Characteristics of evidence in the eyes of the policy maker: What makes scientific knowledge evidence

Master Thesis

by

Simon Couwenberg
s1739131
s.couwenberg@student.utwente.nl

Submitted in fulfilment of the requirements for the degree of Master of Science, program Public Administration, University of Twente

2021

Supervisors:
Prof. dr. René Torenvlied, University of Twente
Dr. Veronica Junjan, University of Twente

Acknowledgement

This thesis began in October 2019 when Prof. dr. René Torenvlied gave me the opportunity to become a student-assistant for a WODC research project into the use of evidence at the Ministry of Justice and Security. This provided me not only with the topic for this thesis but also with a very educational year as part of the research team. I am grateful for this opportunity and to the other members of the research team for involving me in all parts of the research project.

I would like to thank my supervisors Prof. dr. René Torenvlied for his feedback, but especially for his guidance with the statistical analysis of the data and Dr. Veronica Junjan for her feedback and for pointing me towards previous research and thus helping me find my feet in the field of the evidence-based policy literature.

Secondly, I would like to thank the WODC for their permission to use the gathered survey data also for my thesis. I am grateful to the civil servants at the Ministry of Justice and Security that took the time to complete the survey thereby providing me with valuable data.

Finally, I would like my family and my housemates for their support and for lending me a listening ear. I would especially like to thank my sister for proofreading my master thesis.

# Abstract

In the field of evidence-based policy the most attention goes to scientific knowledge as the main form of evidence. However, there are doubts as to the extent to which scientific knowledge is used as evidence. This thesis examines if the use of evidence in the policy making process depends on whether a scientific study can meet several characteristics policy makers consider important for scientific knowledge. An attempt is made to identify these characteristics, in order to determine what policy makers perceive as a high-quality and useful scientific study. This is guided by the first research question; *What characteristics of scientific evidence foster their use by policy makers?* Similarly, the standards scientists set for scientific studies are identified. These characteristics and standards are then compared to see if any major conflicts exist. As such this thesis hopes to be able to identify problematic points which prevent policy makers from using scientific evidence by examining knowledge itself. This identification is done at the hand of the second research question; *To what extent are these characteristics compatible with accepted academic norms about scientific evidence?* It is hoped this will provide a better explanation as to why evidence-based policy making remains so complicated.

## Table of content

# Abbreviations

| | |
|---|---|
| OMT | Outbreak Management Team |
| RIVM | Rijksinstituut voor Volksgezondheid en Milieu (National Institute for Public Health and the Environment) |
| EBP | Evidence-based policy |
| RCT | Randomised controlled trial |
| VSNU | Vereniging van Universiteiten (Association for Universities) |
| NWO | Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research) |
| WODC | Wetenschappelijk Onderzoek- en Documentatiecentrum (Research and Documentation Centre) |
| pcf | Principal component factor (data reduction method) |

# 1 Introduction

One of the expectations that exists of governments and by extension its policy makers is that they are well informed before they create any policies. This helps legitimise the policies which are created but simultaneously places a burden upon the policy makers to inform themselves of all the issues that are relevant to the policy arena they are active in. A good example of how governments remain informed can be found in the current Covid-19 pandemic. In many countries committees were created consisting of medical experts who advised politicians and policy makers about possible measures they could implement and the probable impact of those measures. In the Netherlands the "Outbreak Management Team" (OMT) was created, which consisted of various medical experts and scientists. The president and thus the face of this team was Jaap van Dissel who is the director of the RIVM Centre for Infectious Disease control (Rijksinstituut voor Volksgezondheid en Milieu, in English: National Institute for Public Health and the Environment). An aspect which is of interest later in this thesis is that he is also a professor of Internal Medicine at Leiden University. He is thus active in both the academic field as well as in government, by leading a team which provides policy advice on Covid-19. During the pandemic almost every country had their own medical experts who were very visible, with Anthony Fauci, the American equivalent of van Dissel, probably being one of the most well-known experts worldwide. These experts have become some of the most known figures in their respective countries and provide politicians and policy makers with most of the information upon which decisions are made and policies are written.

However, while the recommendations of medical experts and scientists had a big impact on policy due to Covid-19, the reliability of knowledge, facts and evidence are increasingly under discussion. This includes the knowledge that is available about Covid-19. This has led to the increasing prominence of terms such as alternative facts, which seemingly offer an alternative to previously uncontested facts. Narratives based upon these alternative facts have become increasingly prominent in both social media as well as regular media, offering alternative explanations to current societal issues. Using the Covid-19 pandemic as an example, numerous stories exist which question the (academic) knowledge upon which governments base their measures to minimise the impact of Covid-19. These stories range from questioning the use and efficacy of vaccines to stories on social media which suggest that you will become magnetic after being vaccinated against Covid-19 as one of the more outlandish examples (Goodman, 2021). There are even more extreme versions of such alternative narratives where people do not trust science at all. Instead, they view it as simply another opinion devoid of any reliable factual knowledge. This increasing distrust of knowledge which was previously generally accepted and relied upon has made the discussion of how and what knowledge is used by policy makers to create policies even more relevant.

Decades ago governments and more specifically policy makers often relied on information for their policy making process that was generated internally within the government itself (Head, Ferguson, Cherney, & Boreham, 2014). Eventually this changed to incorporate more external sources of information or evidence as it is called currently. Governments are no longer able to be completely self-sufficient for their information needs in their policy making process and have started to incorporate other external sources of information and knowledge. Head et al. (2014) observe that a wide variety of sources are used and that different government agencies have different ways of using and analysing them. There is an academic field that focusses precisely on what kind of knowledge and how knowledge can be incorporated in the policy making process. This is the field of evidence-based policy making. The core idea in this academic field is that evidence can contribute to better, more efficient and more effective policies (Head, 2010). The assumption is that by providing knowledge in the form of information, ideas and recommendations policy makers can create better policies. As such evidence is thought to provide a firm foundation upon which policies can be based (Weiss, 1977). This raises the question what evidence is and how it differs from knowledge. There is a very large pool of all kinds of knowledge, whether it be scientific, practical, or some other form of knowledge. Evidence, however, is only the knowledge that can actually be used as evidence by policy makers. For example, knowledge about agriculture will likely never be evidence for educational policies, while it may very well serve as evidence for the ministry of agriculture. As such evidence is knowledge of a specific field that is useful for policy makers. That leaves the question what kind of knowledge can become evidence.

The literature identifies several types of evidence which can be distinguished from each other. The first and most accepted type is scientific evidence (Head, 2008). The idea is that reliable, unbiased, and replicable scientific knowledge can offer valuable evidence upon which policy makers can base their policies. Scientific evidence is also the type of evidence which has been studied the most by various authors, mostly through surveys in Canada (Amara, Ouimet, & Landry, 2004; Landry, Lamari, & Amara, 2003), the United Kingdom (Talbot & Talbot, 2014), the United States of America (Hall & Jennings Jr, 2010; Jennings Jr & Hall, 2012) and in Australia (Cherney, Head, Boreham, Povey, & Ferguson, 2012; Cherney, Head, Povey, Ferguson, & Boreham, 2015; Newman, Cherney, & Head, 2016, 2017). The second type of evidence that is discussed in the literature is professional evidence, also called practical evidence, and stems from the experience people accumulate. Both Schön (1984) and Head (2008) argue that practitioners or professionals possess valuable knowledge for the policy making process often derived from extensive experience in their respective fields. The third type of evidence Head (2008) distinguishes is political evidence, which concerns political activities and the impact they may have on the policy making process. Within this type there are “politicians, parties, organised groups, and the public affairs media” that provide the evidence (Head, 2008, p. 5). There is, however, much less agreement within in the academic field regarding whether politics is a type of evidence or rather an important contextual factor in the policy making process. A fourth type that has been suggested by Freiberg and Carson (2010) is the existence of emotions as a type of evidence, though this has not been widely adopted in the academic field.

Besides the field of policy making the term evidence-based also has an important place in the medical world. The field of evidence-based medicine concerns itself with how scientific knowledge can contribute to medical practice. In contrast to policy making the medical field has been able to incorporate evidence very successfully in their everyday practices. One way through which evidence impacts medical practice is at the hand of Cochrane, which is an international not-for-profit network active in more than 130 countries which concerns itself with generating “authoritative and reliable knowledge” to improve evidence-informed “health decision-making” (Cochrane, n.d.). They help medical professionals improve their knowledge and decision making.

In the field of evidence-based policy-making the British Government was among the first to attempt to explicitly incorporate evidence in 1999 by incorporating it in the governments modernisation agenda (Parsons, 2002). The focus was on academic evidence and professional experience as the two evidence types to substantiate policy (Parsons, 2002). This proved to be an important moment for the evidence-based policy movement as it gained increasing traction in both the scientific world as well as in different governments around the world. However, even though this academic field already exists for some time and governments are already attempting to incorporate evidence in their policy making process this does not mean that utilising evidence is easy. One of the issues is that scientific knowledge does not automatically translate into useable evidence which automatically lands upon the desks of policy makers. Policy makers need to have access to sources of scientific knowledge, they also need to be able to understand the knowledge contained within these sources and they need to be able to apply it to their specific policy problems. Weiss (1977) neatly sums up a range of issues which make the uptake of scientific knowledge more difficult.

> *Weaknesses in the research itself, conflicting demands on policy, and disjunctions between the knowledge needs of policymakers and the research outputs of social scientists. On the research side, much of what goes by the name of social science knowledge is flawed, inconclusive, ambiguous, and contradicted by evidence from other studies. Many research conclusions are limited in scope or out of date. (p. 532/533)*

Weiss (1977) names a host of issues that play some role in the process of evidence-based policy making. Gathering evidence for their policies is of course not the only claim to the attention of the policy makers. They also need to account for possible ideological issues inherent to politics which shape the amount of discretion politicians give to policy makers in the policy making process (Weiss, 1977). All these issues and the limited attention policy makers can devote to gathering and incorporating evidence, since for most that is not the sole focus of their jobs, have led academics to the realisation that evidence-based policy making is not as easy as they once thought (Head, 2010).

Already in the 1970s Caplan (1979) suggested that there is a gap or disjunction between the academic field and the policy world. Weiss (1977) did not go as far as a disjunction but she did perceive a certain distance between policy makers and scientists. They both suggested that scientific knowledge is not directly accessible to policy makers, that it needs to bridge a some distance if not a gap which hinders the uptake of scientific knowledge. Though more recently Newman et al. (2016) nuance this observation by Weiss (1977) and Caplan (1979) by arguing that there is a range of interaction between the two worlds and that viewing them as separate entities does not do justice to reality. This more nuanced view does seem to line up with the situation that has developed with the Covid-19 pandemic. As government committees have been created consisting of medical experts and scientist which provide specific policy advice based upon the latest (scientific) knowledge available. Thus apparently, bridging the gap Weiss (1977) and Caplan (1979) observe. Others argue that it is not so much a gap, as much as that scientist can adopt different roles. This ranges from scientists who broker knowledge to others, so called ‘honest brokers’, to ‘pure scientists’ who are only interested in science itself. The limited number of scientists who focus on brokering scientific knowledge to policy makers, also limits the amount of knowledge that can be brokered (Spruijt, Knol, Torenvlied, & Lebret, 2013). However, no firm consensus exists with regards as to why exactly it remains so difficult to incorporate scientific evidence in the policy making process. This has led to more nuanced terms than evidence-based policy such as evidence-aware, evidence-inspired, evidence-influenced and evidence-informed policy (Nutley, Morton, Jung, & Boaz, 2010). These new terms indicate a lesser use of evidence in the policy making process.

This thesis aims to explore the slow uptake of evidence from a different perspective thus contributing to the discussion within the literature as to why it remains so difficult for policy makers to use evidence. It is argued that the characteristics of scientific knowledge do not necessarily match with the characteristics policy makers expect of scientific evidence. Policy makers may have other requirements they set for evidence for their policies, requirements which scientific knowledge may not be able to meet. This thesis focusses on evidence and the perceived characteristics of evidence from the perspective of the policy makers. The main focus of this thesis will be on scientific evidence and not on the other forms of evidence. With the current societal changes and the prominence of alternative facts it is that much more important that there is a good understanding of how the scientific and the policy world interact and how it can possibly even be improved. This will help with substantiating policies using reliable scientific knowledge. Furthermore, it can hopefully address the concerns of some people who question the reasons for specific government policies, by better explaining the process of how knowledge travels from scientists to policy makers. These societal changes demonstrate the importance for the scientific community to more clearly explain how their studies lead to new knowledge and how that knowledge can function as evidence for policy makers. Simultaneously it demonstrates the importance for the policy community to be transparent about what knowledge they base their policies on. Therefore, the following two questions are central to this thesis:

*What characteristics of scientific evidence foster their use by policy makers?*

*To what extent are these characteristics compatible with accepted academic norms about scientific evidence?*

At the hand of these two questions existing literature will be examined to derive characteristics of evidence for policy makers and standards of knowledge for scientists. Afterwards the theoretical compatibility of these characteristics can be discussed. Finally, the characteristics of evidence for policy makers are tested at the hand of a survey of the civil servant of the Dutch Ministry of Justice and Security to establish whether these characteristics are indeed important to policy makers.

First, however, this thesis continues with a discussion of the sub questions which help with answering the main research questions. This is followed by the theory chapter which will start with a more general introduction of evidence based policy and discuss the work of both Weiss (1977) and Caplan (1979). The third chapter is about the methodology of this thesis, including the operationalisation of the variables. This is followed by the data analysis chapter in which the survey results are analysed. Finally, the fifth chapter is the conclusion in which the main research questions are answered.

## 1.1 Research questions

In this section the research questions for this thesis are discussed. The goal of this thesis is to explore whether the extent to which knowledge is used as evidence can be explained based on characteristics of the knowledge. As such the main research questions are:

*What characteristics of scientific evidence foster their use by policymakers?*

*To what extent are these characteristics compatible with accepted academic norms about scientific evidence?*

To answer these questions several sub questions have been formulated to answer parts of the two main research questions. The first two questions explore what characteristics scientific research must meet to be considered of high quality to scientists and useful as evidence to policy makers. This resulted in the following two sub questions:

1. *Which characteristics does (scientific) knowledge, generally, need to have according to policy makers in order to become authoritative evidence to be used by them in policy?*

2. *Which accepted standards does scientific research, generally, have to meet in order to attain the status of authoritative scientific research in the academic forum?*

The first sub question is about what characteristics scientific knowledge needs to have for policy makers to use it in the policy making process. This question will be answered at the hand of the existing literature, though not many academics have researched what important characteristics of evidence are for policy makers. Some papers do however, either refer to or actually point out issues that are important to policy makers (e.g. Cherney, Head, Boreham, Povey, and Ferguson (2011); Cherney et al. (2015); Weiss (1977, 1979).

The second sub question is similar only it focusses on what important standards of knowledge are to scientists. This question is also answered based upon existing literature. The rules, norms and values that guide scientists are examined as these form the basis upon which scientific research is conducted. What makes scientific studies authoritative is not widely discussed in the literature, as such documents such as the Netherlands Code of Conduct for Research Integrity (KNAW et al., 2018) are used in answering this question.

The third sub question is:

3. *Under what conditions are the characteristics perceived by policy makers compatible with standards for scientific research, and under which conditions are they incompatible?*

This question follows from the first two questions and puts the list of characteristics and standards for each profession besides the other. A comparison is made of which characteristics and scientific standards are compatible with each other and which are not. This is done both prior to analysing the empirical data and verifying the requirements of the policy makers and after analysing the data. This provides both a theoretical comparison and a comparison that only takes the characteristics which are significant in predicting the use of evidence in policy making into account. The hope is that this provides further insight into the interactions between both professions. Such a comparison may indicate aspects which could hinder the uptake of scientific evidence by policy makers.

# 2 Theory

This thesis argues that the extent to which evidence is used in the policy making process depends on the applicability of scientific knowledge by policy makers. The idea is that the scientific knowledge that policy makers can use as evidence has several specific characteristics. If scientific studies cannot meet these characteristics or if these characteristics do not match with scientific standards, then this may explain why using scientific evidence remains so difficult. Therefore, in this thesis an attempt is made to identify the characteristics scientific evidence must have according to policy makers. At the same time, an attempt is made to identify the standards or characteristics scientific studies need to meet according to scientists. After which these two sets of characteristics can be compared to find out whether they could form obstacles for the incorporation of evidence in the policy making process.

First however, a short introduction is given on the basics of the evidence-based policy literature, starting with the types of evidence that are discussed in the field. This is followed by a discussion of the main mechanisms that Weiss (1979) describes in her research utilisation models in which she explains how scientific knowledge can reach policy makers. Then the two communities theory is examined and it is explained why this theory is set aside. This thesis attempts to add to the models of research utilisation by Weiss (1979) by better explaining why some scientific studies are more suitable and as such are easier to incorporate as evidence compared to other scientific studies.

## *2.1 Evidence and knowledge*

The idea that evidence can improve policy making has long been discussed but became more prominent over the last fifty years. Through the rising importance of subfields such as research utilisation (Weiss, 1977, 1979), but also due to the attempt by the British government to focus on evidence-based policy making from 1999 onwards. The new British labour government made it an important part of their modernisation agenda to actively use knowledge in the policy making process (Parsons, 2002). Meanwhile, working evidence-based was already very important in the medical field, where evidence-based medicine has significantly impacted medical practice. The field of evidence-based medicine does not have the same discussion regarding what types of evidence exist as is the case for the field of evidence-based policy. In evidence-based medicine it is scientific evidence that is the sole source of evidence. This is largely due to the rigorous methodological requirements that exist when doing medical research. Before the research can be used by medical professionals, it must be rigorously tested and reviewed.

As already shortly discussed in the introduction there are four different types of evidence that are identified in the literature. The idea of both political evidence as well as emotional evidence has so far gained little traction in the academic field. In contrast there is much more agreement about scientific knowledge and professional knowledge serving as evidence to policy makers.

Head (2008) defines scientific evidence as scientific knowledge that is created through a process of systematic analysis. He argues that the way in which research is conducted is very important, since it needs to be in accordance with scientific values such as replicability and independence. One of the difficulties with scientific evidence is that scientific knowledge is not always complete and is only rarely definitive (Weiss, 1977). This uncertainty that still exists in the available scientific knowledge may make it more difficult for policy makers to use scientific evidence especially in less developed academic fields. This uncertainty also illustrates one of the main differences between the field of evidence-based medicine and the use of scientific evidence by policy makers, namely the evidence hierarchy. In the medical field methodological rigour is very important, with the preferred method being randomised controlled trials (RCTs) (Head, 2010). In the field of evidence-based policy making most academics agree that that RCTs are not feasible for studying policy problems (Head, 2010). What is an acceptable level of methodological rigour for scientific evidence is still being discussed in the literature. Nevertheless, most scientists argue that the field of evidence-based policy making should not adhere to the same strict methodological guidelines as is customary in medical sciences. They argue that the level of methodological rigour should be dependent upon the specific field of study. After all, not all fields of policy research lend themselves to quantitative studies with a large N and as such some studies end up more qualitative in nature (Head, 2010).

Head (2008) argues that professionals, also called practitioners, have extensive knowledge because of their practical experience which can be valuable to policy makers. Also Schön (1984) in his book *The Reflective Practitioner* emphasises the tacit knowledge practitioners have. He argues that universities only reflect a particular epistemology which while important often foregoes knowledge that stems from practitioners. He observes that the tacit knowledge of professionals enables them to reflect on their own knowledge and adapt to different situations. It can be argued that policy makers themselves are also practitioners, who accumulate their own tacit knowledge which undoubtedly also impacts the policy making process. Head (2008), focusses more on the experience of professionals which enable them to help policy makers with assessing the feasibility and the effectiveness of specific policies. Professionals who can provide evidence can be seen in a very broad sense, it can encompasses professionals from the private sector, NGOs or others with some form of practical experience which could provide useful experiential knowledge (Head, 2008; Weiss, 1979). According to Parsons (2002) the British government particularly valued professional knowledge in the 1990s, as it was thought that professional knowledge could help with developing systems which assisted with better management of knowledge and enable the government to keep learning.

The third type of evidence Head (2008) identifies is political evidence which centres around notions as prioritising, agenda setting, ideological spin, accountability, coalition forming and negotiations. In this type of evidence "politicians, parties, organised groups, and the public affairs media" are the actors that play a comparable role to professionals and scientists in the above two evidence types (Head, 2008, p. 5). So, in this case policy makers would need to take their knowledge and experience into account. Since this type concerns politics, ideology plays an important role. The implication of politics and the underlying ideology serving as evidence is that there may be some policy areas which are no longer debatable because political agreements determine what is acceptable (Head, 2008). This could result in a situation in which other types of evidence are no longer relevant or there is only a limited amount of discretion policy makers have in making policies. Similarly Schön (1984) observes that politics determine the policy objectives and the social benefits and costs, thereby defining the boundaries within which the policy makers have to work. However, this type of evidence is less agreed upon in academic literature as others see politics more as another factor that plays a role in the policy making process, but not as a type of evidence upon which policy makers base their policies.

A fourth type of evidence that is suggested by Freiberg and Carson (2010) is emotional evidence. They argue that especially within the field of criminal policy, policy makers should also "deal with the roles of emotions, symbols, faith, belief and religion" (Freiberg & Carson, 2010, p. 152). The difficulty in using emotions as evidence is that emotions regarding fairness and justice should not undermine the policy making process, but rather contribute to it. This type of evidence has not been widely adopted by the academic field, as most of the focus remains on scientific and to a lesser extent professional evidence.

Though, multiple types of evidence are discussed in the literature the focus of this thesis is on scientific evidence. This type has been the focus of most of the scientific studies in the field of evidence-based policy literature. Mainly through surveys scientists have examined the impact of scientific evidence on the policy making process (Cherney et al., 2015; Landry et al., 2003; Newman et al., 2016, 2017; Talbot & Talbot, 2014).

In the academic field of evidence-based policy making the expectation until around 2010 was that policies should be based upon evidence. However, over the past decade this expectation has become more nuanced since it became apparent that the uptake of (scientific) evidence remained difficult. This has led to a shift in the terminology from evidence-based to evidence-informed, aware, influenced and inspired (Nutley et al., 2010). This shift represents a new view on how evidence impact policy making, the main difference is the extent to which evidence is used. In more recent articles in the field of evidence-based policy, authors increasingly use these weaker terms as they no longer expect policies to be based solely upon evidence rather they observe that evidence is one of the factors that influence or inspire new policies. Head (2016) points out that they also tend to accept more types of evidence, rather than solely focus on scientific evidence. In this thesis the term evidence-based policy will continue to be used. The reason for this is twofold; one is that this thesis builds mainly upon an older part of the available literature in which the term evidence-based is still used. Second, this thesis does not aim to examine how or to what extent evidence or what types of evidence are used by policy makers. This thesis focusses on what policy makers find important characteristics in scientific evidence.

**2.1.1 Research utilisation**

Weiss (1977) observed a lack of fit between the desires of policy makers and the evidence scientists were able to provide. This led her to examine the problem from a different perspective, namely of how knowledge got transferred from scientists to policy makers. Consequently, she was able to distinguish seven different models describing ways in which knowledge can travel to policy makers. For the purpose of this thesis the interest lies mainly in the mechanisms she describes of how scientific knowledge can travel to policy makers and not in the specific models. Therefore, only the mechanisms specifically about the utilisation of scientific knowledge are discussed.

The most straight forward mechanism is, of course, direct applicability. When a policy maker is faced with a policy problem, a first course of action may be to examine the available scientific knowledge to see whether it can offer a solution. Similarly, policy makers could directly commission scientific studies to address any knowledge deficits they may have. Especially this latter option can provide tailored scientific evidence for the problem the policy makers are faced with.

Another mechanism of how scientific knowledge can reach policy makers is more convoluted. It presupposes that research findings eventually filter through various channels to all citizens and over time this new knowledge is fully absorbed by all people. In this way scientific knowledge shapes how the population perceives the world and thinks about society. In this case policy makers do not directly use scientific evidence, however they have been shaped by it all their lives. As such it may very well have shaped them and how they think about society and the policy problems they face, but the specific impact scientific knowledge had on the policy process is much harder to determine. This mechanism sketches scientific research as the intellectual foundation of a society, as it provides the "background of concepts, orientations, and empirical generalizations that inform policy" (Weiss, 1977, p. 544). This intellectual foundation results in a gradual dissemination within societies of scientific knowledge and a gradual but cumulative change in society.

In the last mechanism Weiss (1979) sees research as completely intertwined with society. She argues that in this case societal changes prompt new scientific studies, through for example the reallocation of scientific funding, and the results of those studies directly influence society. This is a much more direct and interactive process than the mechanism directly above describes. Weiss (1979) suggests that it may be time that scientists pay closer attention to the needs of society and while it may not necessarily help increase the actual use of knowledge, it may contribute in a broader sense to the "wisdom of social policy" (p. 431).

### 2.1.2 Two communities theory

Though Weiss (1977) observes the difficulties of policy makers to utilise scientific knowledge, the two communities theory goes further as it sees the scientific community and the policy community as two completely separate entities. It points to differences in culture, conflicting values, different incentives and that scientists and policy makers speak a different language from each other (Caplan, 1979). According to Caplan (1979) the scientific community is only concerned with ""pure" science and esoteric issues" (p. 459), while the policy community consists of "action-oriented, practical persons concerned with obvious and immediate issues" (p. 459). Caplan (1979) argues that the purposes and the people who work in both communities are so different that knowledge cannot flow from scientists to policy makers. He lists a number of possible causes for this separation between the two, ranging from cultural differences, to misuse by politicians or simply due to the failure of the two communities to communicate well with each other (Caplan, 1979).

The idea of two separate communities has proven persistent over the years. Head (2010) observes that there is often a poor fit between what knowledge scientists can provide and what kind of knowledge policy makers require. He argues that both groups need to be more aware of the needs of the other group. As such, a researcher should focus more on areas and problems that are important to policy makers and they should improve the accessibility of their research by explicitly discussing the practical implications of their findings. At the same time policy makers should be more aware of the availability of existing research and improve their capacity to understand and utilise scientific studies (Head, 2010). In a survey of senior civil servants in the United Kingdom Talbot and Talbot (2014) observed many complaints about the language that scientists used. Respondents thought it contained too much jargon and perceived the used language as a barrier between the two communities. Though, the civil servants also reported that it did not prevent them from using scientific knowledge in their work.

There are several issues with the two communities theory which deserve some more elaboration. Firstly, Caplan (1979) points to cultural differences or the different languages scientists and policy makers speak and write, which attribute to the separation between the two communities. However, with regards to cultural differences this may differ considerably over different countries, which raises the question whether the gap is there in all countries. Perhaps the gap is only present in the cultures that keep more distance between the two professions than may be the case in other countries. Additionally, the issue with speaking and using different languages seems increasingly dated as, at the very least in the Netherlands, an increasing number of civil servants have a master's degree or even a PhD. In 2017 almost 54% of the Dutch national government civil servants had at least a master's degree (Ministry of the Interior and Kingdom Relations, 2017), which makes it unlikely that they are unable to understand and communicate well with scientists. Furthermore, according to a survey of Australian public servants by Head et al. (2014) 77% was confident in their personal skill "to collect and analyse policy-relevant information" (p. 93). Secondly, with regards to the different norms scientists and policy makers develop based upon which they judge scientific knowledge Caplan (1979) may have a valuable point. These norms help determine what both communities perceive as valuable knowledge. This aspect will be more closely examined later, as it is the core of this thesis. Thirdly, Caplan (1979) observes that there are different incentive structures for scientists which do not necessarily promote studying policy problems. This is also confirmed by other research, at least in Australia, by Cherney et al. (2012) who find that institutional processes in the academic world form an impediment to the uptake of research by policy makers. They point out that writing articles is, in many countries, more highly regarded than writing a report commissioned by government organisations. The expectation is that good scientists publish their work in scientific journals for their peers and not so much that they do studies for governments.

In accordance with Caplan (1979), Weiss (1979) also acknowledges that some distance between scientists and policy makers exists and that transferring the knowledge from the scientists to the policy makers is complicated. However, she focusses on the ways scientific knowledge could reach policy makers, as described in the mechanisms above. As such, Weiss (1979) advocates for a more nuanced view of any possible gap between scientists and policy makers, since there are a number of ways scientific knowledge reaches policy makers.

Newman et al. (2016) re-examined the two communities theory in 2016 at the hand of a survey of and interviews with Australian public servants. The survey responses of the civil servants indicated that they were neither ignoring nor structurally incorporating evidence in the policy making process. Instead, they tended to be relatively neutral to the use of academic research. On the question whether academic research was useful in improving the understanding of policy, the mean answer was 'occasionally' (Newman et al., 2016). The responses of the Australian civil servants in the survey however did not match with what would be expected if there was a gap which separated the policy and scientific communities. In the interviews Newman et al. (2016) conducted they found some respondents who had a decidedly negative view of academic research in line with the two communities theory. Academics were said not to produce relevant results in a timeframe that was compatible with the policy world and that it was often difficult to understand as if they wrote in a different language (Newman et al., 2016, pp. 28-29). However, they also interviewed respondents who had a totally opposite view, with a very positive attitude towards the role of science in policy making. The majority of the interviewees fell somewhere between these two positions. According to Newman et al. (2016) the majority was either neutral to the use of scientific knowledge or thought scientific knowledge was occasionally valuable. This led Newman et al. (2016) to argue for a much more nuanced view than the two communities theory provides. They state that a "more realistic approach would acknowledge that there is a range of interactions between policy and academia, with different individual policy makers valuing and using academic research more than others." (Newman et al., 2016, p. 31). Another point of critique from Newman et al. (2016) on the two communities theory is that the theory does not define the policy community clearly. Often both politicians and the civil servants were combined under the header *policy community*, while according to Newman et al. (2016) politicians and policy makers should be seen as separate groups. They point out that politicians function in a different culture, have different incentives and motivations in comparison with policy makers. Additionally, they warn against aggregating all public servants into one group, termed the policy community, as that may not do justice to the differences that exist among policy makers.

The conclusions by Newman et al. (2016) of a more nuanced view of the relation between scientists and policy makers is further supported by Spruijt et al. (2013).They argue that scientists can have different roles; some of which are very much removed from the policy arena, while others are much closer to policy makers and focus on brokering knowledge.

## *2.2 Characteristics of knowledge*

This thesis departs from the perspective of a separation between scientists and policy makers and follows more in line with the perspective adopted by Weiss (1977, 1979), thus putting knowledge itself central. It aims to catalogue what characteristics make a scientific study authoritative for policy makers and for scientists. In the first part the characteristics of scientific evidence are examined, or in other words what requirements policy makers have of scientific evidence. Afterwards the attention is directed at the characteristics of scientific studies that are important from the perspective of the scientists. Based upon (among others) the literature discussed above the characteristics of both are identified. Therefore, in this section the first two sub questions of this thesis can be answered; *Which characteristics does (scientific) knowledge, generally, need to have according to policy makers in order to become authoritative evidence to be used by them in policy?* And in the second part of this section; *Which accepted standards does scientific research, generally, have to meet in order to attain the status of authoritative scientific research in the academic forum?*

### 2.2.1 Characteristics of evidence according to policy makers

Part of the difficulty for policy makers with using evidence is the context in which they function. They often also need to take non-scientific knowledge or contextual circumstances into account. It is the task of the policy makers to weigh scientific knowledge and other sources of information or obligations and reconcile it with each other into a working policy. This may mean that policies are not (fully) based on scientific knowledge, simply because other obligations also played a role or even superseded scientific evidence. This raises the question whether it is always recognised that scientific knowledge is used when it is diluted with non-scientific knowledge or limited by other obligations. At the hand of the literature several characteristics can be distilled which policy makers may find important when they are weighing scientific knowledge against all the different information sources.

*Timeliness*

Both Cherney et al. (2011) and Weiss (1977) suggest that it is important that scientific research needs to be available on time, or in other words it needs to be available when it is needed by policy makers. As policy problems arise, policy makers often do not have the time to wait for new studies to be done, therefore scientific studies on the subject either need to exist already or be nearly finished. Commissioning scientific studies and then performing those studies takes time, which policy makers may not always have especially with urgent issues. Though there may also be problems which do give policy makers the time to commission scientific studies to use as input for later policies.

*Applicability*

The second characteristic is that scientific knowledge needs to be applicable to become evidence for policy makers to use. Scientific knowledge should be able to be applied to specific situations or problems. Partly this applicability characteristic reflects a need of the policy makers for a certain concreteness and specificity of the scientific studies; it needs to provide some clarification of the policy problem (Weiss, 1977). As Weiss (1979) states “Whatever the nature of the empirical evidence that social science research supplies, the expectation is that it clarifies the situation and reduces uncertainty" (p. 427). The evidence-based policy movement is centred around the idea that policy makers use scientific evidence when creating policies. This means that a good scientific study should clarify the situations or problems the policy makers face and possibly even offer feasible suggestions. This feasibility is also an important aspect, because when a scientist makes policy suggestions in their paper that are infeasible, then that paper is only of limited use to the policy maker. If the scientific study is not capable of meeting these aspects of applicability it is of little use to policy makers. Cherney et al. (2011) also name some similar factors that are important with regards to applicability. First is that scientific knowledge should have direct implications for policy, second is that the knowledge should be presented clearly and third is that summaries of the findings of the study are important.

*Accessibility*

Accessibility is a characteristic that should be interpretated in two different ways. One is that a scientific study can only be used by policy makers if it is accessible in the first place (Cherney et al., 2015). If policy makers do not have physical or digital access to specific studies, they can hardly take them into account. This is also the reason that numerous surveys ask respondents about the access and the sources of information and scientific knowledge they have (e.g. Cherney et al., 2015; Jennings Jr & Hall, 2012; Talbot & Talbot, 2014)

Secondly, and this interpretation is more in line with the applicability characteristic, is that the knowledge in the scientific studies should be relatively easily accessible (Cherney et al., 2015). This means that it should be clearly written and understandable. It should be transparent about the methodology and the results as well as their implications should be discussed clearly. This is potentially a difficult point since most scientific papers focus on other scientists as audience and not necessarily on others, such as policy makers, who may be less familiar with the scientific field. A similarity can be drawn to the language difficulty that Caplan (1979) mentions as one of the factors which cause the gap between scientists and policy makers in the two communities theory.

*Methodology*

The methodology parts of scientific studies are also important for policy makers. A part of this comes from the need for reliable scientific knowledge. This seems to be a fairly obvious requirement as scientific studies which are not reliable are of little use to policy makers, but according to Cherney et al. (2011) policy makers attach significant importance to this aspect. Cherney et al. (2011) also show that policy makers attach a considerably high importance to scientific studies being unbiased. Similarly, it seems likely that policy makers attach importance to most of the same methodological requirements as scientists do, whether it concerns the data gathering method, the validity or the design of the study. If policy makers cannot depend upon the methodology of a study, they also cannot depend on the result of the study as input for the policy making process. As such, a good methodology is required for the credibility of the scientific study.

### 2.2.2 Standards of scientific knowledge

Both Caplan (1979) and Newman et al. (2016) discuss the two communities theory, where the gap between scientists and policy makers is discussed. However, in both articles the discussion of the scientific side is somewhat limited. They both focus on the policy makers rather than the scientists. Caplan (1979) does not go much further than describing the scientific community as being concerned "with "pure" science and esoteric issues" (p. 459). With regards to the article by Newman et al. (2016), most of their focus is on the policy community in their examination of the two communities theory.

This section will examine which standards scientific research has to meet to be considered authoritative research. It aims to provide a comprehensive overview of the standards that are required thereby answering the second sub question of this thesis; *Which accepted standards does scientific research, generally, have to meet in order to attain the status of authoritative scientific research in the academic forum?*

The standards scientific studies need to meet to be considered of high-quality stem at least partly from the scientific culture that has been cultivated over the past decades if not centuries. Some of the norms and values regarding research have been committed to paper, often in the form of methodology books. These books and articles codify practice in the scientific world, however, according to De Groot (1961) not all of these norms have been committed to paper. Some of these requirements have become so interwoven into scientific practice that identifying them has become more challenging. For example, while the fact that plagiarism is absolutely not allowed has been codified and even studied, a requirement as originality is not often written down and acknowledged. Nevertheless, the importance of reaching for new knowledge and not endlessly repeating the same studies is taught at universities, thus instilling it into the scientific culture.

*Systematic, verifiable, critical and independent*

There are a few basic norms that scientific studies have to meet for research to be considered valuable at all. Research needs to be systematic, verifiable, critical and independent. These four standards form the basic rules or requirements of scientific research. It is through independent and systematic research that critical checks (through for example peer reviews or repetition of the study) can take place with which the quality of the study can be verified. The standard for research to be systematic mainly refers to the way in which a study is performed. It needs to be replicable and the authors must be clear on how the results were achieved. This is closely related to the verifiability requirement, as other scientists need to be able to verify the results of scientific studies through replication for example. This links up with the need to examine your own research and that of others critically, though this is more of an attitude and not a characteristic of knowledge is itself. For a scientific study to be trusted and for readers to be, at least partly, certain of its quality it needs to be published in a peer-reviewed journal, for which studies are first critically reviewed. Finally independence refers to how scientific studies should be conducted independently and that it and its results should be free of the influence of other parties and as such impartial.

*Methodology*

Dutch institutions, such as the Association for Universities (VSNU) and the Netherlands Organisation for Scientific Research (NWO) have defined five key principles that researchers should adhere to in the Netherlands Code of Conduct for Research Integrity (KNAW et al., 2018). They are thus not directly characteristic for authoritative knowledge, but characteristic often do follow from such principles. The five principles that are defined are honesty, scrupulousness, transparency, independence and responsibility. Some standards can be distilled from these principles and their explanations in the Netherlands Code of Conduct for Research Integrity.

The main characteristic that stems from these principles is the importance of the methodology of scientific studies. It should be clear where the research data comes from and how it was processed and attention needs to be paid to for example the reliability and the validity of the study. It needs to be clear how data was gathered, how it was processed and how it led to the results. As methodology encompasses a considerable number of aspects, it could be further specified by examining some of its constituent dimensions. One is the design of the study, which needs to enable the researcher to answer the research questions. Another is that the methods should be clear as well as the role of third parties and possible stakeholders in your research project (KNAW et al., 2018). The study should be conducted carefully and accurately, the data should be handled correctly and it should enable verification. The results should be reported clearly and the uncertainties that still exists and the line of reasoning should be made clear. An authoritative study should be assessed and peer-reviewed by publishing it in a, preferably high-quality, journal.

*Originality*

Originality is an important characteristic for scientific studies to adhere to. However, this characteristic is not often made explicit in scientific literature, except for scientific work that is more focussed on the philosophy of science. Additionally, the importance of originality is taught at Universities. Kuhn (2012) argues that scientists should not be satisfied, rather they should be committed to improving conceptually, theoretically, instrumentally and methodologically. A scientist should be committed to continually improving and increasing the explanatory power of their work. The more certain knowledge becomes the easier it will travel to policy makers and the bigger the impact it will have on policy (Weiss, 1979). Scientists can improve the certainty of knowledge by reflecting and improving their theoretical frameworks, by improving the instruments they use to gather and measure their data and by improving the methodologies of their scientific work. Here Kuhn (2012) comes close to what it means to be a scientist as creating new knowledge by improving upon previous knowledge is at the very core of the profession. However, Kuhn (2012) focusses on the characteristics of scientists and not of the knowledge itself, this thesis in contrast focusses on characteristic of knowledge and not of scientists. Consequently, the implications of these characteristics need to be distilled to be able to articulate standards scientists set for scientific research. This results in a standard of originality, a scientific study needs to be original in the knowledge it presents, it needs to improve or challenge existing phenomena, views or theories.

## 2.3 Compatibility of the characteristics of knowledge

In this section a start will be made to answer the third research sub question; *Under what conditions are the characteristics perceived by policy makers compatible with standards for scientific research, and under which conditions are they incompatible?* This will be done by comparing the theoretical characteristics of both the scientific world and the policy world with each other and examining whether they are compatible or not. It should be noted that this is a theoretical exercise, as the characteristics of policy makers have not yet been tested in practice. It may be the case that the characteristics which the literature suggest do not play a role in practice. In the first part of this section, it will be discussed whether there are any general conditions which may play a role with regards to whether the requirements are compatible with each other. Any conditions which may impact the compatibility of the characteristics of scientific evidence can be tested at the hand of the survey data, along with the characteristics policy makers find important.

A problem that remains in studies into the uptake of scientific evidence by policy makers is the judgement of the policy makers that are involved. As long as choices need to be made whether a specific scientific paper meets the criteria described above or any other criteria, a certain level of subjectivity remains. Though both scientists and policy makers may, for example, value a well written scientific article this does not mean they have the same ideas about what the best writing style for a scientific article is. Thus, judgement remains necessary but is by nature also subjective. There are a few aspects which somewhat alleviate this issue, one being that the same subjectivity also exists between scientists as some scientists have other ideas about what counts as a well written scientific study than others do. A second aspect is that at least 54% of the national civil servants in the Netherlands had a master's degree in 2017, this means that they are educated by some of the very scientists who provide the scientific knowledge (Ministry of the Interior and Kingdom Relations, 2017). It seems likely that this means that their views about what counts as good science remains relatively close together. This also makes it likelier that they do not have a radically different view of scientific knowledge than scientists do, as they were after all educated by them.

**General conditions**

Besides the subjectivity discussed above two general conditions are theorised to play a role in determining whether policy makers incorporate scientific evidence in their policy making process. The first is the relevance of the available knowledge, while the second is whether the available knowledge is politically acceptable.

This thesis is concerned with what characteristics play a role in the uptake of scientific knowledge, but for policy makers to be able to utilise scientific knowledge in the first place it should be relevant. For scientific studies to be relevant for policy makers they need to be relevant to society and the policy arena of interest. A consequence of this condition is that the pool of scientific knowledge that policy makers can draw upon may be limited. This all depends upon the status of the specific scientific (sub)field. If there are not many scientific studies or the opposite if the subject is so broad and all-encompassing, it may be subject to frequent changes. This has consequences to the pool of readily available knowledge policy makers have at their disposal. A part of the relevance condition is also about the certainty of the knowledge base that is available. Weiss (1977) points out in her paper that sometimes the contradictory nature of scientific studies may pose a problem for the uptake of evidence. Especially in academic fields that are less established scientific studies can be contradictory. Hoppe (2002) has written more extensively on this and called it the extent to which there is certainty about the knowledge base. Scientific papers which contradict each other may not be relevant to policy makers as they only further increase the amount of possible contradictory information and obligations they need to consider and weigh. Thus, a certain level of certainty about the knowledge base is desirable. Head (2010) observed that evidence-based policy seems to have less standing and relevance in tumultuous academic fields. He observed that when there is more certainty about the knowledge base that evidence-based policy making seems to work better. Therefore, the characteristics and standards can only be compatible with each other if the specific scientific research sources are relevant in the first place.

The second condition concerns the political acceptability of the scientific knowledge. As Weiss (1977) points out politics is about reconciling different, often value laden, interests and coming to viable compromises. It is less about coming to the best possible solutions based upon logical reasoning and a thorough understanding of the issue. This may mean that scientific knowledge that could function as evidence to policy makers may be problematic, due to what is politically considered as an acceptable outcome. Consequently, not all knowledge may be politically acceptable, which may prevent policy makers from using all the available knowledge. Such a condition has some similarities to the political evidence type that Head (2008) proposes, however, in this case it is not seen as a type of knowledge but rather as a contextual factor that impacts to what extent policy makers are free to use all available scientific knowledge. Limitations regarding the scientific knowledge that can be used do not necessarily have to be imposed upon policy makers by politicians, it may also be the case that policy makers filter the knowledge they have access to based upon whether they think it is acceptable to politicians.

**Compatibility between scientific standards and the expected characteristics of policy makers**

The compatibility of the characteristics will be discussed at the hand of the characteristics that are defined for the policy makers. Thus, first timeliness will be discussed and whether and how it matches with the requirements that are important to scientists. This will be followed by the other requirements, applicability, accessibility and methodology set out against the scientific standards; systematic, verifiable, critical and independent research, methodological quality and originality.

*Timeliness*

The scientific world sets high standards for itself with regards to the rigorousness of the methodology, but also regarding the basic norms that studies are systematic, verifiable, critical and independent. A consequence of all these requirements, however, is that scientific studies are often not performed quickly. This may lead to a conflict with the requirement of policy makers for a certain timeliness, since there is often no time to delay the policy making process by months to wait for scientists to study a policy problem. This leads to a considerable incompatibility especially when faced with urgent new policy problems, as this could mean that relevant scientific research is often unavailable when a new policy problem comes to light. Consequently, this incompatibility may mean that in such urgent policy problems scientific knowledge can only be used when evaluating and modifying the policies that were created.

The last standard of the scientific profession is originality, this standard seems much more compatible with the timeliness criteria. Mostly since it does not have immediate implications for the time scale on which scientific studies are conducted. New and original studies can only provide policy makers with a better insight into their policy issues.

*Applicability*

The second characteristic concerns the applicability of scientific knowledge. Scientific knowledge should be concrete and specific enough that it offers some extent of clarification to policy makers concerning the policy issues. This clarification criterium is compatible with the first four standards that scientific knowledge should be systematic, verifiable, critical and independent. These characteristics do not conflict with each other as they do not prevent scientists from clarifying issues in the papers and books they produce. This is also the case with regards to the methodological standard of scientists, as this also does not conflict with the desire of policy makers for knowledge that clarifies policy problems. The originality standard is especially compatible as scientific papers that offer new and original insights may be able to better explain policy issues and thus provide more applicable input.

*Accessibility*

The accessibility characteristic is formulated above as having two aspects, one being that the scientific studies themselves need to be physically or digitally accessible to policy makers. The other aspect is that the scientific knowledge should also be accessible in the sense that it is well written and that the results and the implications of the study are understandable.

With regards to the first aspect there exists one conflict in terms of compatibility with the scientific standards. This concerns the need for scientists to be critical, as its focus is that scientific papers should be peer-reviewed. While this is an important safeguard for the quality of the scientific studies, it could also form a roadblock in terms of accessibility if the study is not accessible to policy makers due to payment walls of the peer reviewed journals they are published in. This may complicate the accessibility for policy makers and necessitates a subscription to a variety of different journals. The other scientific standards, systematic, verifiable, independent, methodology and originality are compatible with this aspect of the accessibility characteristic. This is due to the fact that they do not influence the physical or digital accessibility of scientific knowledge.

The second aspect of accessibility is, as stated, more concerned with the comprehensibility of scientific studies. As such, this aspect is especially compatible with the need for scientific studies to be systematic since more systematic studies are often easier to read and thus understand the knowledge it contains. But also the other three standards, verifiability, critical and independent are compatible with this aspect. The same goes for the methodology standard if it is explained clearly it only contributes to the accessibility of the knowledge in the study. The last standard of originality is somewhat less compatible in this case since this standard centres around pushing the boundaries of scientific knowledge. The danger that arises for the compatibility is, however, that theories for example may become ever more complicated to increase their explanatory power or that increasingly complex measurement instruments could be developed and used. As a result, it may become more difficult for someone outside the scientific field to fully understand the study. This could have implications for its possible use as evidence, especially if policy makers are not able to determine the quality of the scientific paper.

*Methodology*

Policy makers attach considerable importance to the methodology of scientific studies as Cherney et al. (2011) show. However, the same standard also exists in the scientific world, where the methodology stands at the basis of a good scientific study. Discussing the compatibility between the methodological requirements of on the one hand policy makers and on the other scientists does not make much sense. Since it is the scientific world which sets the standards for methodological quality, at the very least it meets the policy makers desired methodological quality. As such the requirement of policy makers with regards to the methodological rigour seems to be met. The only case in which it may become problematic is if scientists set stricter standards than policy makers and if these standards prevent them from meeting some of the characteristics which are important to policy makers. An example could be the incompatibility between the timeliness characteristics and the scientific methodology characteristic discussed above.

Similarly, the compatibility between the policy makers methodology characteristic and the other scientific standards is high. This conclusion can be drawn relatively simply, due to the fact that the scientific standards for methodology can coexist without any danger to working systematically, verifiably, critically and independently while producing original studies. As such the compatibility with these other scientific standards and the policy makers methodological requirements are high.

## 2.4 Hypotheses

In this last section of the theory the hypotheses are formulated that have been derived from the discussion of the theory above. More specifically these hypotheses concern both the characteristics that policy makers find important for scientific knowledge, the general relevance and the political acceptability conditions that apply in the process of evidence-based policy making. At the hand of these hypotheses the characteristics that have been derived from the literature are tested as independent variables of the dependent variable the use of scientific evidence by policy makers.

Cherney et al. (2011) and Weiss (1977) both indicate that it is important that scientific research is available on time to policy makers. After all, if scientific research is not available on time policy makers cannot make use of it. The first hypothesis tests if timeliness is a good regressor for the use of scientific evidence in policy making.

*H1: Policy makers who report to make more use of scientific evidence attach more importance to the timely availability of scientific knowledge.*

The second hypothesis tests whether the applicability of scientific knowledge is important to policy makers in determining whether they use scientific knowledge as evidence in the policy making process. Weiss (1977, 1979) and Cherney et al. (2011) argue that scientific knowledge that offers a certain extent of clarification of policy problems or that makes specific policy suggestions is more useful as evidence than scientific research that is much more abstract.

*H2: Policy makers who report to make more use of scientific evidence attach more importance to the applicability of scientific knowledge.*

Then follow two hypotheses concerning the accessibility of scientific knowledge. These hypotheses are split into two separate ones as they encompass different dimensions of accessibility. Hypothesis 3a tests whether the accessibility of scientific knowledge, both digitally and physically, plays a role in the policy making process. Hypothesis 3b tests whether scientific knowledge is also accessible to policy makers in the sense that it is comprehensible.

*H3a: Policy makers who report to make more use of scientific evidence attach more importance to the obtainability of scientific knowledge.*

*H3b: Policy makers who report to make more use of scientific evidence attach more importance to the comprehensibility of scientific knowledge.*

The fourth hypothesis tests whether the methodology of scientific studies is an independent variable. According to Cherney et al. (2011), policy makers attach considerable importance to the methodological rigour of their knowledge sources.

*H4: Policy makers who report to make more use of scientific evidence attach more importance to the methodological quality of scientific knowledge.*

The fifth hypothesis is about the general condition that scientific knowledge needs to be relevant to policy makers in the first place, after all to be able to use scientific knowledge there needs to be relevant knowledge available.

*H5: Policy makers who report to make more use of scientific evidence attach more importance to the relevance of scientific knowledge.*

The last hypothesis is about the second general condition, namely that scientific knowledge needs to be politically acceptable. The expectations politicians have and the goals they set may limit policy makers from using at least some parts of the available scientific knowledge.

*H6: Policy makers who report to make more use of scientific evidence attach more importance to the political acceptability of scientific knowledge.*

**Conceptual model**

The causal relations that have been formulated in the hypotheses directly above can also be visualised at the hand of a conceptual model (see figure 1 below).

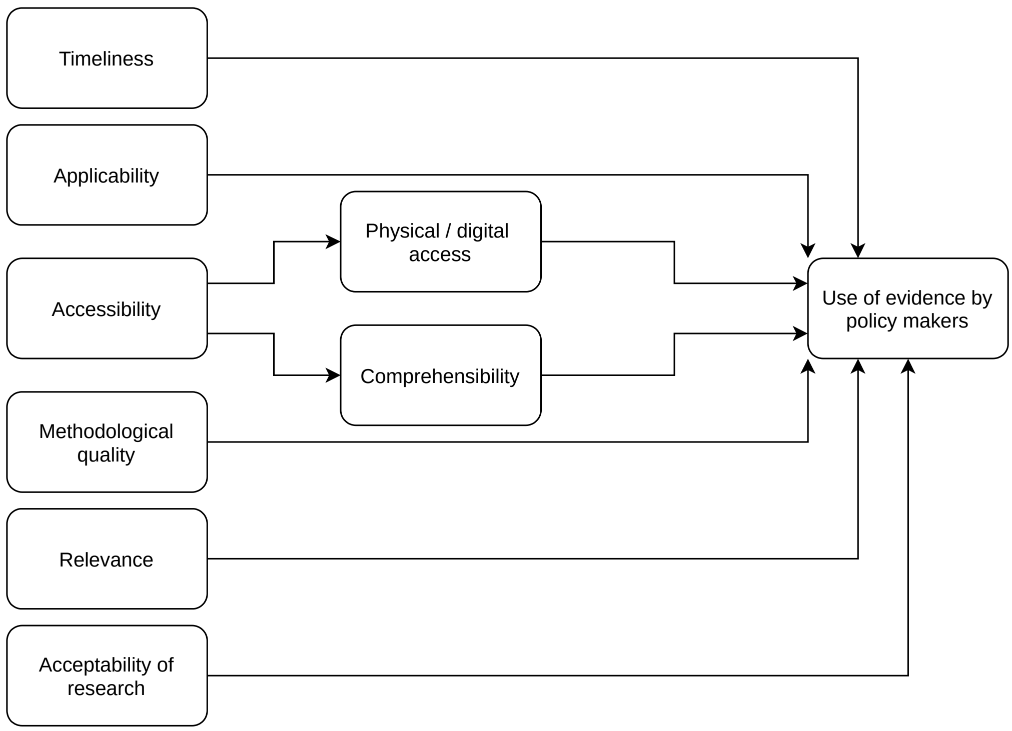


Figure 1: causal model of the theoretical independent variables on the use of evidence.

# 3 Research methodology

In this chapter the research methodology of this master thesis is set out. First the general research design is discussed, after which the sample and sampling that is performed to gather the data for this thesis is explained. This is followed by a discussion of the data collection and analysis as in the section on the conceptualisation and operationalisation of the concepts the results from some basic analyses are already reported.

## *3.1 Research Design*

In 2020 the University of Twente started a research project into evidence-based policy making at the Dutch Ministry of Justice and Security at the request of the WODC (Research and Documentation Centre, in Dutch: Wetenschappelijk Onderzoek- en Documentatiecentrum). This is an organisation that documents and studies, both internally and through external, independent scientists, subjects that are of interest to the Ministry of Justice and Security. The main goal of this research project is to investigate the use of scientific evidence in the policy making process and how this could be improved further. The Dutch Ministry of Justice and Security encompasses a considerable range of different policy areas ranging from more societal and behavioural sciences to more medical fields, with for example forensic psychiatric centres (Torenvlied, 2019). With almost 30.000 full-time equivalent (fte) national civil servants it is one of the bigger ministries in the Netherlands (Rijksoverheid, 2020). This research project serves as the background and inspiration for the subject of this master thesis. The data that is used in this thesis stems from the survey that was set out for the WODC research project. It should be noted that the survey was specifically designed for the WODC research project and not for this master thesis. This means that the survey questions are not specially made to measure the concepts of interest of this master thesis. However, as the subject of this thesis is closely connected to the WODC study into evidence-based policy making there are numerous survey questions which can also be used to test the hypotheses of this thesis.

## *3.2 Sample and Sampling*

The unit of analysis in this study is the Ministry of Justice and Security. As mentioned in the introduction the Ministry encompasses a broad policy area and as such also encompasses a lot of (semi-) independent organisations, institutions and agencies, such as the National Police for example. The WODC research project already limits the scope of the organisations that are examined as some organisational parts of the ministry are not included in the sample. Examples are the National Police and the organisations related to the judicial system, such as the public prosecution service (Openbaar Ministerie) but also the Judicial system Netherlands (which consists of the Dutch courts). However, for this master thesis the sample is further limited to the Directorates-General of the Ministry in 2020. This choice was made for two reasons. One is that this thesis is interested in the policy making process and most if not all of this takes place in these core departments. The second reason is more practical in nature, namely that at the time of writing the surveys set out among (some of) the other organisations of the Ministry have not yet been closed. Finally, the sample is further limited by the decision of the WODC research team to ask only those civil servants who are in the $10^{th}$ salary scale or higher to fill in the survey. This was done since not all functions at the Ministry are about policy making, the inclusion criterium of the $10^{th}$ scale is indicative of a university education.

The civil servants were directed to an intranet page, which contained a link to the survey, through an email by the knowledge coordinators of each Directorate-General in which they were invited to take part in the survey. To increase the response rate reminders were sent after the first few weeks had passed. On the intranet of the Ministry a message from the WODC research team was included in which the survey and its purpose was explained. It was also emphasised that the survey could be filled in anonymously and that the results would not be reported in such a manner that they would be traceable to any individual persons or organisational parts.

In 2017 the selected sample at the Ministry of Justice and Security had 1794 full time equivalents (fte) (Ministerie van Justitie en Veiligheid, 2017). Based on the invitation emails that were sent it is estimated that in the first round 96% of the civil servants were reached. While the reminder reached 78% of the civil servants of the ministry. Estimating the actual response is more difficult as the survey was intended for civil servants who are in the 10th salary scale or higher. However, it is not exactly known how many civil servants within the Ministry are in scale 10 or higher. For the Dutch national government as a whole 53,9% is scale 10 or higher (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2020). It should be noted that it is likely that this percentage is higher within the Ministry of Justice and Security. Using this percentage, it would suggest a maximum sample size of 968 fte's. When the survey was closed 273 responses were submitted, amounting to 15% of the total fte numbers or when taking only those in scale 10 or higher in account 28%. This would suggest that somewhere between 15% and 28% were reached. Though this is probably somewhat too positive as the ministry undoubtedly has parttime employees and probably also has a higher percentage of civil servants in scale 10 or higher than 53%.

## 3.3 Data Collection

The data that is used in this thesis stems from the survey created for the WODC research project into the use of scientific evidence in policy making at the Ministry of Justice and Security. The way in which this survey was created and set out has been discussed above. This survey was set out among the civil servants of the Ministry at the end of 2020 until the end of February 2021. For this thesis only a limited number of survey questions are used to provide the actual data that will be analysed in an attempt to empirically verify the hypotheses that are formulate above.

## 3.4 Data Analysis

The first step in analysing the data is to clean the dataset of variables that are created by the survey programme that was used to create and publish the survey. The survey data is cleaned and analysed with the statistical program Stata version 13.1.

The first step is to determine which questions measure the dependent variable, the use of scientific research. For this a data reduction method called the principal component factor analysis (pcf) is used, which can be run over related questions and will find whether there are any principal component factors underlying the selected questions. Therefore, for this thesis the pcf analysis is performed when analysing the question block which asked the respondents which types of knowledge they used as evidence and to create the control variables. The pcf analysis will then indicate under which questions there is an underlying factor and as such which questions can be combined to create the dependent variable for this thesis. For the pcf analysis items should at least have a factor load of 0,4 or higher and a component should exist of at least three questions.

With regards to the independent variables another method of analysation will be used, which is the calculation of Cronbach's alpha to check the reliability of the scales that can be formed when the variables are operationalised. For Cronbach's alpha a scale should consist of at least 3 items and the cut-off value that is used is 0,6 though the scale reliability is only considered strong from 0,8 or higher. Additionally, a factor analysis is performed over these independent variables to verify whether it also measures the same latent structures in the created scales.

The final step in the analysis of the data is the creation of multiple regression models with the use of scientific evidence as the dependent variable and the other variables as independents. The independent variables are regressed against the dependent variable. This is an explorative process to see which variables are significant or become significant when other variables are added or removed from the regression.

## *3.5 Conceptualisation and operationalisation*

This thesis theoretically identifies and empirically examines several concepts to check whether they play a role in the process of evidence-based policy making. Below these concepts are discussed and their possible constituent dimensions. Additionally, they are operationalised and the results of some basic analyses are shown. Using these basic analyses, the independent variables that policy makers find important can then be tested in the analysis chapter against the use of scientific evidence through multiple regression models. At the end of this section the reliability and the validity are discussed as well as possible control variables. For the dependent and the control variables where the pcf analysis was used the descriptive statistics can be found in appendix 1. For the other variable some basic statistics are given when they are discussed. The survey had an n of 273 as explained in the sample and sampling section above. For the specific data shown below this can differ since the survey also contains partial responses. As a result, the first question blocks have a bigger n than later question blocks. Since the survey was taken in Dutch the questions have been translated to English for use in this thesis. Additionally, it should be noted that the survey often asks questions about the department the respondent performs the largest part of their work and not about the respondent personally. Finally, for the sake of clarity the dependent and the independent variables are written in italics when they are discussed in light of the statistics. This is done to distinguish between when the variable itself is meant and when the concept is more generally discussed.

**Dependent variable the use of scientific evidence**

This thesis explores whether specific independent variables impact the use of scientific evidence by policy makers at the Ministry of Justice and Security. As such, the dependent variable for this thesis is the *use of scientific evidence*. The survey has one block of questions asking the respondents which types of knowledge they make use of. This question block is analysed exploratively to see whether there are any underlying principal components (see also table 1 in the appendix for the descriptive statistics). These twelve questions can be seen below in table 3.1, a pcf analysis is run over this question block which results in three factors underlying specific groups of questions.

**Table 3.1: Principal components underlying the used knowledge sources (component-loads)†**

| | Factor 1 Eigenvalue = 4,70 | Factor 2 Eigenvalue = 1,84 | Factor 3 Eigenvalue = 1,32 |
|---|---|---|---|
| Scientific insights into what works and what does not, such as meta-analyses or systematic reviews | **0,89** | 0,06 | 0,06 |
| Independent quantitative scientific research | **0,89** | 0,07 | 0,03 |
| Independent qualitative scientific research | **0,87** | 0,09 | 0,09 |
| Policy evaluations by independent scientists | **0,79** | 0,24 | 0,01 |
| Results from data-driven work within the Ministry of Justice and Security | **0,68** | 0,11 | 0,08 |
| Policy evaluations carried out by policy departments of the ministry | **0,54** | 0,34 | 0,24 |
| Parliamentary papers | 0,11 | **0,83** | 0,08 |
| Policy documents and regulations | 0,08 | **0,75** | 0,32 |
| Jurisprudence | 0,31 | **0,67** | -0,20 |
| Knowledge and experience of myself / direct colleagues | 0,05 | 0,12 | **0,89** |
| Knowledge and lessons from practice from my professional field | **0,52** | -0,13 | **0,64** |
| Instructions from supervisors | -0,03 | **0,41** | **0,47** |

*Note.* Factor 1 = Use scientific evidence; Factor 2 = Use documents; Factor 3 = Use professional knowledge. †After orthogonal (varimax) Kaiser rotation. *N* = 182.

The first factor is called *use of scientific evidence*, the second use of documents and the third use of professional knowledge. For this master thesis the sole interest goes to the first factor which measures the use of scientific research. Even though the first factor does include some sources which are not purely scientific it is the closest approximation to scientific evidence that is possible. Especially given the fact that the pcf analysis indicates that the scientific sources have the highest components scores in the first factor. Cronbach's Alpha for the factor *use of scientific evidence* is 0,89, which means it is a very reliable scale. For further analysis a sum scale is created based upon the mean score of the first factor by adding all items together and then dividing it by the number of items that make up the factor. This makes the analysis with multiple regression easier to interpret.

**Independent variables**

In this section the characteristics policy makers find important for scientific knowledge to possess according to the theory will be conceptualised and operationalised as independent variables. This is followed by the operationalisation of the two general conditions that are identified.

*Timeliness*

Timeliness as a concept is relatively simple, it refers to the availability of scientific knowledge when it is actually needed by policy makers. Policy problems can arise suddenly and it may be the case that no scientific knowledge is yet available about the specific problem thus preventing it from being available to policy makers when they are faced with the policy issue. The survey contains one question that measures this concept. To be able to draw any conclusions based on a single question the answer categories are recoded to create two dummy answer categories which combine completely disagree with disagree and completely agree with agree. Since this variable is only measured by a single questions a frequency table is fully displayed below in table 3.2. This removes the more extreme cases and creates answer categories that contain a reasonable number of cases as the completely (dis)agree categories only contain a few cases. Due to the way the specific survey question is formulated this variable will be called *research is available too late* in the multiple regression models

**Table 3.2: Variable research is available too late**

| My department does not use scientific research when, or because? | % (completely) agree† | Distribution characteristics | | | Frequency | | | |
|---|---|---|---|---|---|---|---|---|
| | | obs | Mean | s.d. | 1 | 2 | 3 | . |
| The research results are too late | 36,53% | 206 | 3,61 | 1,44 | 19,78 | 19,05 | 22,34 | 38,83 |

*Note*. † Percentaged on categories (1 t/m 3); Answer categories: (Completely) disagree (1); Neutral (2); (Completely) agree (3); I do not know / Not applicable / Missing responses (.).

*Applicability*

The second independent variable is applicability. This concept is somewhat more complicated as applicability is more difficult to demarcate clearly. The main purpose of this concept is to ascertain whether scientific knowledge can be applied by policy makers. For scientific knowledge to be applied it needs to be specific and concrete enough that it has some explanatory power for the policy makers particular policy problem. This variable is operationalised at the hand of four question items which together form a reliable scale as measured by a Cronbach's alpha of 0,80[1]. Since this variable is measured by multiple items only some simplified descriptive statistics are shown below in table 3.3. Since it forms such a reliable scale a sum scale is created by adding the first four questions together and dividing it by four. This sum scale will then be used to represent the independent variable *applicability* in the multiple regression.

**Table 3.3 Variable applicability**

| Motivation for using scientific knowledge: Scientific knowledge allows my department: | % (completely) agree† | Distribution characteristics | | |
|---|---|---|---|---|
| | | obs | Mean | s.d. |
| Gain deeper insight into issues | 93,30% | 216 | 4,36 | 0,71 |
| Make responsible choices | 89,32% | 215 | 4,30 | 0,73 |
| Evaluate existing policies | 88,61% | 216 | 4,25 | 0,81 |
| Develop effective interventions | 84,15% | 216 | 4,21 | 0,82 |

\Note. † Percentaged on categories (1 to 5); Answer categories: Completely disagree (1); disagree (2); Neutral (3); once (4); Completely agree (5); Don't know/ Not applicable (6); Missing responses (.).

*Accessibility*

The accessibility variable consists of two different dimensions. The first dimension of accessibility consists of actual physical or digital access to scientific knowledge. If policy makers do not have or cannot gain access to scientific knowledge, they are also not able to incorporate it in their policy making process. This first dimension, *access,* is operationalised by four questions from the questionnaire. It is operationalised at the hand of questions listed in table 3.4 below. Cronbach's alpha for this scale is 0,75 and as such the scale can be considered reliable[2]. This scale will also be operationalised by creating a sum scale, thus adding the items together and dividing them by four, creating a new variable which can be used in the regression analysis later. This variable is called *access* in the multiple regression models further down.

[1] Factor analysis of the selected questions shows the same latent structure as Cronbach's Alpha.
[2] Factor analysis of the selected questions shows the same latent structure as Cronbach's Alpha.

**Table 3.4: Variable accessibility: Dimension 1 (physical or electronic) access**

| Access channels for knowledge: Within my department we consult the following sources. | % (completely) agree | Distribution characteristics | | |
|---|---|---|---|---|
| | | obs | Mean | s.d. |
| Scientific meetings (such as focus groups, lectures, conferences) | 35,74% | 221 | 3,24 | 0,87 |
| Scientific books, journals or reports | 35,29% | 221 | 3,31 | 1,02 |
| Individual scientists | 22,62% | 221 | 2,92 | 0,99 |
| Scientific search engines (e.g. web-of-science, google scholar, scopus) | 16,29% | 221 | 2,71 | 1,36 |

\Note. † Percentaged on categories (1 to 5); Answer categories: Completely disagree (1); disagree (2); Neutral (3); once (4); Completely agree (5); Don't know/ Not applicable (6); Missing responses (.).

The second dimension concerns the ability of policy makers to access or in other words to comprehend the knowledge that is contained within scientific sources. For the operationalisation of the comprehension aspect no scale can be formed since there are not enough questions (three or more) which could measure this dimension. As such it is operationalised at the hand of a single question. Namely the questions whether "the analyses are too difficult to interpret" for policy makers. For this purpose the answer categories are recoded to create two dummy categories which combine completely agree with agree and completely disagree with disagree (see table 3.5 below for the basic descriptive statistics). This step will remove the extreme cases and form answer categories with a higher number of cases, since the completely (dis)agree categories contain relatively few cases. Due to the way the survey question has been formulated this independent variable will be named *incomprehension* in the multiple regression models.

**Table 3.5: Variable accessibility: Dimension 2 incomprehensibility**

| My department does not use scientific research when, or because? | % (completely) agree† | Distribution characteristics | | | Frequency | | | |
|---|---|---|---|---|---|---|---|---|
| | | obs | Mean | s.d. | 1 | 2 | 3 | . |
| The analyses are difficult to interpret | 32,23% | 206 | 3,28 | 1,41 | 27,47 | 17,22 | 21,25 | 34,07 |

*Note*. † Percentaged on categories (1 t/m 3); Answer categories: (Completely) disagree (1); Neutral (2); (Completely) agree (3); I do not know / Not applicable/ Missing responses (.).

*Methodology*

Methodology as a scientific concept is very complex as it encompasses numerous dimensions that all play a role and contribute to the methodological rigour of a scientific study. For the purpose of this thesis the concept of methodology is somewhat simplified since it needs to be captured in survey questions. Therefore, this concept is operationalised by four questions from the survey which all concern core methodological aspects, such as the verification of knowledge, the methodological quality, whether research is biased or contradictory (see table 3.6 below). These four items are used to construct a scale that measures whether the methodological quality of the research is too low. The Cronbach's alpha for this scale is 0,80, therefore it is a reliable scale. The actual scale is constructed by creating a sum scale for which the four variables are added together and divided by the number of items. Due to the way the specific questions are formulated in the survey this variable is called *methodological quality is too low* in the multiple regression models in the next chapter.

**Table 3.6: Variable the methodology quality is too low**

| My department does not use scientific research when, or because? | % (completely) agree† | Distribution characteristics | | |
|---|---|---|---|---|
| | | obs | Mean | s.d. |
| Knowledge and results are difficult to verify | 45,56% | 206 | 3,57 | 1,34 |
| The methodological quality of the research is too low | 41,95% | 206 | 3,60 | 1,41 |
| The research is too biased | 38,64% | 206 | 3,55 | 1,42 |
| The research results are contradictory | 33,91% | 206 | 3,54 | 1,44 |

\Note. † Percentage on categories (1 to 5); Answer categories: Completely disagree (1); disagree (2); Neutral (3); once (4); Completely agree (5); Don't know/ Not applicable (6); Missing responses (.).

*Availability relevant research*

The first general condition that was discussed in the theory is that scientific research needs to be relevant for it to be of interest to policy makers. Especially in new emerging fields or scientific fields which are still subject to frequent changes this may limit the amount of scientific research policy makers can use as evidence. Since the survey does not contain enough questions that measure this concept it is not possible to create a scale. Therefore, only a single item has been selected to operationalise the availability of relevant scientific research for policy makers. This variable is recoded into a dichotomous variable. This means that two answer categories remain, this has been done by combining completely agree with agree and completely disagree, disagree and neutral into one category. This removes the extreme cases and increase the number of responses, descriptive statistics can be seen below in table 3.7. This variable thus measures the *availability of relevant research.*

**Table 3.7: Variable Availability of relevant research**

| My department does not use scientific research when, or because? | % (completely) agree† | Distribution characteristics | | | Frequency | | |
|---|---|---|---|---|---|---|---|
| | | obs | Mean | s.d. | 1 | 2 | . |
| No relevant scientific research is available | 51,65% | 206 | 3,67 | 1,27 | 32,23 | 34,43 | 33,33 |

*Note*. † Percentaged on categories (1 t/m 2); Answer categories: (Completely) disagree and Neutral (1); (Completely) agree (2); I do not know / Not applicable/ Missing responses (.).

*Unacceptable research*

A second general condition that is discussed in the theory is that scientific research needs to be politically acceptable. Politicians may limit the amount of knowledge that is acceptable based upon the ideological standpoint of their parties. This variable is operationalised at the hand of a single question, since not enough questions are available in the questionnaire to form a reliable scale. Therefore, the questions which measures this variable is recoded into a dichotomous variable. This means that two answer categories remain, this has been done by combining completely agree with agree and completely disagree, disagree and neutral into one category. This removes the extreme cases and increase the number of responses, descriptive statistics can be seen below in table 3.8. This variable is called *unacceptable research*.

**Table 3.8: Variable unacceptable research**

| My department does not use scientific research when, or because? | % (completely) agree† | Distribution characteristics | | | Frequency | | |
|---|---|---|---|---|---|---|---|
| | | obs | Mean | s.d. | 1 | 2 | . |
| The results do not fit well with current policy or are politically-administratively unacceptable | 29,84% | 206 | 3,20 | 1,47 | 46,52 | 19,78 | 33,70 |

*Note*. † Percentaged on categories (1 t/m 2); Answer categories: (Completely) disagree and Neutral (1); (Completely) agree (2); I do not know / Not applicable/ Missing responses (.).

**Control variables**

Based on the questions that are asked in the survey a number of control variables can be constructed. These are questions about background characteristics such as gender and age. Slightly more complicated control variables have been created using explorative pcf analysis on the type of work respondents do and on characteristics which describe their professional field. Through an explorative pcf analysis four underlying factors have been found (see tables 2 and 3 in appendix for descriptive statistics of the control variables below).

Two factors are about the type of work the respondents do, as the term policy maker which has been used until now in the thesis is a very broad term. The pcf analysis suggest separating this into three different factors, one for policy makers with *management* functions, one for policy makers who are active as *policy advisors* and the third one for those who have executive tasks (see table 3.9 below). For this thesis only the first two factors are used as control variables as they can indicate whether the type of function matters with regards to the use of evidence.

**Table 3.9 Principal component factors underlying type of work (component-loads)†**

| | Factor 1<br>Eigenvalue = 2,35 | Factor 2<br>Eigenvalue = 1,72 | Factor 3<br>Eigenvalue = 1,27 |
|---|---|---|---|
| I support policy preparation / decision making | **0,85** | 0,03 | -0,02 |
| I work on policy documents | **0,81** | 0,05 | -0,16 |
| I mainly do advisory tasks | **0,49** | -0,01 | 0,13 |
| I supervise staff | -0,19 | **0,84** | 0,06 |
| I take decisions and determine strategy | 0,05 | **0,82** | 0,15 |
| I have contact with ministers and/or politicians | 0,39 | **0,70** | -0,21 |
| I work directly with target groups (clients, citizens) | 0,04 | 0,08 | **0,71** |
| I do research | 0,07 | 0,08 | **0,70** |
| I mainly perform executive tasks | -0,39 | -0,27 | **0,57** |

*Note.* Factor 1 = Policy advisor; Factor 2 = Management; Factor 3 = Executive tasks. †After orthogonal (varimax) Kaiser rotation. *N* = 201. One item (V7_Q6) no valid cases, dropped.

The other two control variables are factors which measure characteristics of the professional field the respondent works in. The first factor concerns the *uncertainty* of knowledge and goals, while the second factor is a control variable about the mainly political *sensitivity* of issues. In the theory both uncertainty of knowledge and the role of politics have been pointed out as factors which may impact the use of evidence and as such are useful to control for in the multiple regression models (Weiss, 1977). In table 3.10 below the result of the pcf analysis is displayed.

**Table 3.10 Principal component factors underlying characteristics of the professional field (component-loads)†**

| | Factor 1<br>Eigenvalue = 3,27 | Factor 2<br>Eigenvalue = 1,59 |
|---|---|---|
| Is decision-making focused on the short term | **0,72** | 0,15 |
| Factual substantiation is only used selectively | **0,67** | 0,18 |
| Does the available scientific knowledge have an uncertain status? | **0,64** | 0,24 |
| Is knowledge from practice more important than scientific knowledge | **0,62** | -0,12 |
| Is little scientific knowledge available | **0,60** | -0,30 |
| Information and facts often contradict each other | **0,59** | **0,53** |
| Is there little agreement in policymaking about goals and/or means | **0,58** | 0,39 |
| There is a lot of political dynamism | 0,08 | **0,84** |
| Do the media have a major impact | -0,01 | **0,82** |

*Note.* Factor 1 = Uncertainty; Factor 2 = Sensitivity. †After orthogonal (varimax) Kaiser rotation. *N* = 151.

These control variables are tested in the multiple regression models with the use of scientific evidence as the dependent variable. However, in the final regression model below only the significant control variables are included.

**Reliability and validity**

Both the reliability and the validity of this study are at least partly dependent upon the survey performed for the WODC research project. The survey questions are after all formulated specifically for that project and not for this thesis. As a result, the content validity of this thesis is slightly weaker than it would have been if the survey questions were specifically created to measure the concepts that are of interest in this thesis. This problem has been addressed mainly by combining survey questions into factors to be able to measure a concept more reliably. Finally, the independent variables that are measured by single items match closely with each other.

The construct validity is also important to discuss as some concepts partly overlap with each other. One issue in this thesis is that the applicability requirement is very close to the condition that scientific knowledge should be relevant. This is problematic from a construct validity viewpoint as ideally the questions should measure only one specific concept and not different concepts. However, there is not much overlap between the other variables this means that the need for construct validity is largely satisfied.

The last aspect of validity concerns how the survey is grounded in the scientific literature, thus whereupon the survey questions have been based. The questions have been sourced from and inspired by a range of other surveys into evidence-based policy (see Amara et al., 2004; Cherney et al., 2011; Cherney et al., 2015; Jennings Jr & Hall, 2012; Newman et al., 2016, 2017; Rimkutė & Haverland, 2015; Smet, 2013; Talbot & Talbot, 2014; Torenvlied, van der Berg, & Junjan, 2019). The survey was based upon other surveys conducted in the field of evidence-based policy, on a review of the literature in the field and on the specific context of the WODC research project. There is a certain redundancy in the questions of the survey to ensure that the measurements are indeed accurate. In this way the aim is to cover the full range of meanings concepts can have.

With regards to the reliability of the data there are a few issues that should be noted. A possible issue when performing such a survey study is non-response bias. In such a case it is expected that those who are very much in agreement or in disagreement with the study subject are more motivated to participate in the survey than those who are less opinionated. One of the ways by which this issue was minimised was to make sure that all the Ministry departments were surveyed, and that the invitation came from within the ministry itself and was very clear with regards to the purpose of the survey.

Another issue that may arise when surveying is that respondents feel the need to provide socially desirable responses. This issue was minimised by explicitly stating that participating in the survey would be anonymous and that the results would not be traceable to single persons or specific organisational parts. As such some of the background characteristics that are normally part of surveys were either left out or redesigned into categories. Respondents were for example not asked to name their exact age but could instead select the category in which their age fell. A final aspect of socially desirable responses was minimised by making clear through the wording of the survey that the incorporation of scientific evidence in policy is neither specifically good nor bad. In other words, any normative load with regards to possibly desired behaviour for EBP was avoided.

Since the data for this thesis stems from a questionnaire of which only multiple-choice questions are used, reliability in coding the responses is not an issue as would have been the case with interview data.

## *3.6 Ethical Issues*

This master thesis will be based upon the data that has been obtained through a survey set out for the WODC research project. For this WODC research project consent has been obtained from the ethical commission of the Behavioural, Management and Social Sciences (BMS) Faculty of the University of Twente under registration number 200981. Therefore, no separate consent was requested for this thesis as no new data was gathered and the data was handled according to the same terms as was announced to the respondents.

The civil servants of the Ministry of Justice and Security that have been approached with the question to participate in the survey have been informed of the purpose of the survey and the research project. Their consent was requested based on the following terms; they were informed that their participation in the survey was on a voluntary basis and that they could stop at any time, that the results would be anonymised so they cannot be traced back to individual persons or organisational parts. Finally, they were informed that only the researcher team will have access to the unprocessed data and the anonymised data will be saved for ten years after which it will be destroyed. If they agreed with these terms, they could then click on the link to be directed to the survey.

# 4 Data analysis

In this chapter the multiple regression analysis will be discussed and at the hand of those results the hypotheses will be answered.

## *4.1 Multiple regression models use of scientific evidence*

In table 4 below, the results of the multiple regression analysis are shown with the use of scientific evidence as the dependent variable, leading to two regression models. In the first model all independent variables are included as well as the control variable *uncertainty* since this was the only significant control variable. The second model only includes the significant independent variables.

There are three types of independent variables included in the regression below. The first type concerns the variables which are measured by single questions and which were recoded into 3 answer categories. This is the case with the *research is available too late* as well as *research is incomprehensible* variables. For the regression this means that the results are calculated for the neutral (2) and agree (3) categories with the first category disagree as reference. Therefore, two rows of numbers are given for these two variables. The second type are sums scales this concerns the variables *applicability*, *access* and *methodological quality is too low*, these have a number of categories as shown in the table. The third type are dichotomous variables that only have two answer categories and concerns the *availability of relevant research* as well as the *unacceptable research* variable.

**Table 4: OLS-Regression on the use of scientific evidence by policy makers**

| **Independent variables** | Model 1: Full model | | Model 2: only significant independent variables | |
|---|---|---|---|---|
| | B | t | B | t |
| *Research is available too late* (1= reference category (disagree) | | | | |
| - 2 (neutral) | -,02 | -,10 | | |
| - 3 (agree) | ,04 | ,23 | | |
| *Applicability* (sum scale 12 categories) | -,16 | -1,32 | | |
| *Access* (sum scale 19 categories) | ,46 | 4,84** | ,52 | 6,67** |
| *Research is incomprehensible* (1= reference category (disagree) | | | | |
| - 2 (neutral) | -,16 | -1,03 | -,19 | -1,29 |
| - 3 (agree) | -,17 | -,99 | -,31 | -2,05* |
| *Methodological quality is too low* (sum scale 16 categories) | ,24 | 2,44* | ,22 | 2.76** |
| *Availability relevant research* (dichotomous) | -,08 | -,55 | | |
| *Unacceptable research* (dichotomous) | -,04 | -,24 | | |
| **Control variable** | | | | |
| *Uncertainty* | -,37 | -3,49** | | |
| Constant | 3,47 | 4,92** | 1,34 | 4,22** |
| Adj $R^2$ | | ,33 | | ,30 |
| N | | 111 | | 141 |

*Note.* *p<0,05; **p<0,01.

In the first model there are three variables that have a significant relationship with the reported *use of scientific evidence* by policy makers at the Ministry of Justice and Security. These are the variables *access* and *methodological quality is too low* from the independent variables and the control variable *uncertainty*. This indicates that in the case of *access* the more access policy makers have to scientific sources the more they use scientific knowledge in the policies they make. With regards to the methodological quality, the worse the methodological quality of scientific knowledge the more scientific knowledge policy makers use. This may be because scientific knowledge with a lower methodological quality may be easier to use by policy makers compared to knowledge with a (very) complicated methodology. Neither *late availability of research* nor the *applicability* of scientific knowledge has a significant effect on the use of scientific evidence. The *research is incomprehensible* variable is also not significant in the first regression model. The conditions that were theoretically expected, namely that scientific knowledge should be both relevant and acceptable are not significant in relation to the use of scientific evidence. *Availability relevant research* is not significant and as such does not have any impact on the use of research by the respondents. Also the fact that *unacceptable research* is not significant is notable since Weiss (1977) argues that politics plays a role in the use of scientific evidence by policy makers. These results would suggest the opposite. Finally, only *uncertainty* has been included as control variable as it is the only one which has a highly significant relation with the reported use of scientific evidence. In the first model the outcome indicates that the higher the uncertainty the less scientific knowledge is used by policy makers, since the relation is negative. As such this control variable is an explanation for the non-use of scientific evidence.

In the second model all the independent variables that are not significant were left out so that only the significant variables remained. The result is that both *access* and the *methodological quality is too low* remain significant variables as is the case in the first model. The major change is that the third category of the *research is incomprehensible* variable becomes significant. This means that respondents who think that research is incomprehensible use less scientific evidence. In contrast to the other significant independent variables the relation between *research is incomprehensible* and *use of scientific evidence* is negative. This means that it explains the non-use of scientific evidence. Since incomprehension only becomes significant after the other variables are no longer included in the regression model, this means that its significance is explained away by other variables in the first model.

## 4.2 Hypotheses testing

*Timeliness*

The first hypothesis concerns the timeliness variable. Cherney et al. (2011) and Weiss (1977) suggest that it is important for policy makers to have the scientific evidence they need be available on time. In the multiple regression models the variable *research is available too late* is however not significant. The first hypothesis is:

> *H1: Policy makers who report to make more use of scientific evidence attach more importance to the timely availability of scientific knowledge.*

Based on this data whether scientific knowledge is available on time does not have any relation with the reported use of scientific evidence by policy makers. Therefore, this hypothesis is rejected. This is contrary to what both Cherney et al. (2011) and Weiss (1977) expected, as they argued that the direct availability of scientific knowledge was important for policy makers.

*Applicability*
Weiss (1979) emphasises the importance of scientific knowledge that is directly applicable by policy makers. The second hypothesis is:

*H2: Policy makers who report to make more use of scientific evidence attach more importance to the applicability of scientific knowledge.*

This hypothesis is also rejected, as the multiple regression models above show that applicability does not have a significant relation with the use of scientific evidence by policy makers. Therefore, this does not match with the argument that Weiss (1979) makes that scientific knowledge should clarify the situation and resolve uncertainty.

*Accessibility*
The third variable accessibility is further split into two dimensions. One consists of physical or digital access to sources of scientific knowledge and the other consists of whether scientific knowledge is comprehensible to policy makers. This first aspect resulted in hypothesis 3a:

*H3a: Policy makers who report to make more use of scientific evidence attach more importance to the obtainability of scientific knowledge.*

The hypothesis is confirmed, since *access* is highly significant in all the regression models shown in table 4. This in turn indicates that policy makers who report to make more use of evidence do indeed attach more importance to their access to sources of scientific knowledge. Accessibility is also in the literature recognised as an important factor as almost all existing studies into the use of evidence in the policy making process ask questions regarding the sources of knowledge that policy makers use.

The hypothesis that was formulated to test the other dimension of access, namely comprehensibility, is:

*H3b: Policy makers who report to make more use of scientific evidence attach more importance to the comprehensibility of scientific knowledge.*

This hypothesis is more complicated. The independent variable that was tested was the incomprehensibility of scientific knowledge against the use of scientific evidence by policy makers. This variable was not significant in the first full model, where all the independents were included as well as the control variable *uncertainty*. One of the answer categories (agree) of incomprehensibility did become significant however, when all the other variables were removed. Furthermore, it is important to note that the variable *incomprehensibility* explained the non-use of scientific evidence as the regression analysis shows the relation with the use of scientific evidence is negative. The negative relation suggests that the more incomprehensible scientific knowledge becomes the less policy makers report to use it. Or if you were to turn it around, the more comprehensible scientific knowledge becomes the more policy makers report to make use of it.

*Methodology*
The fourth variable, *methodological quality is too low* measures the concerns of policy makers regarding the methodological quality of scientific knowledge. This resulted in the fourth hypothesis:

*H4: Policy makers who report to make more use of scientific evidence attach more importance to the methodological quality of scientific knowledge.*

The results with regards to the methodology variable are complicated. Due to the wording of the relevant questions of the survey what was measured was whether the methodological quality was too low. The variable *methodological quality is too low* has a positive significant relation with the use of scientific evidence by policy makers. This means that policy makers attach more importance to a low methodological quality instead of a high methodological quality when it comes to the use of scientific evidence. However, based upon the theory it was expected that a high-quality methodology would contribute to the use of scientific evidence and not a low-quality methodology. However, this is not what the regression models show. The relation between the variable the *methodological quality is too low* and the *use of evidence* is opposite of what was expected. As such it can be concluded that there is indeed a significant relation between the methodology of scientific knowledge and the use of scientific evidence, however it is not in the direction that was expected.

*Relevance*

With regards to the general conditions that were expected to be important for policy makers to use scientific evidence, the first that was measured concerns the relevance of scientific knowledge. This was done at the hand of the following hypothesis.

> *H5: Policy makers who report to make more use of scientific evidence attach more importance to the relevance of scientific knowledge.*

This hypothesis is rejected as this variable is not significant in both models. Therefore, based upon this dataset it appears that the relevance of scientific knowledge does not relate to the use of scientific evidence.

*Unacceptable research*

This leaves the last hypothesis, which concerns the acceptability of scientific knowledge. Both Head (2008) and Weiss (1979) point out the importance of politics and ideology in the policy making process which led to the formulation of the sixth and final hypothesis.

> *H6: Policy makers who report to make more use of scientific evidence attach more importance to the political acceptability of scientific knowledge.*

This last hypothesis is also rejected. The regression analysis does not show any significance of the unacceptability of research in relation to the use of scientific evidence. This indicates that whether scientific knowledge is politically acceptable or not does not impact the reported use of scientific evidence by the respondents.

# 5 Conclusion and discussion

This goal of this thesis was to provide a different explanation for the difficulty in the uptake of scientific knowledge as evidence in the policy making process. The problem was approached not from the standpoint of the policy makers nor the scientists but rather from the perspective of the knowledge itself. This resulted in the following two research questions:

*What characteristics of scientific evidence foster their use by policy makers?*

*To what extent are these characteristics compatible with accepted academic norms about scientific evidence?*

This was further split into three sub questions which were answered at the hand of the academic literature. The outcomes of these inquiries in the form of characteristics and conditions were then tested at the hand of survey data gathered from the civil servants of the Dutch Ministry of Justice and Security. These results will be further discussed below to answer the two main research questions. However, they do come with some limitations and suggestions for further research, which will be discussed thereafter.

The first multiple regression model shows two independent variables which have a significant relation with the reported use of scientific evidence. The first is *access*, which Cherney et al. (2015) pointed out as a possible characteristic scientific knowledge needed to meet to be considered evidence by policy makers. This variable was shown to have a positive impact on the use of scientific evidence. The more access respondents have the more they report to use scientific knowledge in their work. The second significant variable is the *methodological quality is too low*, which also has a significant relation with the reported use of scientific evidence. However, this relation is positive as well, which means that the more the *methodological quality is too low*, the more the respondents reported to use scientific knowledge. The respondents are more inclined to use scientific knowledge when the methodological quality of the knowledge is worse. This suggests that scientific knowledge with a lacking methodology is more easily used in the policy making process, though further research is needed to confirm the nature of this relation. In the second model, which includes only the significant independent variables, incomprehensibility also becomes significant. The negative relation explains non-use indicating that the more incomprehensible scientific knowledge is the less it is used.

This significance of access and incomprehensibility is in line with the theoretical expectations that were discussed in the theory chapter and as formulated in the hypotheses. With regards to the methodology variable this too was significant though the relation is in the opposite direction as to what was expected. It was theorised that policy makers attached more importance to scientific knowledge with a good methodology, instead it turned out that policy makers attach more importance to scientific knowledge with a low-quality methodology. With regards to the other variables, timeliness, applicability, relevance, and acceptability these are all not significant in relation to the reported use of scientific evidence. Though these were theoretically expected to be significant, no evidence for this was found in the dataset. As a result, in answer to the first research question; *What characteristics of scientific evidence foster their use by policymakers?* Access, comprehensibility, and the lack of methodological quality in scientific knowledge foster the use of evidence by policy makers.

With regards to the second main question; *To what extent are these characteristics compatible with accepted academic norms about scientific evidence?* The access variable is compatible with the academic norms about scientific evidence, though the need for publishing in peer-reviewed journals does limit access to scientific knowledge somewhat. However, publications are not the only source of evidence for policy makers and as such this characteristic is compatible. The second dimension of *accessibility* concerned the comprehensibility of scientific knowledge, which is compatible with academic norms. Nevertheless, one caveat closely related to the next variable *methodological quality is too low*, is that scientific knowledge does need to be systematically written and offer clear conclusions. This leads directly to the last significant variable, namely *methodological quality is too low*, which is not compatible with the academic norms that exist about the methodological quality and rigour. The regression analysis shows that policy makers value a low methodological quality, which is contrary to the academic norm that the methodological quality should be high. An explanation could be that policy makers do not so much want scientific knowledge with a low-quality methodology rather than they just want knowledge that is easily understood. Which especially in more complex studies with complicated methodological methods may pose a challenge for the uptake of knowledge. In conclusion, at least based upon the regression analysis done above, the policy makers characteristic methodology and the academic standard for methodology are not compatible with each other.

There are some limitations to these conclusions which should be discussed here. First is that the survey was set out at the Ministry of Justice and Security, which encompasses a very wide range of policy areas. This means that the analysis for this thesis is based on the aggregate across a wide variety of policy subjects. As a result, this possibly distorts the view of the individual policy subjects, as for some policy areas scientific knowledge may be much more plentiful, the field much more settled and thus the knowledge more definitive and therefore easier to use. In contrast in other policy areas scientific knowledge may be scarce and of low quality. Consequently, where you to examine the policy areas individually the view may vary considerably depending on the subject. This could impact for example the significance of variables such as relevance and the acceptability of scientific knowledge. If a considerable group of respondents works on policy areas which are much more politically sensitive and for which scientific knowledge is relatively scarce it could distort the data and thus the results.

A second limitation is that the characteristics of evidence and the academic standards have been solely based upon the available literature. This could be further improved by conducting interviews with policy makers and scientists to find out whether the characteristics and standards that stem from the literature indeed match with reality. The results of such interviews could then again be tested to see whether that is also more generally the case, or whether the characteristics are very particular to the policy maker or the policy issue at hand.

A third limitation is that the survey was not specifically designed to measure the concepts that are central to this thesis. Therefore, for the analysis the survey questions through which the concepts have been operationalised are the closest approximations of the concepts that were possible given the available questions. As such, the construct validity of this thesis is not as high as it would have been when a survey had been specifically created to measure the concepts of interest.

As a result, a definitive conclusion regarding whether the uptake of scientific knowledge can be explained not on the basis of the people but rather on the basis of the knowledge cannot be reached. It does not become clear based upon the analyses performed in this thesis whether the characteristics of scientific knowledge and the compatibility with academic norms are indeed capable of explaining the difficulty in the uptake of scientific knowledge as evidence by policy makers. No uniform outcome is reached, this in combination with the limitations means that plenty is left to study and explain.

Finally, with regards to further research on this subject a few suggestions can be made based upon the limitation identified above. First is that the characteristics of scientific knowledge according to policy makers and the academic standards should be studied more closely. This in order to find out whether they indeed match with what the limited literature suggests or whether especially policy makers find other aspects of scientific knowledge much more important. Secondly, these new findings should then be tested again preferably with a survey specifically designed to measure the concepts as that would eliminate some of the concerns regarding the construct validity that are present in this thesis. Third, such a study should be performed preferably on a policy area which is less diverse or in which the responses can be split based upon the policy field. This way it becomes easier to account for differences in the policy areas regarding how they approach and handle scientific evidence.

# 6 References


Amara, N., Ouimet, M., & Landry, R. (2004). New evidence on instrumental, conceptual, and symbolic utilization of university research in government agencies. *Science communication, 26*(1), 75-106.

Caplan, N. (1979). The two-communities theory and knowledge utilization. *American behavioral scientist, 22*(3), 459-470.

Cherney, A., Head, B., Boreham, P., Povey, J., & Ferguson, M. (2011). The utilisation of social science research in policy development and program review. *Preliminary report: Phase, 1*.

Cherney, A., Head, B., Boreham, P., Povey, J., & Ferguson, M. (2012). Perspectives of academic social scientists on knowledge transfer and research collaborations: a cross-sectional survey of Australian academics. *Evidence & Policy: A Journal of Research, Debate and Practice, 8*(4), 433-453.

Cherney, A., Head, B., Povey, J., Ferguson, M., & Boreham, P. (2015). Use of academic social research by public officials: exploring preferences and constraints that impact on research use. *Evidence & Policy: A Journal of Research, Debate and Practice, 11*(2), 169-188.

Cochrane. (n.d.). About us. Retrieved from https://www.cochrane.org/about-us

De Groot, A. D. (1961). Methodologie: grondslagen van onderzoek en denken in de gedragswetenschappen.[Methodology: Foundations of research and thinking in the social sciences]'s-Gravenhage: Mouton.

Freiberg, A., & Carson, W. G. (2010). The Limits to Evidence-Based Policy: Evidence, Emotion and Criminal Justice 1. *Australian Journal of Public Administration, 69*(2), 152-164.

Goodman, J. (2021). Covid Vaccines: No, your jab isn't magnetic. *BBC*. Retrieved from https://www.bbc.com/news/av/57207134.

Hall, J. L., & Jennings Jr, E. T. (2010). Assessing the use and weight of information and evidence in US state policy decisions. *Policy and Society, 29*(2), 137-147.

Head, B., Ferguson, M., Cherney, A., & Boreham, P. (2014). Are policy-makers interested in social research? Exploring the sources and uses of valued information among public servants in Australia. *Policy and Society, 33*(2), 89-101.

Head, B. W. (2008). Three lenses of evidence-based policy. *Australian Journal of Public Administration, 67*(1), 1-11.

Head, B. W. (2010). Reconsidering evidence-based policy: Key issues and challenges. In: Taylor & Francis.

Head, B. W. (2016). Toward more "evidence-informed" policy making? *Public administration review, 76*(3), 472-484.

Hoppe, R. (2002). Cultures of public policy problems. *Journal of Comparative Policy Analysis: Research and Practice, 4*(3), 305-326.

Jennings Jr, E. T., & Hall, J. L. (2012). Evidence-based practice and the use of information in state agency decision making. *Journal of public administration research and theory, 22*(2), 245-266.

KNAW, NFU, NWO, TO2-federatie, Hogescholen, V., & VSNU. (2018). Nederlandse gedragscode wetenschappelijke integriteit [Netherlands Code of Conduct for Research Integrity]. *DANS*. Retrieved from https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:110600/tab/2.

Kuhn, T. S. (2012). *The structure of scientific revolutions*: University of Chicago press.

Landry, R., Lamari, M., & Amara, N. (2003). The extent and determinants of the utilization of university research in government agencies. *Public administration review, 63*(2), 192-205.

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2020). *Jaarrapportage Bedrijfsvoering Rijk 2019 [Annual Report for Central Government Operations 2019]*. Retrieved from https://www.rijksoverheid.nl/documenten/rapporten/2020/05/01/jaarrapportage-bedrijfsvoering-rijk-2019

Ministerie van Justitie en Veiligheid. (2017). *Introductiedossier Ministerie van Justitie en Veiligheid*. Retrieved from https://www.rijksoverheid.nl/documenten/publicaties/2017/11/21/introductiedossier-ministerie-van-justitie-en-veiligheid

Ministry of the Interior and Kingdom Relations. (2017). Een doorkijk in het personeelsbestand van het Rijk, *Wat vertellen de cijfers?* [An examination of the government's workforce, what do the numbers say?]. Retrieved from https://kennisopenbaarbestuur.nl/media/255579/eenplusdoorkijkplusinplushetpluspersoneelsbestandplusvanplushetplusrijk.pdf

Newman, J., Cherney, A., & Head, B. W. (2016). Do policy makers use academic research? Reexamining the "two communities" theory of research utilization. *Public administration review, 76*(1), 24-32.

Newman, J., Cherney, A., & Head, B. W. (2017). Policy capacity and evidence-based policy in the public service. *Public Management Review, 19*(2), 157-174.

Nutley, S., Morton, S., Jung, T., & Boaz, A. (2010). Evidence and policy in six European countries: diverse approaches and common challenges. *Evidence & Policy, 6*(2), 131-144.

Parsons, W. (2002). From muddling through to muddling up-evidence based policy making and the modernisation of British Government. *Public policy and administration, 17*(3), 43-60.

Rijksoverheid. (2020). Jaarrapportage Bedrijfsvoering Rijk 2019 [Annual Management Report for Central Government 2019]. Retrieved from https://www.rijksoverheid.nl/documenten/rapporten/2020/05/01/jaarrapportage-bedrijfsvoering-rijk-2019

Rimkutė, D., & Haverland, M. (2015). How does the European Commission use scientific expertise? Results from a survey of scientific members of the Commission's expert committees. *Comparative European Politics, 13*(4), 430-449.

Schön, D. A. (1984). *The reflective practitioner: How professionals think in action* (Vol. 5126): Basic books.

Smet, V. (2013). Evidence-based policy en de doorwerking van sociaalwetenschappelijk onderzoek: are we there yet? *Beleidsonderzoek Online, 12*.

Spruijt, P., Knol, A. B., Torenvlied, R., & Lebret, E. (2013). Different roles and viewpoints of scientific experts in advising on environmental health risks. *Risk analysis, 33*(10), 1844-1857.

Talbot, C., & Talbot, C. (2014). Sir Humphrey and the professors: What does Whitehall want from academics. *Manchester: University of Manchester*.

Torenvlied, R. (2019). OFFERTE; Naar meer evidence-based beleid - Stand van zaken van evidence-based beleid binnen JenV [TENDER; To more evidence-based policy - Current situation evidence-based policy inside Justice and Security]. *Universiteit Twente*.

Torenvlied, R., van der Berg, C., & Junjan, V. (2019). Top-Level Civil Servants' Use of Scientific Studies and Professional Knowledge: The Mediating Effect of Strategy Orientation. *Manuscript in preparation*.

Weiss, C. H. (1977). Research for policy's sake: The enlightenment function of social research. *Policy analysis*, 531-545.

Weiss, C. H. (1979). The many meanings of research utilization. *Public administration review, 39*(5), 426-431.

# Appendix 1 Basic descriptive statistics

**Table 1: Knowledge sources that are used: Distribution of the response**

| We base work or policies within my department on: | % (completely) agree | Distribution characteristics | | |
|---|---|---|---|---|
| | | obs | Mean | s.d. |
| Policy documents and regulations | 95,71% | 238 | 4,40 | 0,67 |
| Knowledge and experience from myself / direct colleagues | 92,76% | 238 | 4,21 | 0,66 |
| Knowledge and lessons from my professional field | 90,48% | 238 | 4,20 | 0,73 |
| Parliamentary documents | 81,53% | 238 | 4,20 | 0,99 |
| Instructions from supervisors | 73,70% | 238 | 3,84 | 0,86 |
| Case law | 68,84% | 238 | 3,96 | 1,20 |
| Independent qualitative scientific research | 64,26% | 237 | 3,77 | 1,14 |
| Policy evaluations by independent scientists | 63,39% | 238 | 3,73 | 1,09 |
| Independent quantitative scientific research | 56,69% | 237 | 3,67 | 1,22 |
| Policy evaluations carried out by the Ministry's policy departments | 54,67% | 238 | 3,53 | 1,10 |
| Scientific insights into what works and what does not, such as meta-analyses or systematic reviews | 53,42% | 238 | 3,64 | 1,23 |
| Outputs of data-driven work within the Ministry of JenV | 43,54% | 238 | 3,52 | 1,36 |

\Note. † Percentaged on categories (1 to 5); Answer categories: Completely disagree (1); disagree (2); Neutral (3); once (4); Completely agree (5); Don't know/ Not applicable (6); Missing responses (.).

**Table 2 control variables type of work**

| I perform the following activities: | % (completely) agree | Distribution characteristics | | |
|---|---|---|---|---|
| | | obs | Mean | s.d. |
| I mainly do advisory tasks | 78,22% | 202 | 3,87 | 0,67 |
| I support policy preparation / decision making | 64,36% | 202 | 3,59 | 0,88 |
| I work on policy documents | 54,46% | 202 | 3,37 | 1,03 |
| I give treatments or therapies | 35,65% | 202 | 1,06 | 0,33 |
| I take decisions and determine strategy | 34,16% | 202 | 2,94 | 1,10 |
| I mainly perform executive tasks | 18,32% | 202 | 2,56 | 0,98 |
| I have contact with ministers and/or politicians | 17,91% | 201 | 2,51 | 1,07 |
| I supervise staff | 17,82% | 202 | 2,35 | 1,28 |
| I work directly with target groups (clients, citizens) | 9,41% | 202 | 1,67 | 1,03 |
| I do research | 6,93% | 202 | 1,97 | 0,98 |

\Note. † Percentaged on categories (1 to 5); Answer categories: Never (1); Rarely (2); Sometime (3); Usually (4); Always (5); Missing responses (.).

**Table 3 control variables professional field**

| In my professional field: | % (completely) agree | Distribution characteristics | | |
|---|---|---|---|---|
| | | obs | Mean | s.d. |
| There is a lot of political dynamism | 86,50% | 201 | 4,29 | 0,91 |
| Do the media have a major impact | 83,99% | 202 | 4,22 | 0,89 |
| Is decision-making focused on the short term | 57,49% | 202 | 3,53 | 1,08 |
| Factual substantiation is only used selectively | 49,73% | 201 | 3,51 | 1,24 |
| Is knowledge from practice more important than scientific knowledge | 47,93% | 202 | 3,49 | 1,02 |
| Information and facts often contradict each other | 38,55% | 202 | 3,31 | 1,12 |
| Is little scientific knowledge available | 30,53% | 202 | 3,13 | 1,17 |
| Is there little agreement in policymaking about goals and/or means | 27,81% | 202 | 3,11 | 1,22 |
| Does the available scientific knowledge have an uncertain status? | 22,29% | 202 | 3,13 | 1,43 |

\Note. † Percentaged on categories (1 to 5); Answer categories: Completely disagree (1); disagree (2); Neutral (3); once (4); Completely agree (5); Don't know/ Not applicable (6); Missing responses (.).