledge users are also more susceptible to the influence of geographical and social proximity (Chapter 5). In relation to geographical proximity, this idea is supported by Ponds et al. (2007). Other studies also suggest that geographical proximity is much more important for public research networks than for private research networks; the concepts of public research and private research are operationalised in such a way that they are similar to our concepts of knowledge users and knowledge producers (Scherngell and Barber, 2011). This is probably also related to an earlier finding that geographical proximity plays a more prominent role in applied research than in basic research because knowledge users are usually more involved in applied research (d'Este and Iammarino, 2010). For applied knowledge, it may make more sense to collaborate with geographically proximate partners, as the research questions are more contextualised and localised. It may also be that knowledge users are less embedded in knowledge production networks, and thus are only aware of organisations that are proximate to them in an organisational and geographical sense.

Proximity has an effect on collaboration choices, both in networks where the focus is on informal collaborations and in networks with formal collaborations and a large share of distant relationships. As I argued in the introduction, one of the contributions of this thesis is that it triangulates findings across different methods and across different types of data, each with a specific scope. The value of such triangulation becomes clear in a comparison of Chapter 3 and Chapter 5. In Chapter 3 I showed that most collaborations are proximate, while I demonstrated in Chapter 5 that many collaborations are not proximate in any of the dimensions tested. The analysis in Chapter 3 was survey based and included

many collaborations for which no explicit (hard) outcomes were reported; moreover, the collaboration in many cases had no formal status (such as being partners in a consortium). The analysis in Chapter 5, on the other hand, was based on joint participation in FP projects. In many cases the FPs give a direct incentive to include distant collaborations, for example by explicit requirements to have participants from a minimum of two countries. In addition, it is inherent to the data that all partners are formal collaborators and at least have the intention to develop explicit outcomes of collaboration. This explains why one dataset contains many proximate relationships, while the other contains many distant relationships. In fact it is more or less analogous to the concepts of 'local buzz' and 'global pipelines'. Local buzz concerns collaborative knowledge production processes in local communities, created simply by being part of the community, and also being largely unplanned. Global pipelines consist of communication channels that are used to select collaborators outside the local community (Bathelt et al., 2004). Both for local buzz, such as the informal collaborations analysed in Chapter 3, and for global pipelines, such as the consortia in Chapter 5, proximity influences the propensity of actors to collaborate.

6.2.4 What is the influence of proximity on the reported outcomes of collaborative knowledge production?

Social and cognitive proximity have a positive effect on the occurrence of reported outcomes of research collaboration for all six outcomes analysed. Geographical and organisational proximity have a negative effect on explicit (hard) outcomes of collaboration, but a weak positive effect (if any) on tacit (soft) outcomes of collaboration. The effect of geographical proximity on explicit outcomes becomes weaker when controlling for the other dimensions of proximity.

Social and cognitive proximity have a positive effect on all of the outcomes examined: more proximate collaborators report these outcomes more often. For geographical and organisational proximity, there is a difference between what I have labelled 'explicit' and 'tacit' outcomes of collaboration. Explicit knowledge is clearly articulated and codified (such as a publication), while tacit knowledge is based on action, experience, or involvement in a specific context (Alavi and Leidner, 2001). In Chapter 3 I have operationalised these abstract concepts using the terms 'hard' (explicit) and 'soft' (tacit) outcomes of collaboration.

Geographical and organisational proximity have a negative effect on explicit outcomes: innovation, joint publications, financial turnover. Organisational proximity has a weak positive effect on soft outcomes: support for ideas, collaboration programmes and shared knowledge. Geographical proximity has no effect on these soft outcomes. The finding that explicit outcomes such as joint publications have a negative correlation with geographical and organisational proximity may seem to contradict the findings of studies that have analysed patterns of joint publications based on bibliographical data (e.g. Hoekman et al., 2010; d'Este and Iammarino, 2010). However, the approach in Chapter 3 differs in an

important way from that of bibliographical studies. Bibliographical studies show – generally speaking – that collaborators are geographically and organisationally more proximate than expected, based on a random distribution. This is also corroborated in my study. However, my survey-based study also showed that, of the collaborations they have, actors report more explicit outcomes from the relatively distant collaborations. This is in line with Weterings and Ponds (2009), who found that although actors generally have more proximate than distant relationships, they consider the distant relationships more valuable (in the sense that the knowledge exchange directly contributes to the problem for which the relationship was established). There is a popular idea in the literature that the relationship between collaboration and proximity takes the form of an inverted U shape, where both too little and too much proximity hinder collaboration (Boschma, 2005). I did not find any evidence for such patterns with regard to reported outcomes, but did find that the optimum for an actor consists of a mix of proximate and distant collaborations. I also tested whether the dimensions of proximity interact with each other (Breschi and Lissoni, 2003; Broekel and Boschma, 2012) and indeed found that – in contrast to the effect on the propensity to collaborate – the effect of geographical proximity on the outcomes of collaboration becomes smaller for the explicit outcomes when controlling for the other dimensions. For two of the three outcomes the effect even disappeared. The effect of social proximity on the tacit outcomes also weakened slightly.

6.3 Main conclusions

Before I proceed to answer my main research question, I will first address a finding that overarches the sub-questions of my research. The conceptual framework made clear that I was essentially testing two relationships between three concepts: how proximity shapes patterns of collaborative knowledge production, and how these patterns influence the reported outcomes of the collaboration process. Thus far, I have treated these relationships separately. Combining the findings on both shows that although geographical and organisational proximity increase the probability of collaboration, they cause a decrease in the probability of delivering explicit outcomes. Actors have proximate relationships which entail mainly tacit outcomes; but they only engage in distant relationships if they expect clear pay-offs in the form of explicit outcomes. Nevertheless, even in the case of explicit outcomes, the collaborators are more proximate than if collaborations would be distributed at random.

There are caveats in comparing the different chapters, as the underlying data have different scopes in time, space and the background of actors. Nevertheless, as I have also elaborated upon above, the finding that geographical and organisational proximity increase the propensity to collaborate is very robust for such scope differences. The finding that explicit outcomes are reported less often in geographically and organisationally proximate collaborations (Chapter 3) can be reasonably expected to also hold at geographical scopes greater than a small country. In terms of local buzz and global pipelines, the former entails many

forms of collaborative knowledge production, some with explicit, most with tacit, outcomes. Collaborations through the global pipelines, however, come with much higher transaction costs and investments. Such collaborations are only initiated if there is a reasonable expectation of pay-offs in the form of explicit outcomes. This was shown in Chapter 3 for the national level, but it is likely that it holds to an even greater degree at European level, as the costs for long distance collaboration would be even higher. This also explains why there is no negative effect from geographical and organisational proximity on tacit outcomes of collaboration: tacit outcomes are probably also reported from largely unplanned and informal collaborations in local communities. Altogether, this may seem to suggest that Chapter 2, which is based on the explicit outcome of joint publications, should conclude that geographical proximity leads to fewer joint papers, which is not the case. This is because Chapter 2 – like most bibliographical studies – shows that collaborators are geographically and organisationally more proximate than what one expects based on a random distribution of collaborations across space. This was also confirmed by Chapter 3, based on the survey data. However, the latter chapter puts that view into a new perspective by showing that of all collaborations people have the relatively distant ones more often result in explicit outcomes.

The main research question of this thesis is: How is collaborative knowledge production in the water field influenced by the geographical, organisational, social and cognitive proximity of the actors involved? The studies in this thesis each shed light on a specific element of this rather broad question. It is now time to combine the insights from the sub-questions to answer the main question, encompassing the findings of all of the constituting studies.

To begin with, there is an emergent knowledge production network on water in Europe that is both geographically and organisationally more diverse than the overall European knowledge production network to which it can be compared. This network contains a stable core of central actors, consisting of universities and research institutes from countries that joined the EU early. Geographical, organisational and social proximity increase the propensity of actors to collaborate. I found no evidence that these three dimensions of proximity interact with each other directly in the sense that they complement or substitute for each other regarding the propensity to collaborate. However, professional roles matter: knowledge users are more susceptible to the effect of proximity than knowledge producers. The relationship between proximity and the reported outcomes of collaboration is a little more complicated. Social and cognitive proximity promote the likelihood of all reported outcomes. Geographical and organisational proximity increase the likelihood of tacit outcomes, such as knowledge sharing (if they have any effect at all), but they decrease the likelihood of explicit outcomes such as joint innovations or joint publications. In contrast to these findings regarding the propensity to collaboration, I did find an interaction between the dimensions in their effect on the outcomes of collabora-

tion: the negative effect of geographical proximity on explicit *outcomes* becomes weaker or disappears when controlling for the other dimensions.

6.4 Policy recommendations

As was explained in the introductory chapter of this thesis, in addition to its academic contribution, the research aims to provide evidence that can be used as guidance in research policy. Collaborative knowledge production and proximity have been major themes in both national and European research policy in recent decades, with particular attention being paid to collaboration across different types of organisations and collaboration across geographical distances and borders. In this section, I recapitulate four main points from the developments in policy, and link them to four recommendations derived from the empirical results of this thesis.

First, funding instruments with large generic components, such as the Framework Programmes, should include more specific measures for different research fields. The introduction of innovation policies in the late 1970s was accompanied by more generic policies for all research fields (Caracostas and Muldur, 2001). The specific features of the water network revealed here illustrate that applied research areas with strategic relevance need tailored incentives and have specific requirements that must be recognised by funding instruments. Nevertheless, the FPs have generic participation rules. The largest element of Horizon 2020 (the recently launched new FP) is the societal challenges pillar, consisting of thematic calls. The calls are categorised into societal challenges. While this seems to leave room for incentives and requirements that are tailored to the configuration of the knowledge production network around that theme or challenge, the rules for participation consist of a generic set of regulations for all societal challenges. Exceptions are possible for each call: all calls in the societal challenge of health, demographic change and wellbeing are open to actors from the US for example. Nevertheless, this is not based on a thorough analysis of the needs and peculiarities of a specific field. Horizon 2020 incorporates new initiatives to ensure wider participation, which mainly aim at improving the participation of low performing RDI regions. In addition, Horizon 2020 aims to create synergies with EU cohesion policy funding for the upgrading of infrastructure and equipment (EC, 2014). This may well contribute to a more balanced participation across regions, but it may overlook the specific needs of research fields.

Previous studies corroborate the need for tailored instruments for specific fields. Hessels (2010) showed that the interaction of researchers with actors from other organisational backgrounds varies across fields. Other studies show that the implications of changing science policies differ across research fields (Bonaccorsi, 2008; Whitley, 2000). Moreover, in the analysis of Chapter 3 I showed that the composition and dynamics of the water knowledge production network differs from the generic research network. There are more small, specialised research organisations in the backbone of the network. In addition, several small and new

member states have a relatively large share in water research. A specific lesson for policy on knowledge production in the water sector in this respect might be to focus less on incentives that increase the involvement of actors from small and new member states (as they seem to be involved anyway), and more on stimulating the potential role of the smaller research organisations in the backbone of the network to improve the network governance, as this is an important aspect of the water network which requires attention (Biswas, 2004).

I also showed in chapters 2 and 5 that organisations which are primarily knowledge users are more susceptible to dimensions of proximity than knowledge producers when it comes to collaboration. The involvement of knowledge users differs across research fields. This may be a reason to give research fields with a high representation of knowledge users additional incentives to ensure that distance does not obstruct the search for relevant partners, for example by organising more networking events where organisations can meet potential project partners.

Second, regulation can be used to create knowledge needs and steer knowledge production. The introduction of innovation policies has brought greater awareness of the broad range of policy instruments available to stimulate innovative performance, including environmental regulation (Tindemans, 2009). My analysis shows that water projects are relatively accessible to new member states of the European Union. One possible explanation for this is that the extensive EU legislation on water, such as the Drinking Water Framework Directive and the Water Framework Directive, has induced a need for the development of new knowledge in this field. Previous studies confirm that new regulations can play a prominent role in inducing knowledge production and innovation, for example in environmental conservation regulation (Den Hertog, 2000; Jaffe and Palmer, 1997). By imposing rules and regulations for the entire European Union, regulation may act not only as an incentive to knowledge production, but also as a catalyst for the participation of new countries, and in this way promote the development of the ERA.

Third, in the case of water research, universities and public research institutes should be stimulated by policymakers to play a stronger role in the governance of the knowledge production network. The policies to stimulate collaboration across organisations from different backgrounds have complicated the governance of the resulting networks, as the division of responsibilities has become less clear (Biswas, 2004). In chapter 3 I showed that the core of the water research network is relatively stable and consists of universities and research institutes: the organisational diversification of the network occurs in the more weakly embedded periphery of the network. The stability and relative uniformity of the core may imply that the most central organisations are best suited to play a stronger role in the governance of the network. A known problem of participatory and integrated networks is that a clear division of responsibilities and accountabilities is lacking,

which prohibits effective and integrative collaboration (Muro and Jeffrey, 2012; Biswas, 2004). The identified stable core of the network may assist in addressing such issues and further developing the network. This is especially important for the water sector, given the need for stronger governance in integrated water management. Across the different contexts worldwide, an important common challenge is the fragmented institutional structure, combined with the unclear allocation of roles and responsibilities (OECD, 2011; OECD 2014). This thesis suggests that for the knowledge production network, policymakers should allocate governance responsibilities to public research institutes and universities, preferably those with longstanding experience in the network. This can be done, for example, by involving them in the governance of the networks that are built through funding instruments aimed at collaborative knowledge production.

Finally, research policies aiming to stimulate collaborative knowledge production by promoting the proximity (or distance) of potential collaborators should take the difference between explicit and tacit research outcomes into account. In more general terms, the focus of policy on collaborative knowledge production is insufficiently based on evidence regarding the outcomes of such collaborations (Velzing, 2013). In chapter 3 I have shown that the effect of proximity can differ for explicit and tacit outcomes of collaboration. Explicit outcomes such as joint papers are more often realised in geographically and organisationally distant collaborations, while tacit outcomes such as the exchange of ideas are more frequent in proximate relationships. This also suggests that science parks and other policy initiatives that promote geographical proximity will only be effective in promoting explicit outcomes of collaboration insofar as they bring together actors who would collaborate anyway but thanks to the policy initiative can do so at lower transaction costs. This finding was corroborated by Gurney (2014), who, in a case study of science parks, showed that there are few joint explicit outcomes from organisations in a science park, while there are many explicit outcomes with collaborators outside the park. It is also in line with the work of Ganesan et al. (2005), who, in relation to innovation and new product development, recommended the development of collaborations with relevant knowledge providers regardless of geographical proximity. My research suggests that stimulating geographical and organisational proximity will contribute to tacit outcomes of collaboration.

6.5 Contribution to the literature

In the introductory chapter of this thesis I have argued that the literature on proximity could be enriched by a multi-level and multi-dimensional approach (studies at different levels of geographical aggregation and studies that test several dimensions of proximity simultaneously). I also showed that there is a need for empirical studies on proximity in applied fields of research with strategic relevance. While there is an extensive body of literature on the object of study, water, and knowledge production in this field, it pays little attention to the

configuration of research networks and factors that shape research collaboration. Below, I describe in more detail what insights this thesis contributes to the existing literature about on proximity and water as a research field. I also elaborate upon the contribution to the Triple Helix literature, which is closely related to the concept of organisational proximity.

Proximity

The contribution of this thesis to the ongoing discussions in the literature on proximity is fivefold and concerns:

- the relevance of proximity at different levels of geographical aggregation;
- proximity in applied fields of research with strategic relevance;
- the difference in the relationships between proximity and outcomes versus proximity and the propensity to collaborate;
- the existence of an inverted U-shaped relationship between proximity and innovative performance;
- possible interactions among the dimensions of proximity.

First, this thesis shows that geographical and organisational proximity in the water sector have a similar effect on the propensity to collaborate at the level of a small country, the Netherlands, and at the transnational level of Europe. To the best of my knowledge, there are no studies of a particular research field that combine findings at several geographical levels. The evidence at the level of a small country is particularly important, because empirical evidence is scarce at this level – with the exception of, for example, the work of Ponds et al. (2007). The finding that the effect of geographical proximity is similar for the Netherlands and Europe suggests that it is not only distance itself that has an effect on the propensity to collaborate but that the perception of distance also plays a role, and that this perception differs between small and large countries.

Second, this thesis contributes to the empirical findings on proximity in applied fields of science with strategic relevance. Previous literature has suggested that the effect of proximity differs for applied and basic research, particularly because the actors in applied research are organisationally more diverse (e.g. d'Este and Iammarino, 2010). This thesis has shown that the water sector is indeed organisationally more diverse than the overall European research network, with a larger representation of knowledge users. It has also been shown that knowledge users are more susceptible to the effect of organisational, geographical and social proximity on the propensity to collaborate.

Third, it has been shown that geographical, organisational and social proximity all contribute to the propensity of actors to collaborate, but that geographical and organisational proximity do not contribute to the likelihood of explicit outcomes being reported. This adds empirical evidence to the suggestion in previous literature that there is a proximity paradox: proximity promotes collabora-

tion among actors, but having proximate collaborations does not necessarily promote the innovative performance of the actors themselves (Boschma and Frenken, 2010). However, for tacit outcomes of collaboration, there is no such paradox: all four dimensions of proximity contribute to reported tacit outcomes. This provides empirical evidence in support of the conceptual ideas of global pipelines and local buzz. Bathelt et al. (2004) suggested that proximity-related factors such as similar language and technology attitudes play a role in local buzz, while global pipelines involve intentional efforts to overcome distances. This thesis shows that proximity shapes collaborations both in local buzz and global pipeline contexts.

Fourth, in this thesis I also contribute to the discourse about optimal levels of proximity. There has been a debate in the proximity literature about the existence of an inverted U shape between proximity and innovative performance (Boschma, 2005; Nooteboom, 1999), where the optimum for innovative performance would lie between too much and too little proximity. When testing this in terms of the relationship between geographical, organisational, social and cognitive proximity on the one hand and the reported outcomes of collaboration on the other, I did not find any evidence for the existence of an inverted U shape, but only linear positive and negative relations. The combination of my findings about the effect of proximity on outcomes of collaboration – greater geographical and organisational proximity is associated with less explicit and more tacit outcomes – and on the propensity to collaborate – social, organisational and geographical proximity promote collaboration – suggests that, from the perspective of the actor, optimal collaborative knowledge production consists of a mix of proximate and distant relations in all dimensions. This corroborates Boschma and Frenken's (2010) notion of the proximity paradox: although proximity drives actors to initiate collaborations, too much proximity may harm their innovative performance. Determining the level that could be considered optimal may depend on the goals and needs of the actors (and hence also on their organisational backgrounds) and on the field of research, where more applied and contextualised fields of knowledge probably benefit from higher levels of proximity than fundamental research fields, given the need to adapt knowledge in applied fields to local conditions and contexts.

Finally, my results do not provide evidence for direct interactions between the dimensions of proximity in relation to the propensity for collaboration. Recently, there has been growing interest in the possibility that the dimensions of proximity do interact with each other. Several recent studies corroborate our finding that there are no direct interaction effects in relation to the propensity for collaboration (Broekel and Boschma, 2012; Cunningham and Werker, 2012). The literature is not unequivocal on this point; other studies find interaction effects between geographical and social proximity (Breschi and Lissoni, 2003; Ter Wal, 2009). However, the combination of organisational background and organisational proximity does interact with other dimensions: pairs of knowledge users are

more susceptible to the effect of social and geographical proximity than knowledge producers or mixed pairs. I do find direct interactions between the proximity effects on reported outcomes. The effect of geographical proximity on reported outcomes weakens or even disappears when controlling for organisational, social and cognitive proximity. Moreover, the effect of social proximity on tacit outcomes is weaker in a multidimensional model. To the best of my knowledge, this is the first study to conduct a multidimensional analysis of the effect of proximity on outcomes of collaboration.

Triple Helix

This thesis contributes to the Triple Helix literature by providing empirical evidence about the organisational diversification of the knowledge production network and by showing the different effects of proximity on explicit and tacit outcomes. The literature about the Triple Helix model is based on the conceptual idea that universities, industry and government are increasingly interdependent in knowledge production. The knowledge infrastructures of many countries are said to converge into a model where these three 'spheres' overlap, take each other's form, and where hybrid organisations can emerge at the interfaces (Etzkowitz and Leydesdorff, 2000). The empirical research in this thesis shows that the organisational background of actors in the water knowledge production network is indeed diversifying. This is not limited to the spheres mentioned, but also includes NGOs. However, the actors with the most central positions form a stable and homogeneous group that consists solely of universities and public research organisations, so the spheres of government and industry are not (yet) represented among the most central actors in the network. Furthermore, in line with the idea of the Triple Helix, this thesis shows that organisationally distant collaborators report more explicit outcomes in their collaboration. This suggests that intensive collaborations between the spheres can contribute to explicit outcomes of collaboration, but that initiatives that aim to eliminate the organisational distance between actors, such as hybrid organisations – new actors situated between two spheres with characteristics of both – may be counterproductive: it may be thanks to the difference in organisational structures, aims and incentives that the actors from the different spheres can jointly realise explicit outcomes.

Water

The contribution of this thesis to the literature on knowledge production in water consists of two elements: I provide evidence for and insight into the configuration of the European knowledge production network, and I demonstrate that EU research policy contributes to overcoming fragmentation of the network. Previous literature states that the water sector consists of a multiplicity of actors, from many different organisational backgrounds, and that collaboration across these actors is lacking (EIP, 2014; OECD, 2011). Despite this, studies that use actor analysis and related approaches that may shed new light on collaborative knowledge production remain scarce in this field. The studies that have been

done often focus on a particular project or region (Hermans, 2005). This thesis contributes systematic evidence on the configuration of the European knowledge production network in the water sector. I have shown that the network is indeed diversifying over time in terms of the organisational backgrounds of the actors. However, I have also shown that the diversifying network has a stable and relatively uniform core. This sheds new light on the ways in which the governance of integrative water networks can be arranged. The homogeneity and stability of the central actors makes them relatively easy to address. They may therefore act as catalysts to promote integrative knowledge production in the entire network.

This relates to the second point. This thesis explains how dimensions of proximity shape the patterns of collaboration, revealing that most collaborators in the European knowledge production network are not organisationally proximate. This is remarkable, given that previous literature observed a high degree of organisational and institutional fragmentation (Thomas and Ford, 2005; EIP, 2014; OECD, 2011). This perceived fragmentation has led to calls for a more integratative approach in water management. In relation to knowledge production in particular this has resulted in attempts to develop more participatory approaches, where actors from different backgrounds develop new knowledge together (Pahl-Wostl, 2007; Frijns et al., 2013). My finding that most collaborators in the FP projects are not organisationally proximate suggests that EU research policy is already contributing to the overcoming of organisational fragmentation and promoting integrative approaches.

6.6 Future research questions

I close with a set of questions and topics that deserve attention in future research.

First, the dynamic character of proximity is almost a blind spot in the literature to date. I have demonstrated in Chapter 4 that the network has changed over time, with many new actors becoming involved, and the water network has diversified in terms of both the national and organisational backgrounds of the participants. Moreover, the proximity of collaborators is probably not static. Especially for social and cognitive proximity, one can even argue that it can be built up by collaborating. At the same time, actors may become less susceptible to the effect of proximity in their choice of collaborators should the introduction of modern IT technologies substantially decrease the costs of overcoming distance. There is, however, no empirical evidence regarding this phenomenon (Hoekman et al., 2010; Chessa et al., 2013; Bouba-Olga et al., 2012). Future research could build on the conceptual idea that proximity and collaborations co-evolve and influence each other (Balland, 2012; Ter Wal, 2009). This is not only important for the research on the relationship between proximity and choice of collaborators, but also for research on the relationship between proximity and reported outcomes. Here, as well, the direction of causality could be questioned: the proximity between collaborators influences the expected outcomes of their

collaboration, but if collaborators have jointly realised specific outcomes, this may in turn increase their (perceived) proximity.

Second, more attention needs to be paid to the influence of actor-specific characteristics on the role of proximity. I have shown that there is a difference between organisations that can be characterised as knowledge users and those that are knowledge producers, and that proximity has a greater effect on the propensity of knowledge users to collaborate. It is likely that there are other factors that determine to what extent an actor is susceptible to the influence of proximity in collaborative knowledge production; for example, factors such as the size of an organisation, its age, in-house facilities and the capacity it has to develop consortia and collaborations (universities, for example, often have their own liaison office with dedicated means to initiate projects, write proposals and apply for grants). First steps in this direction suggest – at least for geographical proximity – that organisation size and academic research quality play a role in collaborations between firms and universities, and that there are differences between disciplines (d'Este and Iammarino, 2010; Bouba-Olga et al., 2012). Including such factors in a model will not only provide a better insight into how much of the variance in the propensity to collaborate should be attributed to proximity, but will also help in explaining how susceptible an organisation is to the effect of proximity. This, in turn, may contribute to developing tailored policies for research fields.

Third, it would be very useful to the further interpretation of the concept of proximity to further investigate the reasons why actors collaborate with proximate or distant partners. Is there a kind of trade-off between higher transaction costs and more relevant collaborators? Are distant partners chosen more deliberately than proximate collaborators? My analysis shows that although geographical, organisational and social proximity increase the propensity for collaboration, explicit outcomes of collaboration are less likely in geographically and organisationally proximate collaborations. This raises the question of what factors are decisive in the perception of the actors themselves when selecting a specific collaborator. With regard to the geographical dimension, Hoekman et al. (2009) found suggestions of what they called 'elite structures', where actors from economically strong regions tend to collaborate. For publications in economic top journals it has been shown that the quality of the organisation of the co-author is an important factor in collaboration (Sutter and Kocher, 2004). Many studies also find a bias to domestic collaborations (Frenken et al., 2009). In this respect, it also seems promising to collect information on all the collaborations a specific actor is involved in, to obtain an overview of all the collaboration choices they make. This would also give us more insight into the extent to which the collaborations of an actor arising from a specific funding instrument differ, in terms of proximity, from their other collaborations.

Fourth, it would contribute to both research and policy purposes to analyse the

effects of the growing intensity of collaborations at the system level. Is there, for example, a convergence of research topics as a result of more intensive collaborations across borders within Europe? Or does each region still focus on its own research strengths? I have shown that there has been a diversification in the national and organisational backgrounds of the actors on the periphery of the network, but the core has remained stable and consists of research organisations and universities from 'old' EU countries. It is unknown whether the new entrants have brought in their own research topics and cognitive specialisations, or whether they only build on the research strengths of the stable core. Langfeldt et al. (2012) have provided an interesting basis for such research questions with their analysis of the interactions between Norwegian and European priority setting in research. Also, the recently developed toolboxes to generate cognitive 'overlay' maps of scientific outputs against a benchmark are a promising start in this regard (Rafols et al., 2010). With respect to the European research network, Barber and Scherngell (2013) showed how research communities can be distinguished. Policy initiatives such as *smart specialisation* – the idea that regions in Europe should choose specific thematic specialisations, meaning that not every region or nation needs to cover all research areas – would benefit enormously from more evidence on the cognitive convergence that may result from collaborative knowledge production.

6.7 References

Alavi, M. & D. E. Leidner (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, 107-136.

Balland, P.A. (2012). Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies*, 46(6), 741-756.

Barber, M. J. & T. Scherngell (2013). Is the European R&D network homogeneous? Distinguishing relevant network communities using graph theoretic and spatial interaction modelling approaches. *Regional Studies*, 47(8), 1283-1298.

Bathelt, H., A. Malmberg & P. Maskell (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31-56. doi: 10.1191/0309132504ph469oa

Biswas, A. K. (2004). Integrated water resources management: a reassessment: a water forum contribution. *Water international*, 29(2), 248-256.

Bonaccorsi, A. (2008). Search regimes and the industrial dynamics of science. *Minerva*, 46(3), 285-315.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional Studies*, 39(1), 61-74.

Boschma, R. & K. Frenken (2010). The spatial evolution of innovation networks. A proximity perspective. In R. Boschma & R. Martin (Eds.), *Handbook of evolutionary economic geography* (pp. 120-135). Cheltenham, UK: Edward Elgar Publishing.

Bouba-Olga, O., M. Ferru & D. Pépin (2012). Exploring spatial features of science-industry partnerships: A study on French data. *Papers in Regional Science*, 91(2), 355-375.

Breschi, S. & F. Lissoni (2003). Mobility and social networks: Localised knowledge spillovers revisited. *Milan: University Bocconi, CESPRI Working Paper* (142).

Broekel, T. & R. Boschma (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, 12(2), 409-433.

Caracostas, P. & U. Muldur (2001). The Emergence of a new European Union Research and Innovation Policy. In Laredo, P. & P. Mustar (Eds.), *Research and Innovation Policies in the New Global Economy*. Cheltenham, UK: Edward Elgar Publishing.

Chessa, A., A. Morescalchi, F. Pammolli, O. Penner, A. Petersen & M. Riccaboni (2013). Is Europe Evolving Toward an Integrated Research Area? *Science*, 339(6120), 650-651.

Cunningham, S. W. & C. Werker (2012). Proximity and collaboration in European nanotechnology. *Papers in Regional Science*, 91(4), 723-742.

EC. (2014). *Enabling synergies between European Structural and Investment Funds, Horizon 2020 and other research, innovation and competitiveness-related Union programmes: Guidance for policy-makers and implementing bodies*. Brussels: European Commission.

EIP. (2014). *Barriers and bottlenecks for innovation in the water sector*. Brussels: European Innovation Partnership on Water.

Este, P. d' & S. Iammarino (2010). The spatial profile of university-business research partnerships. *Papers in Regional Science*, 89(2), 335-350.

Etzkowitz, H. & L. Leydesdorff (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy*, 29(2), 109-123.

Frenken, K., S. Hardeman & J. Hoekman (2009). Spatial scientometrics: Towards a cumulative research program. *Journal of Informetrics*, 3(3), 222-232.

Frijns, J., C. Büscher, A. Segrave & M. van der Zouwen (2013). Dealing with future challenges: a social learning alliance in the Dutch water sector. *Water Policy*, 15(2), 212-222.

Ganesan, S., A. J Malter & A. Rindfleisch (2005). Does distance still matter? Geographic proximity and new product development. *Journal of Marketing*, 69(4), 44-60.

Gurney, T. (2014). *The Intellectual Salmon Run: Knowledge Transfer and Dynamics between Academia and Industry*. Den Haag: Rathenau Instituut.

Hagstrom, W. (1965). *The Scientific Community*. New York: Basic Books.

Heller-Schuh, B., M. Barber, L. Henriques et al. (2011). JRC scientific and technical reports: Analysis of Networks in European Framework Programmes (1984-2006).

Hermans, L. (2005). *Actor analysis for water resources management: putting the promise into practice*: Eburon Uitgeverij BV.

Hertog, P. de (2000). Knowledge-intensive business services as co-producers of innovation. *International Journal of Innovation Management*, 4(04), 491-528.

Hessels, L. (2010). *Science and the struggle for relevance*. (PhD Thesis), Utrecht University, Utrecht.

Hoekman, J., Frenken, K. & Tijssen, R. J. W. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662-673.

Hoekman, J., K. Frenken & F. van Oort (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43(3), 721-738.

Jaffe, A. B. & K. Palmer (1997). Environmental regulation and innovation: a panel data study. *Review of Economics and Statistics*, 79(4), 610-619.

Kraut, R., C. Egido & J. Galegher (1988). *Patterns of contact and communication in scientific research collaboration*.

Langfeldt, L., Godø, H., Gornitzka, Å. & Kaloudis, A. (2012). Integration modes in EU research: Centrifugality versus coordination of national research policies. *Science and Public Policy*, 39(1), 88-98.

Muro, M. & P. Jeffrey (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.

Nagpaul, P. (2003). Exploring a pseudo-regression model of transnational cooperation in science. *Scientometrics*, 56(3), 403-416.

Nooteboom, B. (1999). Innovation, learning and industrial organisation. *Cambridge Journal of Economics*, 23(2), 127-150.

OECD. (2011). Water Governance in OECD Countries: A Multi-level Approach. *OECD Studies on Water*.

OECD. (2014). Water Governance in the Netherlands: Fit for the Future? *OECD Studies on Water*.

Pahl-Wostl, C., M. Craps, A. Dewulf et al. (2007). Social learning and water resources management. *Ecology and society*, 12(2), 5.

Ponds, R., F. van Oort & K. Frenken (2007). The geographical and institutional proximity of research collaboration*. *Papers in Regional Science*, 86(3), 423-443.

Rafols, I., A. L Porter & L. Leydesdorff (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 1871-1887.

Scherngell, T. & M.J. Barber (2011). Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU framework programme. *The Annals of Regional Science*, 46(2), 247-266.

Sharp, M. (1998). Competitiveness and cohesion—are the two compatible? *Research Policy*, 27(6), 569-588.

Sutter, M. & M. Kocher (2004). Patterns of co-authorship among economics departments in the USA. *Applied Economics*, 36(4), 327-333.

Thomas, D. A. & R.R. Ford (2005). *The crisis of innovation in water and waste-water*: Edward Elgar Publishing.

Tindemans, P. (2009). Post-war research, education and innovation policy-making in Europe. In H. Delanghe, U. Muldur & L. Soete (Eds.), *European Science and Technology Policy* (pp. 3-23). Cheltenham, UK: Edward Elgar Publishing.

Velzing, E.J. (2013). *Innovatiepolitiek: Een reconstructie van het innovatiebeleid van het ministerie van Economische Zaken van 1976 tot en met 2010*: Eburon.

Wal, A. L. J. ter (2009). *The Structure and Dynamics of Knowledge Networks: a Proximity Approach*. (PhD Thesis), Utrecht University, Utrecht.

Weterings, A. & R. Ponds (2009). Do Regional and Non-regional Knowledge Flows Differ? An Empirical Study on Clustered Firms in the Dutch Life Sciences and Computing Services Industry. *Industry and Innovation*, 16(1), 11-31.

Whitley, R. (2000). *The Intellectual And Social Organisation Of The Sciences* (Second Edition ed.). Oxford: Oxford University Press.

Summary

The question if distance matters for collaborative knowledge production has intrigued scholars already for a long time. The question has become all the more urgent as knowledge production has increasingly become a collaborative process, not only among individuals but also among organisations. Both in the scientific literature and in the policy discourse there is a strong interest for this phenomenon. The literature on the role of distance first focused on geographical distance, but in the past two decades it has broadened its scope to other dimensions of distance (or proximity) such as an organisational, cognitive and social dimension. The central premise of this literature is that more proximate people have a tendency to collaborate, as it easier and less costly to collaborate with close partners. On the other hand, if people become too proximate, this may lead to lock-in and lower the innovation potential of collaboration.

Despite the growing body of literature on this topic, there are important questions that are still left unanswered. I have identified a combination of three aspects that together form a niche in the literature. First, there is a need for empirical analysis of strategic and applied fields of research. Second, there is the question if proximity has the same effect on the propensity of collaboration across different spatial levels of aggregation. Third, recent contributions to the literature prove the relevance of analyses that include more dimensions of proximity. This also sheds light on potential interactions among these dimensions.

My thesis contributes to filling gaps in this this niche, with a multi-level and multi-dimensional approach that analyses the effect of different dimensions of proximity at several geographical levels of aggregation. I have chosen the water sector as object of study. Water is an applied field of research, where knowledge is often contextualised for specific local conditions. It is also a field with large strategic relevance, also given the prominent role of water in several grand societal challenges. Water is an interesting field for proximity research because of its high organisational and cognitive diversity, and an interesting geographical configuration as water management is organised in national systems, but the challenges are typically specific to local environmental conditions, but not bound by administrative and cultural borders.

This thesis also aims to provide evidence and guidance to policies for collaborative knowledge production. Collaborative knowledge production and the dimensions of proximity analysed in this thesis feature very prominently in the national and transnational research policies of the past decades, especially in Europe. It is all the more remarkable that are blind spots in the knowledge about the effect of proximity on collaborative knowledge production. This thesis contributes to filling that gap as it provides additional evidence regarding the effect of proximity on collaborative knowledge production.

This brought me to the following research question: how is collaborative knowledge production in the water field influenced by geographical, organisational, social and cognitive proximity of the actors involved?

In order to answer this relatively broad question, I have disentangled and addressed several underlying concepts. This has resulted in four sub-questions:

1. How has the configuration of the European water knowledge production network developed over time?
2. What explains the variation in centrality of the actors in the network?
3. What is the influence of proximity on the propensity of actors to collaborate in knowledge production in an applied field of research, water?
4. What is the influence of proximity on the reported outcomes of collaborative knowledge production?

I have addressed these questions in four empirical chapters. In chapter 2 I start to address question 3. I fit a gravity model, based on a sample of 2247 publications on water research from the ISI Web of Science. I analyse the impact of geographical and organisational proximity on knowledge production in the water sector in a small country, the Netherlands. In chapter 3 I answer question 4, examining the effect of geographical, social, organisational and cognitive proximity on knowledge-related outcomes of collaboration. I use data from a survey among water professionals in the Netherlands. The set contains complete data on 1020 ego-alter relationships, in a heterogeneous population with people from a wide variety of organisational backgrounds. In chapter 4 I address the first two questions, analysing the configuration of the European knowledge production network on water by applying social network analysis to data from projects in the European Union's Framework Programme 1 to 7. In chapter 5 I extend the answer on question 3. I investigate the effect of geographical, organisational and social proximity on the propensity of organisations to collaborate. I apply both multivariate and bivariate analyses on data from water research projects in Framework Programme 1 to 7. Below I present the main conclusions regarding each of these sub-questions.

How has the configuration of the European water knowledge production network developed over time?

The European water knowledge production network has grown strongly over time. Organisations from the first fifteen member states of the European Union have a strong representation in the network; the number of projects without a partner from specifically one of these countries increases only very slowly. Still, small countries – in terms of overall R&D budget and number of scientific publications on water – are better represented in the water research network than in the overall European research network. In contrast to the overall network, the water research network has become organisationally much more diverse over time.

What explains the variation in centrality of the actors in the network?
At least three factors explain the variation in centrality of the actors in the network. First, public research institutes and universities generally have a much higher centrality than organisations from other types. Second, organisations with a higher centrality are often from an EU15 country. Third, organisations with a higher centrality have so far continued joining new projects since the first time they participated, whereas many organisations with a lower centrality joined an FP project only once.

What is the influence of proximity on the propensity of actors to collaborate in knowledge production in an applied field of research, water?
Geographical, organisational and social proximity all three have a positive effect on the propensity of actors to collaborate. For geographical and organisational proximity, this effect is very robust across different spatial levels (national and transnational) and across collaboration types (co-authored publications and joint project participation). The dimensions of proximity do not influence each other directly. Knowledge users are more susceptible to the effect of the dimensions of proximity than knowledge producers.

What is the influence of proximity on the reported outcomes of collaborative knowledge production?
Social and cognitive proximity have a positive effect on the occurrence of reported outcomes of research collaboration for all six outcomes analysed: socially and cognitively proximate collaborators report all six outcomes more often. Geographical and organisational proximity have a negative effect on explicit (hard) outcomes of collaboration, but a weak positive effect (if any) on tacit (soft) outcomes of collaboration. The effect of geographical proximity on explicit outcomes becomes weaker if controlling for the other dimensions of proximity.

There is a hence remarkable difference between the effect of proximity on the propensity to collaborate and on the reported outcomes. While geographical, organisational and social proximity promote the propensity to collaborate, geographical and organisational proximity lead to less reported explicit outcomes of collaboration. This suggests the need for a mix of distant and proximate collaborations. Combining the chapters also shows that the effect of geographical and organisational proximity on the propensity to collaborate is very robust for differences in the spatial scope and the composition of the network. Proximity plays a role both in networks where the focus is on informal collaborations and in networks with formal collaborations and a large share of distant relationships.

Based on these main conclusions I formulate four implications for policy. Specifically, I recommend that:

- Funding instruments with large generic components that aim to build up and strengthen collaborative research networks –whether at

national or at European level– should include tailored measures for specific research fields and for actors of various organisational backgrounds, distinguishing knowledge users and knowledge producers;
- Governments and policy makers should use regulation outside the field of research policy –such as environmental regulation– to create knowledge needs and steer knowledge production, regulation has the potential to induce knowledge production but also act as a catalyst for the participation of new countries ;
- In knowledge production networks with a stable and homogeneous core such as water, policy makers should address this core should be addressed to further develop the network and improve the governance of the collaborations in the network;
- If research policies for collaborative knowledge production promote or discourage the proximity of collaborators, they should be aware that geographical and organisational proximity may result in more tacit outcomes but at the same time may decrease the number of explicit outcomes of collaboration. Initiatives that stimulate geographical proximity, such as science parks, are probably only effective in promoting explicit outcomes in the sense that they lower the transaction costs for actors that would jointly realise such outcomes anyway.

This thesis contributes to several bodies of literature. The contribution to the literature on proximity is fivefold. First, this thesis shows that the effect of geographical and organisational proximity in the water sector is similar at national and transnational level. Second, it contributes to the empirical evidence of proximity in applied fields of research with strategic relevance. It specifically shows that the water network is organisationally more diverse than the overall European knowledge production network, and that knowledge users are more susceptible to the effect of proximity on collaboration than knowledge producers. Third, I show that geographical, organisational and social proximity all three contribute to the propensity of actors to collaborate, but that geographical and organisational proximity cause a decrease in the possibility that collaborators report outcomes with explicit knowledge in their collaboration. Fourth, my research suggests that from the perspective of the actor the optimal collaboration network consists of a mix of distant and proximate collaborations, at least in the geographical and organisational dimension. This contributes to the idea of a proximity paradox that although proximity is a driver of innovation, too much proximity may harm the innovative performance of the collaborators. Fifth, I contribute to the discussion about interactions between the dimensions of proximity. I find no evidence for direct interactions in the effect on the propensity of collaboration, but the combination of organisational proximity and organisational background has an indirect interaction effect: pairs of knowledge users are more susceptible to the effect of social and geographical proximity than knowledge producers and mixed pairs. I also do find that the effect of geographical proximity on explicit

outcomes of collaboration weakens when controlling for the other dimensions. My research contributes to the Triple Helix literature – which is closely related to the concept of organisational proximity – because it shows that the organisational background of actors in the water sectors is indeed diversifying; however the actors with the most central positions in the network are all universities and public research institutes. I show that organisationally distant collaborators report more explicit outcomes. This corroborates the triple helix idea that intensive collaborations among government, industry and universities can contribute to knowledge production, but also that overlap between these spheres and hybrid organisations might work counterproductive.

I contribute to the body of literature on knowledge production in the water sector by providing evidence about the configuration of the European water knowledge production network. I show that it is organisationally more diverse than the overall network, and that countries with small R&D budgets have a relatively strong representation in the FP knowledge production network on water. I also show that the EU research policies are effective in overcoming the perceived organisational fragmentation in the sector and promoting integrative approaches.

Nederlandse samenvatting

De vraag of afstand bepalend is voor het al dan niet gezamenlijk produceren van kennis, houdt onderzoekers al geruime tijd bezig. De vraag is alleen maar urgenter geworden omdat kennisproductie in toenemende mate een gezamenlijk proces is geworden, niet alleen van individuen, maar ook van organisaties. Zowel in de wetenschappelijke literatuur als in het beleidsdiscours is er een sterke belangstelling voor dit fenomeen. De literatuur over het belang van afstand richtte zich in eerste instantie op geografische afstand, maar in de afgelopen twee decennia is de blik verbreed naar andere dimensies van afstand (of nabijheid), zoals een organisatorische, cognitieve en sociale dimensie. De gemeenschappelijke premisse in de literatuur over dit onderwerp is dat mensen die zich in elkaars nabijheid bevinden, de neiging hebben om onderling samen te werken, aangezien het eenvoudiger en minder kostbaar is om samen te werken met nabije partners. Aan de andere kant, als mensen te dicht bij elkaar komen, kan dit leiden tot *lock-in* en het potentieel om gezamenlijk te innoveren verkleinen.

Ondanks de groeiende hoeveelheid literatuur over dit onderwerp, resteren er nog belangrijke vragen die onbeantwoord zijn. Ik heb een combinatie van drie aspecten geïdentificeerd die een gezamenlijke niche in de literatuur vormen. Ten eerste is er behoefte aan empirische analyses van strategische en toegepaste velden van onderzoek. Ten tweede is er de vraag of nabijheid hetzelfde effect heeft op verschillende geografische aggregatieniveaus. Ten derde laten recente bijdragen aan de literatuur de relevantie zien van analyses die meerdere dimensies van nabijheid in één model testen. Dit draagt ook bij aan inzichten over mogelijke interacties tussen de verschillende dimensies van nabijheid.

Mijn proefschrift draagt bij aan het opvullen van leemtes in deze niche, met een multi-dimensionale benadering op meerdere niveaus die het effect van meerdere dimensies van nabijheid toetst op verschillende geografische aggregatieniveaus. Ik heb de watersector gekozen als studieobject. Water is een toegepast onderzoeksveld, waar kennis wordt gecontextualiseerd voor specifieke plaatselijke condities. Het is ook een veld met een grote strategische relevantie, mede vanwege de prominente rol van water in diverse grote maatschappelijke uitdagingen. Water is een interessant veld voor het onderzoek naar nabijheid vanwege de hoge organisatorische en cognitieve diversiteit, en vanwege de interessante geografische configuratie, omdat waterbeheer doorgaans is georganiseerd in nationale systemen, terwijl de uitdagingen doorgaans specifiek zijn voor eigenschappen van de lokale omgeving, maar niet gebonden aan bestuurlijke en culturele grenzen.

Dit proefschrift beoogt ook om empirisch bewijs te leveren om het beleid voor gezamenlijke kennisproductie te voeden. Gezamenlijke kennisproductie en de

dimensies van nabijheid die geanalyseerd worden in dit proefschrift hebben een zeer prominente rol in het nationale en transnationale onderzoeksbeleid van de afgelopen decennia, met name in Europa. Het is des te opmerkelijker dat er blinde vlekken zijn in de kennis over het effect van nabijheid op gezamenlijke kennisproductie. Dit proefschrift draagt eraan bij die leemte te vullen door het leveren van meer empirisch bewijs aangaande het effect van nabijheid op gezamenlijke kennisproductie.

Dit leidde mij tot de volgende onderzoeksvraag: hoe wordt gezamenlijke kennisproductie in de watersector beïnvloed door geografische, organisatorische, sociale en cognitieve nabijheid van de betrokken actoren?

Om deze relatief brede vraag te kunnen beantwoorden, heb vier deelvragen geformuleerd die elk een aspect van de vraag belichten:

1. Hoe heeft de configuratie van het Europese water kennisproductienetwerk zich ontwikkeld over de jaren heen?
2. Wat verklaart de variatie in de centraliteit van de actoren in het netwerk?
3. Wat is de invloed van nabijheid op de geneigdheid van actoren om samen te werken in het produceren van kennis in een toegepast onderzoeksgebied, namelijk water?
4. Wat is de invloed van nabijheid op de gerapporteerde uitkomsten van gezamenlijke kennisproductie?

Ik heb deze vragen geadresseerd in vier empirische hoofdstukken. In hoofdstuk 2 maak ik een aanvang met het beantwoorden van vraag 3. Ik stel een zwaartekrachtmodel op, gebaseerd op een *sample* van 2247 publicaties over water-gerelateerd onderzoek, uit het *ISI Web of Science*. Ik analyseer de invloed van geografische en organisatorische nabijheid op gezamenlijke kennisproductie in de watersector in een klein land, Nederland. In hoofdstuk 3 beantwoord ik vraag 4, door het effect te toetsen van geografische, sociale, organisatorische en cognitieve nabijheid op kennis-gerelateerde uitkomsten van samenwerking. Ik gebruik daarvoor data van een enquête onder professionals uit de Nederlandse watersector. De dataset omvat complete gegevens over 1020 ego-alter relaties, in een heterogene populatie van mensen met een grote diversiteit aan organisatorische achtergronden. In hoofdstuk 4 beantwoord ik de eerste twee vragen. Ik analyseer de configuratie van het Europese kennisproductienetwerk door sociale-netwerkanalyse toe te passen op data over projecten uit de Europese Kaderprogramma's 1 tot en met 7. In hoofdstuk 5 breid ik het antwoord op vraag 3 verder uit. Ik onderzoek het effect van geografische, organisatorische en sociale nabijheid op de geneigdheid van organisaties om samen te werken. Ik pas daarvoor multivariate en bivariate analyses toe op data van wateronderzoeksprojecten in Kaderprogramma 1 tot en met 7. Hierna beschrijf ik de belangrijkste conclusies op elk van de deelvragen.

Hoe heeft de configuratie van het Europese water kennisproductie-netwerk zich ontwikkeld over de jaren heen?
Het Europese wateronderzoeksnetwerk is door de jaren heen sterk gegroeid. Organisaties uit de eerste vijftien lidstaten van de Europese Unie zijn sterk vertegenwoordigd in het netwerk; het aantal projecten zonder een partner uit een van die lidstaten stijgt slechts heel langzaam. Toch zijn kleine landen – in termen van algeheel R&D budget en aantal wetenschappelijke publicaties op het gebied van water – beter vertegenwoordigd in het wateronderzoeksnetwerk dan in het algehele Europese onderzoeksnetwerk. In tegenstelling tot het algehele netwerk is het wateronderzoeksnetwerk organisatorisch gezien veel diverser geworden door de jaren heen.

Wat verklaart de variatie in de centraliteit van de actoren in het netwerk?
Ten minste drie factoren verklaren de variatie in de centraliteit van de actoren in het netwerk. Ten eerste hebben publieke onderzoeksorganisaties en universiteiten doorgaans een veel hogere centraliteit dan andere organisatietypes. Ten tweede zijn organisaties met een hoge centraliteit vaak afkomstig uit een van de eerste vijftien lidstaten. Ten derde hebben organisaties met een hoge centraliteit steeds nieuwe projecten gestart sinds de eerste keer dat ze hebben meegedaan, terwijl veel organisaties met een lage centraliteit maar één keer in een project hebben geparticipeerd.

Wat is de invloed van nabijheid op de geneigdheid van actoren om samen te werken in het produceren van kennis in een toegepast onderzoeksgebied, namelijk water?
Geografische, organisatorische, en sociale nabijheid hebben elk een positief effect op de geneigdheid van actoren om samen te werken. In het geval van geografische en organisatorische nabijheid is dat effect zeer robuust voor verschillende ruimtelijke niveaus (nationaal en transnationaal) en voor verschillende vormen van samenwerking (gezamenlijk geschreven publicaties en gezamenlijke participatie in projecten). De dimensies van nabijheid hebben geen directe invloed op elkaar. Kennisgebruikers zijn vatbaarder voor het effect van de dimensies van nabijheid dan kennisproducenten.

Wat is de invloed van nabijheid op de gerapporteerde uitkomsten van gezamenlijke kennisproductie?
Sociale en cognitieve nabijheid hebben een positief effect op het vóórkomen van gerapporteerde uitkomsten van onderzoekssamenwerking, voor alle zes geanalyseerde uitkomsten: actoren die samenwerken en sociaal en cognitief nabij zijn, rapporteren alle zes uitkomsten vaker. Geografische en organisatorische nabijheid hebben een negatief effect op expliciete (harde) uitkomsten van samenwerking, maar een zwak positief effect (en in sommige gevallen geen effect) op ontastbare (zachte) uitkomsten van samenwerking. Het effect van geografische nabijheid op expliciete uitkomsten is zwakker wanneer het wordt gecorrigeerd voor de andere dimensies van nabijheid.

Er is dus een opmerkelijk verschil tussen het effect van nabijheid op de geneigdheid om samen te werken en op de gerapporteerde uitkomsten van samenwerking. Terwijl geografische, organisatorische en sociale nabijheid de geneigdheid om samen te werken bevorderen, leiden geografische en organisatorische nabijheid ook tot het minder vaak rapporteren van expliciete uitkomsten van samenwerking. Dit suggereert de noodzaak van een mix van verre en nabije samenwerkingen. Het combineren van de bevindingen uit de verschillende hoofdstukken laat ook zien dat het effect van geografische en organisatorische nabijheid op de geneigdheid om samen te werken heel robuust is voor verschillen in de ruimtelijke schaal en de samenstelling van het netwerk. Nabijheid speelt zowel een rol in netwerken waar de nadruk ligt op informele samenwerking als in netwerken met formele samenwerkingen en een groot aandeel van verre samenwerkingen.

Gebaseerd op deze hoofdconclusies formuleer ik vier implicaties voor beleid. Ik beveel in het bijzonder aan dat:

- Financieringsinstrumenten met grote generieke componenten, die erop gericht zijn om onderzoeksnetwerken op te bouwen en te versterken – of dat nu op nationaal of op Europees niveau is – zouden op maat gemaakte maatregelen moeten bevatten voor specifieke onderzoeksvelden en voor actoren van verschillende organisatorische achtergronden, met een onderscheid tussen kennisproducenten en kennisgebruikers.
- Overheden en beleidsmakers dienen regulering buiten het veld van onderzoeksbeleid – zoals bijvoorbeeld regulering op milieugebied – te gebruiken om kennisbehoeften te creëren en kennisproductie te sturen. Regulering heeft het potentieel om kennisproductie te induceren maar ook om als katalysator te fungeren voor de participatie van nieuwe landen.
- In kennisproductienetwerken met een stabiele en homogene kern zoals water, moeten beleidsmakers deze kern inschakelen om het netwerk verder te ontwikkelen en de aansturing van de samenwerkingen in het netwerk te verbeteren.
- Als onderzoeksbeleidsmaatregelen voor gezamenlijke kennisproductie de nabijheid van samenwerkende actoren bevorderen of juist ontmoedigen, moeten de ontwerpers van beleid zich er bewust van zijn dat geografische en organisatorische nabijheid kan resulteren in meer ontastbare uitkomsten, maar tegelijkertijd ook kan leiden tot een afname in van expliciete uitkomsten van samenwerking. Initiatieven die geografische nabijheid stimuleren (zoals *science parks*) zijn waarschijnlijk uitsluitend effectief in het bevorderen van expliciete uitkomsten in die zin dat ze de transactiekosten verlagen voor actoren die toch al zulke uitkomsten zouden realiseren.

Mijn onderzoek draagt bij aan verschillende bestaande literatuurstromingen. De bijdrage aan de literatuur over nabijheid is vijfledig. Ten eerste laat dit proefschrift zien dat het effect van geografische en organisatorische nabijheid in de watersector op het nationale en transnationale niveau vergelijkbaar is. Ten tweede draagt het proefschrift bij aan het empirische bewijs over de rol van nabijheid in toegepaste onderzoeksgebieden met een strategische relevantie. Het laat met name zien dat het waternetwerk organisatorisch diverser is dan het algehele Europese onderzoeksnetwerk, en dat kennisgebruikers vatbaarder zijn voor het effect van nabijheid dan kennisproducenten. Ten derde laat ik zien dat geografische, organisatorische en sociale nabijheid elk bijdragen aan de geneigdheid van actoren om samen te werken, maar dat geografische en organisatorische nabijheid ook een afname veroorzaken van de waarschijnlijkheid dat actoren uitkomsten met expliciete kennis melden als resultaat van hun samenwerking. Ten vierde suggereert mijn onderzoek dat optimale samenwerking vanuit het perspectief van de actor bestaat uit een mix van verre en nabije relaties, ten minste in de geografische en organisatorische dimensie. Dit draagt bij aan het idee van een nabijheidsparadox, die luidt dat hoewel nabijheid een bepalende factor is voor innovatie, teveel nabijheid schadelijk kan zijn voor de innovatieve prestaties van de samenwerkende actoren. Ten vijfde draag ik bij aan de discussie over interacties tussen de verschillende dimensies van nabijheid. Ik vind geen bewijs voor directe interacties tussen de dimensies van nabijheid in hun effect op de geneigdheid om samen te werken, maar de combinatie van organisatorische nabijheid en organisatorische achtergrond heeft wel een indirect interactie effect: paren van kennisgebruikers zijn vatbaarder voor het effect van sociale en geografische nabijheid dan kennisproducenten en gemengde paren. Ik stel ook vast dat het effect van geografische nabijheid op expliciete uitkomsten van samenwerking minder sterk wordt als er wordt gecorrigeerd voor het effect van andere dimensies.

Mijn onderzoek draagt bij aan de literatuur over de *Triple Helix* – die sterk verbonden is aan de literatuur over organisatorische nabijheid – omdat het laat zien dat de organisatorische achtergrond van actoren in het wateronderzoek inderdaad diversifieert; de actoren met de meest centrale posities in het netwerk zijn echter allemaal universiteiten en publieke onderzoeksorganisaties. Ik laat zien dat samenwerkende actoren die organisatorisch ver van elkaar staan meer expliciete uitkomsten rapporteren. Dit ondersteunt het triple helix idee dat intensieve samenwerkingen tussen overheid, industrie en universiteiten kan bijdragen aan kennisproductie, maar ook dat overlap tussen deze sferen en hybride organisaties contraproductief kunnen werken.

Ik draag bij aan de literatuur over kennisproductie in de watersector door ideeën over de configuratie van het Europese water kennisproductienetwerk te staven met bewijs. Ik toon aan dat het organisatorisch diverser is dan het algehele netwerk, en dat landen met een klein R&D budget een relatief sterke representatie hebben in het kennisproductienetwerk over water dat ontstaat uit de

Kaderprogramma's. Ik laat ook zien dat het EU onderzoeksbeleid effectief is in het slechten van de ervaren organisatorische fragmentatie in de sector en het bevorderen van integratieve benaderingen.

Appendix A

List of keywords just to select all potentially relevant projects from the EUPRO database:
water OR *desalin* OR *hydrolog* OR *sewer* OR *sewage* OR *drought* OR *flood* OR *sludge* OR (*arid area*) OR (*arid region*) OR *irrigation* OR *erosion* OR *estuar* OR *wetland* OR *coast* OR *lagoon* OR (*river* AND NOT *driver*) OR *tidal* OR *aquatic* OR *brackish* OR (*sea level*) OR (*catchment area*) OR (*sea defence*) OR (*sea-defence*) OR *hydropower* OR *desertification* OR *dyke* OR *dike* OR *desiccation* OR *eutrophication* OR *acidification* OR *alga* OR (*pond* AND NOT *spond*) OR *dredg* OR *biomanipulation* OR (*acid rain*) OR *aridification*

List of keywords just to select the final set of water projects:
water treat*; water quality; drinking water; waste water OR wastewater; desalin* water; hydrolog*; water cycle*; water system*; water management; sewer* OR sewage; water distribution; water suppl*; water sanitation; water resource*; water quantity; water demand; water policy; water sustainab*; water climate change; water global warming; water recycl; water reuse; ; water recovery resource; energy 'water us'; water governance; water scarc*; water drought; water stress; water deficit; water technolog*; water sludge; water framework directive; water meter*; groundwater OR ground water; surface water; water consumption; water us* OR water-us*; water protection; arid area OR arid region; grey water; irrigat*; water conservation; water shed OR watershed; ; water contaminat*; water utilit*; 'water research'; 'water sector'; erosion; estuar*; wetland; blue energy; river water; freshwater; tidal; aquatic; brackish; salin* water; 'asset management' water; sea level; lake; catchment area; watershed; sea defence; hydropower; desertification; arid; eutrophication; biomanipulation; precipitation; dredg; algae water; aridification; pond; acidification; acid rain; bathing water; aquifer; reverse osmosis; water reclamation; flood disaster; flood hazard; biogas digestion; 'available technology' water; flotat*; biofilm water; protozoa water; 'water network'; rehabilitation water; ecosystem service water; 'hydraulic fracturing' water; 'potable water'; legionella water; bottled water; 'urban water'; 'water pump'; wave energy; rainwater OR rain water.

Dankwoord

In dit proefschrift bouw ik voort op de observatie dat kennisproductie vaak een gezamenlijke activiteit is. Mijn onderzoek vormt daarop zeker geen uitzondering. Ik ben dan ook vele mensen dank verschuldigd voor hun hulp bij de totstandkoming van dit proefschrift.

Wim, jij bent eigenlijk de grondlegger van dit proefschrift. Je droeg veel ideeën aan, en liet me ook de ruimte om die naar eigen goeddunken in te passen in mijn werk. Je gaf me ook vele ingangen in de waterwereld. Dank daarvoor.

Mariëlle, het aantal concepten van mijn hand waar jij je doorheen hebt geworsteld, is niet meer te tellen. Dank voor al je commentaar en kritische vragen. Laurens, jij raakte betrokken bij dit project toen ik al een poosje bezig was, en je wist het nieuwe schwung te geven. Je hebt me veel bijgebracht over de ambachtelijke kant van wetenschap bedrijven. Ik ben heel blij dat je ook na mijn vertrek bij het Rathenau nauw betrokken bent gebleven bij de afronding van mijn proefschrift.
Edwin, jij hebt me ingewijd in de wondere wereld die *science studies* heet. Je bijdrage aan dit proefschrift ligt op velerlei vlak, van conceptuele ideeën tot slimme methodes om data te verwerken. Ik heb veel van je geleerd.

Wim Heijman, jou wil ik bedanken voor je bijdrage aan hoofdstuk twee van dit proefschrift, maar vooral ook voor de mogelijkheden die je me hebt geboden om mijn eerste stappen te zetten in het onderzoek en kennis te maken met het wetenschappelijke bedrijf.

Het Rathenau Instituut is een fantastische plek om te mogen werken. Ik dank dan ook alle collega's dat ze samen deze inspirerende omgeving vormen. In het bijzonder mijn dank aan de JuScis. Pleun , Thomas, Stefan, Roos, Keelie, Tjerk, Bei, bedankt dat we al onze proefschriftervaringen samen konden delen. Een speciaal woord van dank natuurlijk voor mijn kamergenoten Tjerk en Bei. Tjerk, jij hebt waarschijnlijk het meeste gejeremieer over mijn proefschrift in wording verdragen, bedankt daarvoor. Ik mis onze discussies en groteske muzikale performances zeer. Bei, het was een genoegen om met jou samen te werken. Op naar jouw promotie! Clara wil ik bedanken voor al haar hulp bij het vormgeven en drukken van dit boekje. Barend, bedankt dat je als afdelingshoofd de voortgang van het complexe waterproject hebt bewaakt en bevorderd. Ik heb het erg gewaardeerd dat je ook met me hebt meegedacht over mogelijke nieuwe carrièrestappen na het Rathenau.

De collega's bij KWR Watercycle Research Institute wil ik danken voor de interessante inkijkjes die ze me geboden hebben in de praktijk van kennisproductie in de watersector. In het bijzonder wil ik de collega's van de kennis-

groep Kennismanagement noemen. Bedankt voor de interessante discussies, het meedenken en alle praktijkvoorbeelden die jullie me gegeven hebben! Theo, bedankt dat je jouw kennis over hoe de Brusselse hazen lopen met me wilde delen.

Jasper en Evelyn, jullie boden me de kans om zelf een radertje te worden in het ingewikkelde mechaniek van het onderzoeksbeleid. Dank voor jullie vertrouwen en de ruimte die jullie boden om mijn proefschrift af te ronden! Ook de andere collega's bij EZ wil ik danken voor hun begrip als ik weer eens op donderdag ondergedoken zat.

Mijn familie wil ik danken dat ze altijd voor me klaar staan. Mijn ouders dank ik dat ze me altijd de ruimte hebben gegeven om mijn eigen weg te gaan en me een kritische blik lieten ontwikkelen. Ken, bedankt voor je hulp bij het maken van een webscraper. Het is niet meer direct te zien, maar het heeft me erg geholpen bij het ontwerpen van hoofdstuk 4 en 5. Herman, jij hebt me voorgedaan hoe je een proefschrift moet afronden, bedankt daarvoor.

Marjon, in de eerste plaats bedankt voor je hulp bij het opmaken van mijn proefschrift en het maken van een werkbare planning, zonder jou was ik nu waarschijnlijk nog bezig met fonts en stijlen. Maar meer nog wil ik je bedanken voor je liefde, geduld en opofferingsgezindheid. Woorden schieten tekort om te omschrijven wat ik jou verschuldigd ben. Tineke, door jou heb ik meer over gezamenlijke kennisproductie geleerd dan door dit proefschrift. Bedankt dat jij mijn werk in het juiste perspectief plaatst.

Soli Deo Gloria

Curriculum Vitae

Pieter Heringa was born in Borsele on 5 December 1985. He attended secondary education at Calvijn College in Goes, graduating in 2004. Pieter got his Bachelor's degree in International Development studies with a specialisation in Economics of Rural Development from Wageningen University in 2008 (*cum laude*). He continued his studies at Wageningen University in the same area and specialisation and obtained his Master's degree in June 2010 (*cum laude*). From January 2010 to December 2013 he worked for the Science System Assessment department of the Rathenau Instituut (Den Haag) as a junior researcher.
Pieter was involved in a research project about knowledge production in the water sector, resulting in this thesis. This was a joint project with KWR Watercycle Research Institute in Nieuwegein, where Pieter was affiliated as guest researcher during the project. In February 2014 he has joined the department of Innovation and Knowledge at the Ministry of Economic Affairs (Den Haag) as senior policy officer.

Who was Rathenau?

The Rathenau Instituut is named after Professor G.W. Rathenau (1911-1989), who was successively professor of experimental physics at the University of Amsterdam, director of the Philips Physics Laboratory in Eindhoven, and a member of the Scientific Advisory Council on Government Policy. He achieved national fame as chairman of the commission formed in 1978 to investigate the societal implications of micro-electronics. One of the commission's recommendations was that there should be ongoing and systematic monitoring of the societal significance of all technological advances. Rathenau's activities led to the foundation of the Netherlands Organization for Technology Assessment (NOTA) in 1986. On 2 June 1994, this organization was renamed 'the Rathenau Instituut'.

Does distance matter for collaborative knowledge production? This question has intrigued scholars and policy makers for a long time. Distance does not only have a geographical dimension, but also an organisational, cognitive and social dimension. These dimensions feature prominently in the national and transnational research policies of the past decades. The object of study in this thesis is the water sector. This field has a large strategic relevance, and it is an interesting field for proximity research because of its high organisational and cognitive diversity, and because its challenges are specific to local conditions, but not bound by administrative and cultural borders. The research question of this study is: How is collaborative knowledge production in the water field influenced by geographical, organisational, social and cognitive proximity of the actors involved?

This study shows that geographical, organisational and social proximity all three have a positive effect on the propensity of actors to collaborate. Knowledge users turn out to be more susceptible to this effect than knowledge producers. Social and cognitive proximity also have a positive effect on the occurrence of reported outcomes of collaboration. Geographical and organisational proximity have a negative effect on explicit outcomes, but a positive effect on tacit outcomes of collaboration.

This study ends with recommendations for research policy to build and strengthen collaborative knowledge production networks.

ISBN 978-90-77364-61-1