Bias in theory and practice: a literature review of bias types and a case study of bias views at the Dutch Safety Board

Master Thesis (public version)

Comparative Politics, Administration and Society (COMPASS), Public Administration

Thesis coordinator: Pieter Zwaan

Wybe Janssen, 4783360

October 26, 2018

Note to the reader about the public version

Before this thesis was made public, I agreed that the Dutch Safety Board would review it, in order to check for correctness and to secure the anonymity of the participants. From this review, it was pointed out that I incorrectly referred to the organization as ‘research organization’ and to its employees as ‘researchers’. This implied that the organization does research with a scientific purpose. Therefore, I changed this in the text to ‘investigation-oriented organization’ and ‘investigators’, except for in the Q methodology’s statements and tables, as this would not fit. No changes were made to the content or data of the thesis, but the mentioned changes might still cause some confusion. I hope acknowledging and clarifying these changes here and adding a brief explanation of the difference in the introduction section mitigates some of this confusion.

Index

0. Management summary

This thesis investigates bias in both theory and practice with the research question "*What views on the existence, desirability and solutions regarding types of bias exist at the Dutch Safety Board and how can these views be explained?"*. The Dutch Safety Board serves as a case study.

In the theoretical framework, different kinds of bias are investigated. The types of bias discussed in the framework are issue bias, technical bias, methods bias and publication bias. Through the research question, the bias types are discussed for their existence and what defines them, their desirability, the solutions that exist for them and how views on each can be explained.

For the case study, views on bias are investigated at the Dutch Safety Board using the Q methodology. The Q methodology is the perfect method for investigating clusters of subjectivity. Based on the theoretical framework, 36 statements are created and shuffled, forming the Q-set. During the next step, called Q-sorting, the participants in the study, the P-set, are asked about their background first. Next, participants are tasked to divide the 36 statements over a score sheet, ranging with nine categories from 'least agreed' to 'most agreed'. Lastly, participants are asked to explain why they placed the statements on the far ends of the score sheet, to learn more about the participants opinion.

In the analysis of the Q methodology, four groups of participants are discovered through factor analysis. The factors are interpreted in the results section by using the explanations participants have given during the Q-sorting sessions. The first group is shown to think openness and issue bias are both very important. The second group are scientists by the book, who seem to hold scientific principles near and dear. However, this group is difficult to interpret, as it only contains two participants. The third group finds output and publication very important, but does not think publication bias is very problematic in context. The final group thinks that technical bias is the most important problem.

The differences between the groups of opinion are shown to not be explained by background variables and years of employment, as the groups contain participants of all backgrounds. However, there does seem some connection between time pressure and opinions on publication bias. There is also a connection between issue bias and topic complexity.

In the conclusion the research question is answered. It also shows that the bias types can be found to a limited degree in the groups of opinions and explanations of the participants. Lastly, there is discussion about the bias of this study to showcase its own flaws, the Q methodology as a research tool, problems with the statements and some other general positive and negative points. The thesis ends by making recommendations for future research.

1. Introduction

"Ninety percent of everything is crud" (Sturgeon, 1958, p. 66) and scientific research is no exception to this. Of course, scientific research has been known for its many virtues for humanity. It has guided all sorts of developments in many fields, ranging from Theoretical Physics to Sociology to Public Management and so on. Scientists and their research are also often consulted by both private and public organizations. This academic consulting involves actions like giving advice, solving problems and transferring knowledge (Fudickar, 2016). However, both scientists and their research are far from perfect. Due to all sorts of factors, it is simply impossible to do 'perfect' research. Factors that were found to influence this are, among others, study size, financial and other interests, effect sizes, and bias (Ioaniddis, 2005). In another study it was reportedly found that ninety percent of research could not be replicated (Begley and Ellis, 2012). With all of this in mind, the goal of doing research should not be to find a final answer and remain completely uncontested. The real goal is to make mistakes and keep learning from them in order to improve the methods of research (Castensson, 2015).

1.1 The role of individual bias

Of the aforementioned factors that can negatively impact research quality, bias is a particularly interesting one. Everything an individual scientist has experienced in their life is unique. No other person has the same background and social circumstances. However, this also means that every individual scientist has their own normative and societal ideas and views. All of these social circumstances and ideas influence the decisions made by individual scientists in their research. This influence goes from the chosen subject to the research question, from choosing theories and what to measure to selecting the right tools for the analysis. All of these choices influence the results and conclusions of research (Tholen, 2017).

1.2 The role of bias in context: the case of the Dutch Safety Board

The ways in which results and conclusions depend on individual choices shows how important bias is in doing research. However, the choices of an individual and their background are not the only source of bias. Context can also form a certain bias. Within a certain social environment or culture, for example within an organization, exist endogenous assumptions and preferences. Shared practices and choices by people within this environment are influenced by these assumptions and preferences (Thompson and Wildavsky, 1986). In research by Bohnet and Morse, it is even recommended to change processes within organizations themselves to tackle bias, instead of doing diversity training programs to change people on an individual level

(Bohnet and Morse, 2016). On the other hand, contextual bias could also be considered a good thing. Unique bias can provide an organization with new knowledge that is not found by others. By having a unique perspective, similar problems can be looked at from different angles, providing this new information (Pauleen and Murphy, 2005). This is also the case for the contextual bias of an organization like the Dutch Safety Board.

For this thesis, the Dutch Safety Board serves as case study material for what role bias plays within the context of an organization by asking employees about their views and experiences. The Safety Board serves a certain role in society and has a specific goal with their 'investigations', different than scientific 'research'. It is an independent organization aimed at improving the safety of people in the Netherlands (Dutch Safety Board, 2018). The unique role of the Safety Board makes it a great case to investigate with a case study.

As bias is everywhere, those who work at the Dutch Safety Board have to deal with bias as an organization as well. The results of the organization's investigations are, in a way, biased to look at events, factors and facts that are in some way related to safety. However, this bias provides the Safety Board with unique results that bring the perspective of safety improvement to the attention of the public, media, organizations and politicians alike. As such, this contextual bias is not negative per se (Pauleen and Murphy, 2005). Also, the role of biases and the importance of neutrality and independence are well understood (Dutch Safety Board, 2018).

1.3 Research question

There has already been much research on bias in general, from investigating what types of bias exist to how to deal with bias. However, it would be very interesting to see how bias is seen in practice and how previous theoretical research holds up to case study investigation. Therefore, this thesis is focused on investigating the views on bias that can be found at the Safety Board, after having reviewed existing literature for theoretical types of bias. Following the goal of this thesis, the following research question is posed:

What views on the existence, desirability and solutions regarding types of bias exist at the Dutch Safety Board and how can these views be explained?

To answer the research question, four sub-questions are asked:

What does existing literature teach us about views on the existence, desirability, solutions regarding different types of bias?

What factors explain the differences in the views on the existence, desirability and solutions regarding types of bias?

How can the Q methodology work to uncover views on the existence, desirability and solutions regarding types of bias at the Dutch Safety Board?

What views on the existence, desirability and solutions regarding types of bias are uncovered from using the Q methodology at the Dutch Safety Board and how can they be explained?

1.4 Relevance

Research on this topic has scientific relevance and societal relevance, which are briefly discussed in this section.

Scientific relevance

The scientific relevance of this research is to learn more about bias, its effects on research and how it is viewed within the context of an organization. The existence of bias is well known, but there is less knowledge about what types of bias exist and how it is experienced and viewed in practice. By learning more about views on bias in practice, this research explores how well theoretical views on bias translate to the practical field of an investigation-oriented organization, especially one with a unique role, like the Dutch Safety Board. This presents a different angle than one solely focused on theory and adds a new perspective to the field.

Societal relevance

As the case study presents new perspective to the science surrounding bias, the reverse also applies. By reviewing the theoretical field and comparing it to practice, organizations like the Dutch Safety Board can learn more about bias and how it applies to their work. The insights gained can spread awareness about bias and help improve research and investigation quality.

1.5 Reading guide

To answer the first two sub-questions, a literature review explores different types of bias that are discussed in scientific literature and what explanations exist for different views. Second, the *Data and methods* section mainly discusses the research approach of the Q methodology for the investigation of bias at the Dutch Safety Board and answers the third sub-question. Next, the *Analysis* of bias at the Dutch Safety Board is presented, explaining how the analysis is performed. The *Results* section answers the fourth sub-question. It shows how bias is viewed and explores explanations for these views. As *Conclusion*, the research question is answered by looking at the results of the sub-questions, followed by a *discussion* about the thesis.

2. Literature review

The literature review aims to see what types of bias are found in previous research on the subject. This review itself exists of four parts, in the same order as presented in the main research question, each structured by bias types. First, several main types of bias are distinguished from the literature. This section also discusses some of the explanations why certain bias occurs, which is also discussed in the fourth section. Second, the desirability of the types of bias are discussed. As was briefly discussed in the *Introduction* chapter, bias is not necessarily bad. The example of uniquely gathered information shows that sometimes bias can be a good thing as well, though often some bias is deemed less problematic and would be too much of a hassle to deal with. In the third part of the review, it is discussed in what ways it is possible to cope with the different types of bias, as certain problems ask for certain solutions. The fourth section of the review discusses what factors can explain the differences in the views on bias that people have, also showing why bias occurs. At the end of the chapter, all types of bias are put together in an overview, showing the types of bias, their desirability, possible solutions for the types of bias and the factors that influence people's views on them.

2.1 Types of bias

There are many names for bias types in literature. However, many are kind of similar or even completely the same. Therefore, for every type of bias reviewed in this section, the similar types of bias are discussed together with each other. It is made clear what the similar types of bias are originally called and explained why these are put together with another type of bias. The reviewed types of bias are issue bias, technical bias, methods bias and publication bias.

Issue bias

Issue bias concerns the choice of researchers to investigate certain issues over others. This also means that only certain types of evidence are selected and looked at when investigating an issue. As there are many issues to choose from, there also are many kinds of evidence to choose from. The issues and evidence a researcher decides to look at greatly influence the outcome of research and what kind of knowledge is gained (Parkhurst, 2016; Parkhurst, 2017). It basically means that you cannot learn about things you do not ask about. This goes for researched topics, theories and the choice of what type of evidence is looked at (Parkhurst, 2017; Tholen, 2017). The simplification can create a distorted view of reality. If policymakers draw their conclusion from this distorted view, the policy based on it may not be adequate in practice (Ferretti, 2018).

Another problem with issue bias is that it can also be purposefully used to obscure values that are not researched. This way, attention will be focused on desired issues and factors that will show only one the side of the story. This is especially problematic if it is done systematically. And even in cases in which this is not done on purpose, it is very possible for a society to, for example, systematically not ask questions concerning certain minority groups, leading to a lack of evidence created about the minority groups (Parkhurst, 2017).

Next to choices of topics and evidence being subject to issue bias, the same is also possible for the interpretation and use of research results. After research has been performed, the biased choice can still be made to prioritize certain parts of an outcome over others, for example for political means. An interested party could cherry-pick the one piece of subject that they want to lobby for or is favourable to their cause and use only that for the interpretation of research outcomes. In reality, however, the research outcomes could be way more nuanced as a whole or the chosen piece of evidence could be of minor importance among the other pieces of evidence (Parkhurst, 2017).

There are several other types of bias that can be seen as a part of issue bias. These are confirmation bias, content bias and political bias. Confirmation bias is the seeking or interpreting evidence in such a way that it conforms with, supports and confirms your own beliefs. Researchers can search specifically for evidence that supports their hypotheses or interpret it as such, or select certain topics or theories to support their ideas, just like with issue bias. During a research project, information gained at the start of a research project has more weight for a researcher than what is found later. The researcher will be likely to want to seek confirmation for the findings early on in the research project. This same goes for policies and politicians or policymakers defending them by seeking one-sided information (Nickerson, 1998). Just like with issue bias, evidence of a certain issue or type is sought or interpreted in a way that conforms with the preferred outcome.

Content bias refers to favouring one side of a story over another (Entman, 2007). This, again, fits within issue bias, as this concerns favouring an issue. It basically refers to the same.

Political bias reflects a decision-maker who has their own ideas of how things should be done, different than what the society they lead would want. They might even see some own benefits in war, while their society would not support this (Jackson and Morelli, 2007). Again, this corresponds with issue bias, as in this case the specific personal beliefs of the biased political leader are used in their decision making.

In short, issue bias concerns the conscious or unconscious preference for one issue over another. This goes for creation of evidence based on biased issues, biased selection of issues to pay attention to and biased interpretation of research results. People with this bias can seek confirmation for what they already believe in, explicitly favour one side of a story over another or be biased through their political perspective. In this subsection, it was made clear that there are many forms of bias that come down to preferring one issue over another. All variations discussed in this subsection are referred to as issue bias for the rest of the thesis.

Technical bias

Technical bias is similar to issue bias in that it can concern the biased creation of evidence, selection of evidence and the use or interpretation of it. The way this type of bias works, however, is different. Technical bias concerns performing research incorrectly, possibly leading to flawed results and conclusions. This can, for example, be due to a lack of impartiality, as good research practices have to be impartial, regardless of the researchers' own personal values. Researchers can also do things such as modifying results to create more positive outcomes, have an unrepresentative population sample, or cherry-picking by highlighting certain evidence for the sake of creating desired results. Whereas with issue bias the issues on which to select evidence is cherry-picked, with technical bias it is the evidence itself that is cherry-picked from what is found with no regard for what issue it is selected for. The described technically biased practices can mean that evidence and research design are manipulated in such a way that it will lead to the desired conclusions of an interested party (Parkhurst, 2016; Parkhurst, 2017).

In the interpretation and use of evidence it is possible to have invalid conclusions drawn from research results. This and all other biased practices are not exclusively done by malicious intent. Drawing invalid conclusions can, for example, be done because people do not understand the statistics or methods well enough to draw the correct conclusions from their research results. An example of this is that many people will interpret a correlation as if it is the same as causation (Parkhurst, 2016; Parkhurst, 2017). Due to the high frequency of this mistake, the phrase "correlation does not imply causation" has become famous. It is a good reflection of how important technical bias is, even outside the scientific realm.

Selection bias can be seen as a form of technical bias. Selection bias concerns selecting cases or participants that are not representative of the population because of the selection process itself. Researchers can be tempted to select cases that are extreme cases, as these can give new knowledge in an unknown situation. However, these results should not be generalized to the larger population, as they are biased and not representative of the larger population

(Collier and Mahoney, 1996). An example of this is if you would use a sample that is only filled with people that go to college to say something about the entire population of young people, including the people who do not go to college. Young people going to college have been pointed out to be different than those who do not, so this would provide you with a biased sample, which is not representative of the larger young population (Smart, 1966). The different types of selection biases can also be seen as a kind of confirmation bias, when the selections are biased towards confirming the researchers' own viewpoints or hypotheses (Nickerson, 1998).

This subsection has shown that technical bias is a way to create, select or interpret evidence incorrectly by having flawed research design. Examples of how this happens are a lack of statistical knowledge, cherry-picking positive outcomes to get some results or selecting a biased sample. The focus of technical bias is on the reliability of research. Technically biased research tends to be less reproducible and less representative than its less biased counterparts (Golafshani, 2003).

Methods bias

There are two important concepts related to methods bias, namely common methods bias (CMB) and common methods variance (CMV) (Podsakoff et al., 2003). CMB is "*the degree to which correlations are altered (inflated) due to a methods effect"* (Maede et al., 2007, p. 1). This means that the results of research are influenced by the way methodology is used in research. However, results should reflect a population and what goes on within it. CMV is "*a form of systematic error variance and can cause observed correlations among variables to differ from their population values"* (Maede et al., 2007, p. 1). In other words, systematic problems in the measurement of variables causes faulty results, that do not match the population. The difference between the two is that CMV is a potential cause of CMB, but does not necessarily lead to significant problems. CMB, on the other hand, is the presence of bias problems an sich (Maede et al., 2007).

Methods bias can be seen as part of technical bias, but is discussed separately, as it is a very distinct aspect of technical bias (Parkhurst, 2017). Whereas technical bias mainly discusses flawed research design as the cause for bias, methods bias concerns methods causing variance in results instead of independent variables. It is always present to a degree and does not mean research is flawed (Podsakoff et al., 2003; Maede et al., 2007). Technical bias has problems in its reliability. Methods bias has problems in internal validity, as methodically biased research tends to have results influenced by methodological factors instead of by independent variables

(LoBiondo-Wood and Haber, 2017). Methods choice can cause type 1 errors, i.e. false positive results, or type 2 errors, i.e. false negative results, to appear (Tholen, 2017).

Podsakoff and his colleagues (2007) discuss four types of possible sources of CMV. These are common rater effects, item characteristics effects, item context effects and measurement context effects. Common rater effects mean that the CMV is caused by the fact that the same respondent is providing the measure of the variables (Posdakoff et al., 2007). An example of this is social desirability, which is the "*tendency of an individual to convey an image in keeping with social norms and to avoid criticism in a 'testing' situation"* (Hebert et al., 1995, p. 389). In practice, this bias has been shown to affect what people report when it comes to, for example, their dietary intake. People would report a lower dietary intake than they actually had, as a higher intake is not socially desirable (Hebert et al., 1995).

The second type, item characteristics effects, is concerned with variance caused by respondents picking items for the way the items are presented (Podsakoff et al., 2003). For example, the positive or negative wording of items in a questionnaire can cause respondents to answer more positively or negatively, thus skewing results (Schriesheim and Hill, 1981).

Third, item context effects make respondents connect certain items to each other and respondent accordingly to those connections (Podsakoff et al., 2003). An example of this is the effect of item embeddedness. Respondents tend to use contextual cues, such as the questions surrounding the one at hand, and assess other questions by using those cues. If a question is embedded in negative questions, respondents tend to see that question more negatively as well. The same goes for when a question is embedded in positive questions (Harrison and McLaughlin, 1993). In other words, respondents are biased to respond to questions in a way that the item context unintentionally conditions them to.

The last type of source for CMV is measurement context effects. This refers to variance resulting from the context in which measurements are taken (Podsakoff et al., 2003). For example, it can make a difference if an interview is done face-to-face or through a telephone interview. In a study by Holbrook and her colleagues, respondents were more suspicious and gave more socially desirable answers when interviewed through a telephone interview than face-to-face. However, they do add that these results should not be blindly generalized, as the length of the interview and other factors may have had an effect on this as well (Holbrook et al., 2003). The choice to do research in a more quantitative or qualitative way is also relevant for this, as qualitative research gives more room for people to answer questions. This is similar to choosing the type of evidence, like with issue bias (Parkhurst, 2017).

With all these types of sources explained, it is clear that the intricacies of methods can act as the cause of the variance in results instead of real connections of factors in the population. Obviously, it is not the intention of researchers to have the methods themselves account for correlations and results. CMV and CMB show how important methodological choices are in the research process and that researchers should pay very close attention to every methodological choice they make, instead of just using a standard technique and assume everything should be fine.

Publication bias

Publication bias is often used as a collective term for publication bias, reporting bias (i.e. selectively reporting research outcomes), inclusion bias (i.e. selectively including research in databases) and some other types of bias (Bax and Moons, 2011; Turner, 2013). However, it would be strange for the collective term to be the same as one of the parts it is describing. A better collective term is 'dissemination bias' (Bax and Moons, 2011). To avoid confusion, this thesis focuses on publication bias as something separate from reporting bias and inclusion bias.

Publication bias by itself concerns selectively publishing studies, with a tendency to feature positive or extreme results (Nicholas et al., 2000). This means that some research is more likely to be published than others. The problem with this is that researchers will be pressured to provide significant results or other results that would lead to publication. This is likely to lead to an increase of other kinds of bias in research. Some researchers can even be pushed to report some arbitrary significant effects in order to have their research published (Gerber and Malhotra, 2008). This means that publication bias can lead to reporting bias, as arbitrarily reporting significant results is a form of reporting bias (Bax and Moons, 2011). Publication bias has also been attributed to preventing the advance of a field, as studies with nonsignificant results are often not published (Nagakawa, 2004).

The media's bias in reporting scientific discoveries has been an annoyance for many scientists. Sensationalism, inaccuracy and under-reporting are some of the things that can be problematic. Journalists need something news-worthy and thus there is a bias in what they will end up reporting in, for example, a newspaper (Bauer and Bucchi, 2008). Sensationalism is also seen in a problem in newer developments. As science is popularized, the boundaries between science and opinion is blurred and bias has an increasingly open field in the scientific and semi-scientific realms. This is especially the case with the freedom on the internet to claim things as science that are not actually science and are in reality very biased (Brumfiel, 2009).

There are also bias problems in policy advise. Policy advisors are expected to deliver relevant results for policymakers and have the results be understandable for these policymakers. Advise has to involve a clear distinction between positive and negative and dumbing down complex cases. The pressure of publishing advise for certain policymakers makes researchers provide less neutral and complex results, as these are otherwise less desired (Tholen, 2017).

Publication bias is a problem both in the scientific community as the science reported in news sources and the internet, but also to policymakers. Publication bias by any source can limit the developments of a certain field or topic and produce poorly executed research. These developments can even pressure researchers to produce significant results, even if there are none. In news sources and the internet, the accuracy of science can be a major problem and cause the general public to have grave misconceptions about science.

2.2 Desirability of bias

As was briefly addressed in the *Introduction* chapter of the thesis, bias does not have to be very problematic. There are several scenarios in which bias can be a good thing, or at least less problematic. This section discusses under what circumstances having bias might not be a completely bad thing after all.

Desirability of: Issue bias

As was briefly discussed in the *Introduction* chapter, the biased choice in what to gather evidence about can also be a good thing. Bias for certain topics can create an opportunity to develop much knowledge on that topic. Especially interesting is to see how a certain culture can develop a unique kind of knowledge by being culturally biased. By having certain culturally biased standards and thoughts, different knowledge can be gained in one culture than the other. This should be seen as an advantage rather than a disadvantage. Culturally biased influence is precisely where new ideas come from (Pauleen and Murphy, 2005).

As it is not possible to completely get rid of bias, it would be better to use it as an advantage. By having responsible bias, researchers can write down their biases, so that others can put their stances against them. This way, bias stays and is beneficial instead of a problem (Feinberg, 2007). This is discussed further in the discussion of how to deal with issue bias.

Desirability of: Technical bias

It is difficult to argue that technically biased and flawed research practices are desirable. However, the results and findings of this kind of research are usually not completely bogus. Although the research on its own might not be useable, it is possible to look at evidentiary

fragments. In other words, research is evaluated for what it does have. It is very possible to find pearls of wisdom. These pearls of wisdom can, for example, be used in a synthesis of research and findings and be worth something in the process. Of course, this implies that problems in the flawed research are taken into account with its evaluation (Pawson, 2006).

Desirability of: Methods bias

Methods bias is argued to badly influence research results. However, attempting to get rid of all methods bias will cost more time than it is worth (Luttrell, 2000). For example, even with an infinite sample size, there is still possibility of effects to not correspond to the population (Serlin and Lapsley, 1985). It is better to keep some methods bias in a research project by applying 'good enough' methods and simply name encountered methodological problems, instead of trying to eliminate all of them. In fact, it is impossible to eliminate all bias, so trying to do this completely would be a fool's errand. Researchers should accept their mistakes, so they and others can learn from them (Luttrell, 2000). Of course, this does not mean methods bias and methodological issues should be ignored.

Desirability of: Publication bias

When it comes to publication bias, there are not always problems. In some cases, publication bias only leads to very few studies not being included. It can also be the case that there is some sort of selection in what gets published and what does not, without any negative consequences and little bias as result (Bax and Moons, 2011). For example, with publication bias, smaller studies with lower power are often filtered out. However, these studies are less likely to measure true effects of a population, so it may in some cases be beneficial that there is some sort of publication bias (Button et al., 2013). In other words, some preference in what is published is perhaps desired, as, for example, it may not always be desirable to publish low quality research.

2.3 Dealing with bias

As bias is often a problem in doing research, there have been many people investigating how to deal with bias and its negative impact on research. There are ways to deal with each of the previously described types of bias. Some of these prescriptions are the same or similar for different types of bias. This goes for issue bias and technical bias, and for technical bias and methods bias.

Dealing with: Issue bias

There are at least three concepts that can help to deal with issue bias. The first concept that helps dealing with issue bias is the so-called "*good governance of evidence*" (Parkhurst, 2017).

The good governance of evidence is concerned with governance arrangements and processes as a means to reach collective decisions. This good governance exists of eight parts. Five of these parts are particularly well applicable to issue bias. The first of these is appropriateness. After a policy decision has been chosen, evidence is selected based on its appropriates to the subject. This concerns whether it addresses relevant social concerns, is useful for policy goals and is applicable to the local context. Second, the people who set the rules for evidence selection should take a stewardship role. This means they are authorized, accountable to the public and will resist the influences of those without mandate or accountability. Third, the final decisions should stay with a public representative figure. Fourth, information should be transparent for the public. This includes clear and accessible ways for the public to learn what evidence is used, how and why. Fifth, deliberation should take place in order to give attention to multiple and different opinions and values. This must also include the concerns that are not included in the final decision (Parkhurst, 2017).

The second possible option to deal with issue bias deliberative inquiry. Not only is it similar to Parkhurst's fifth point discussed in the previous paragraph, but it has also received much attention recently in the scientific community. Deliberative inquiry "*is to deliberate about the issues as perceived by diverse stakeholders, and provide an opportunity to challenge ideas, reveal misconceptions and establish where mutual understandings exist"* (Kanuka, 2010, p. 102). Thusly, it can guide the choices of policymakers, politicians and researchers in deciding what issues to focus on. Next to knowing what stance people take, it can also be used to explore the reasons behind the positions of people. This way, light is shed on more issues from more sides (Carcasson and Sprain, 2016). However, researchers must be careful with this solution, as the general public is often uninformed about scientific practices or may want to lobby for specific interests (Tholen, 2017).

The third option to deal with issue bias relates to a point made in the section about the desirability of it. This point is to deal with issue bias in a responsible and beneficial way by writing it down and letting others compare it with their own biases and standpoints, as it is not possible to get rid of it completely (Feinberg, 2007). This option works in tandem with the other two. Any remaining bias can be named and discussed, in order to use it responsibly.

Dealing with: Technical bias

The arrangements and processes of the good governance of evidence can also be applied to deal with technical bias. Four of these apply to technical bias. First, there is rigor, as in the rigorous

gathering of evidence, whilst avoiding cherry-picking. This method tries to make sure that all relevant evidence is collected and not a biased and selective portion (Parkhurst, 2017).

Second, evidence quality can be used to deal with technical bias. The criteria for the governance of evidence quality should be in line with the kind of research that is being undertaken and the nature of gathered data. The criteria would otherwise not fulfill their purpose, which is checking the quality of the research being undertaken. To illustrate, qualitative and quantitative research have to meet completely different criteria (Parkhurst, 2017).

Third, transparency involves research practices being completely open to public, so that everyone can see how it was done. This allows people to see whether or not the research process was handled correctly (Parkhurst, 2017).

Lastly, contestability adds that technical evidence and research choices should be open to be challenged for their quality and correctness. This requires transparency in order to work. Then, people can see how research is handled and make remarks (Parkhurst, 2017).

Dealing with: Methods bias

There are many possible recommendations for dealing with methods bias. To be sure variable measures are correct, it is possible to use different data sources to obtain these measurements by ways of triangulation (Podsakoff et al., 2012; Johnson et al., 2007). It is also possible to triangulate by letting multiple researchers give feedback at research, including different theories (Johnson et al., 2007), or using different kinds of methods (Oliver-Hoyo and Allen, 2006) in order to avoid problems with methods bias.

Second, a researcher can separate questions temporarily with a time-delay, proximally by having increased physical distance between different measures, or psychologically by creating distance between measures with a story. This can prevent item context effects, in which people link different questions with each other. Scales can be changed so that they are different from scales of other questions and do not include ambiguous questions. Item wording can be changed so to have no negativity, positivity or social desirability embedded within a question that could push respondents to answer accordingly. The way these techniques are designed aims to make sure respondents are able and motivated to answer the questions and minimize the difficulty of tasks respondents have to perform (Podsakoff et al., 2012; Podsakoff et al., 2003).

Third, it is possible to statistically check if there is methods bias by looking at the interactions between measures and by using factor analysis. For example, if items are very strongly correlated with each other, while they should not be, it could mean there is methods bias (Podsakoff et al., 2012; Podsakoff et al., 2003).

Methods bias can also be treated similarly to technical bias by having good governance of evidence. By being transparent and open for contestation, it is possible to have peers give constructive criticism (Parkhurst 2017). This is especially good during the research process by using triangulation of people, as was stated at the beginning of this subsection.

Dealing with: Publication bias

A first way to deal with publication bias is to check how many published studies have significant results. If the percentage is high, chances are that there is a bias for publishing studies with significant results (Thornton and Lee, 2000; Gerber and Malhotra, 2008). It is possible to use a funnel plot in order to check if small studies with nonsignificant results have not been left out (Thornton and Lee, 2000). It is also possible to use statistical tests for the funnel plot, instead of just seeing how it looks (Thornton and Lee, 2000; Peters et al., 2006).

It is also possible to try preventing publication bias altogether. Registries can be used to search for both published and unpublished research trials. Registration of trials happens before results are written, so chances of publication bias are low (Thornton and Lee, 2000; Turner, 2013). However, as this is time consuming, it would be better to have some sort of editorial policy that aims to publish all articles of good quality (Thornton and Lee, 2000). This could go together well with the good governance of evidence, as both will need some sort of intermediary to perform such a task. This can, for example, be an independent governmental actor, or it can be done by self-regulation of journal editors or reviewers (Thornton and Lee, 2000; Parkhurst, 2017; Turner, 2013).

2.4 Factors explaining views on bias

There are several factors that explain the views on bias people have. As was addressed in the *Introduction* chapter, not only the past of people themselves influence their views on bias, but the context in which they work as well. This section explores how people and context could explain the varying presence of views on different types of bias. First, a general explanation for the types of bias is discussed. Then, separate explanations for the four types of bias are provided.

General factors explaining views on bias

The way people act is influenced by their surroundings. By interacting with other people in their surroundings, people are socialized. This means that they receive and share ideas, values and ways of life. This is often done in social groups, but can also happen through more loose interactions (Ochs, 2000). As such, people are socialized in their field of study and gain and share social values with their peers. Every discipline has its own idea on what is important to

know and what the most important things are in doing research (Tholen, 2017; Sarewitz et al., 2004). As such it can be said that all researchers are biased in their views on bias, based on past experiences of their own and the surroundings of their field of study. This applies to people's views about the importance of issue choice, technical experience, methodological choices and the choice of what is and is not published. When awareness of the harmfulness of certain types of bias are encouraged within a study field, this is ingrained in people's understanding. The people within a study field socialize with each other and strengthen their awareness and views on certain types of bias. Of course, the opposite is true as well. If certain bias is ignored or seen as not very important in a study field, this will influence the people within this study field through socialization as well. In this latter case, it means that these people are less aware of bias or see it as less important.

Lastly, it is assumed that people will be more aware of bias and its problems if the bias is more present. To explain, imagine a situation in which argumentation is very one-sided, suffers of tunnel vision or is otherwise very biased. These biased aspects will be very clear in the way people interact with each other, do their work and write their results and conclusions. With more bias come more of these signs of bias. With more signs of bias, people are more likely to notice it, be it in conversations or in their work. Therefore, the factors that explain the presence of bias and are discussed here also help to explain more awareness of the importance of certain bias. Therefore, this section discusses factors explaining people's views and awareness of bias and factors explaining bias altogether.

Factors explaining views on: Issue bias

One reason that certain views on issue bias have been appearing is the recent upsurge and popularity of evidence-based policy (EBP). Because of the idea that all policy decisions should be based on research, in order to combine policy and science, scientific research can become biased to address specific topics and analyze the effects of specific factors. Policymakers and politicians will want to find evidence for their policy, while opposing the policies and ideas of their opponents (Wesselink et al., 2014; Parkhurst, 2017). It is the political choice that decides discussed issues and, consequently, which side of the story is shown (Parkhurst, 2017).

Against the background of EBP, there are three mechanisms that contribute to a greater awareness for issue bias and technical bias. These mechanisms are complexity, contestation and polarization. Complexity applies to multifaceted problems, which gives room for issue bias. With more possible sides of evidence, the chance of evidence getting excluded grows. It also means that more uncertain pieces of evidence could get ignored (Parkhurst, 2017). In practice,

deadlines can push people to have less room to address these multiple possible sides of evidence, creating more room for bias.

Contestation concerns the diversity of involved stakeholders present in the policy arena. As these problems are deemed more important, chances are greater that stakeholders will want their own view taking the upper hand, pushing their own take on the issue (Parkhurst, 2017). This could also be seen as a necessary part of choosing a relevant issue, though not supported from a neutral scientific standpoint.

Polarized problems are problems with a lot of options with highly polarized problems having little to no middle ground for people to position themselves. Especially with more highly contested issues, it is more likely that issues will be handled from one of the more extreme positions (Parkhurst, 2017). Same as with the contestation mechanisms, people could find it a good thing to choose a strong position, while a more scientifically based researcher would prefer a more neutral standpoint.

Complexity, contestation and polarization can lead people to view issue bias as something that has to be present for contextual reasons.

Factors explaining views on: Technical bias

Just like issue bias, the political playing field has a lot of effect on technical bias and the way it is shaped and viewed. Evidence can be manipulated in order to achieve certain results, depending on party of interest. Interest groups have a certain policy position they want to have defended. For this purpose, things an interest group can do are funding and publishing technically flawed research that supports their cause or criticize and suppress research that opposes their position. Second, research can also be manipulated in order to show results, if there is pressure to have to show these. On the other hand, certain evidence that is disadvantageous to the government can also be hidden, if it would hurt interest groups' cause. Third, political actors or stakeholders can actively undermine research if it would work in the favor of their desired agenda. This way, they can have their own views and policies, based on flawed research, be the only ones available (Parkhurst, 2017). So, for political benefits, this bias can be viewed as something that is a necessary evil. However, scientifically, practices like this will be frowned upon.

As mentioned before, complexity, contestation and polarization apply to technical bias as well, though differently than they did for issue bias. Complex cases can push people to search for shortcuts in order to be able to finish research. Furthermore, much uncertainty can cause researchers to refer to past results, while these may not really address the issues at hand

(Parkhurst, 2017). These shortcuts can be preferred by certain researchers when dealing with deadlines, though many do not condone these practices.

Contested cases, in which issues are more important for stakeholders, make research quality important. As certain issues get pushed forward, the scientific accuracy and research quality gets left in the dust. Lastly, polarized problems will have more incentive for manipulation, as an issue becomes a competition of one extreme against the other (Parkhurst, 2017). So, in short, political involvement and the involvement of interest groups can cause incentive to create flawed research and have people see this as acceptable.

As complexity, contestation and polarization of issues increase, it is more likely for research to become one-sided. People become prone to collecting evidence from a certain point of view and bias becomes very noticeable. This noticeability of bias is likely to lead to more awareness and stronger views.

Factors explaining views on: Methods bias

As explained earlier, it is common methods variance (CMV) that functions as a cause for common methods bias. All kinds of conditions and forms explain why variance is created, which could ultimately lead to bias (Podsakoff et al., 2003). However, this does not explain how this process works and what the mechanisms are to explain it. The four mechanisms generally accepted to explain methods bias are the capability of a respondent, task difficulty, the motivation to answer accurately and satisficing (MacKenzie and Podsakoff, 2012).

If a respondent has low capabilities, for example bad verbal skills or lack of experience, the respondent will have more difficulty understanding a question and giving an accurate response. Second, task difficulty can make it difficult for respondents to respond accurately to a question. Examples are complex or ambiguous questions. This shows the importance of item characteristics. Third, the motivation to respond to questions can, for example, be affected by the respondent having little personal connection to questions or having lengthy scales. Lastly, if similar scales are repeatedly used, items are grouped together, or previously given responses are available to respondents, respondents are tempted to take less time to think about questions and just give satisficing responses (MacKenzie and Podsakoff, 2012).

So, in short, of the four sources of common methods bias can be explained by the four mechanisms named here. Common rater effects can be explained by demotivated respondents. Item characteristics effects can be explained by respondents' abilities and task difficulty. Item context effects can be explained by respondents giving satisficing respondents. Measurement context effects can be explained by respondents' abilities and task difficulty. Overall, these are

not things that researchers try to perfect. As was explained with the desirability of methods bias, there is a certain degree of acceptance of imperfection in research design.

Next to these four sources of methods bias, as this methods bias concerns caring about having a good research setup, the factors that explained views on technical bias can also be applied here. If the political playing field creates incentive for flawed research, it is very likely that incentive for ignoring methods bias will appear as well, as it is well connected to research quality. This means researchers who have certain political incentive could care less about working on making sure the research setup is correct. However, similar to technical bias, more presence of methods bias could also lead to stronger opinions about its problems.

Time pressure could also play a role preventing a good research setup. It can lead to less accuracy of human judgement and less effort in finding alternative strategies (Edland and Svenson, 1993). It is possible that time pressure plays a role with views on methods bias. As preventing methods bias means doing something about the four described mechanisms, time constraints might have people see these as less important.

Factors explaining views on: Publication bias

There are several explanations for views on publication bias. One possible scenario is that reporting bias leads to publishing bias. It is very possible that researchers want to publish significant results. For this end, they can keep changing their methods in order to find some sort of significance and cherry-pick the trials that were significant afterwards. If the peers who review the manuscript do not know about this, it is likely that the biased report is published (Turner, 2013). It is possible such fraud is not discovered after both peer review and replication (Thornton and Lee, 2000). Even with this in mind, people may still prefer to report biased results in order to get their research published and receive attention for their work.

A reason that research with no significant results or negative results are not published can be because it is not very exciting reading material. Negative research that opposes previous research is also less likely to be published and will be more heavily criticized. Furthermore, it is possible for sponsors to oppose negative research about their products or lobby (Thornton and Lee, 2000). Therefore, some may prefer publishing with bias in order to uphold a journal, by keeping readers and sponsors positively interested.

2.5 Review overview

This section has reviewed four types of bias, namely issue bias, technical bias, methods bias and publication bias. The findings of the review are summarized in table 1.

Table 1. Summary of bias types

Bias types	Description	Desirability	Solutions	Explanations of views
Issue bias	Biased preference for one issue over others	Unique knowledge through bias	Good governance of evidence and deliberative inquiry	Evidence-based policy (EBP); complexity, contestation and polarization; primacy effect and belief persistence
Technical bias	Bias due to flawed research design	Finding pearls of wisdom in flawed research	Good governance of evidence and reweighting samples	Political processes, interested parties; complexity, contestation and polarization
Methods bias	Bias due to a methods effect instead of the data	Bias is desirable over time spent trying to get rid of it. Adopt 'good enough' methods	Triangulation; creating distance between items; statistical checks; transparency	Common method variance, task difficulty, and respondent's motivation and capabilities; political processes; time pressure
Publication bias	Tendency to selectively publish some articles over others (e.g. only significant effects or large studies)	Bias might only miss out on a few studies. Smaller studies less likely to measure true effects	Check amount of significant published studies, funnel plot, prevention with registration or editorial policies	Reporting bias; pressures to deliver exciting reading material; sponsorship

Source: author's summary

3. Data and methods: measuring with Q at the Dutch Safety Board

This chapter discusses the Q methodology, validity and reliability. The first section discusses the steps involved in doing research with the Q methodology. The second section discusses gathering data and contains a brief discussion on the Dutch Safety Board with my own relevant experiences at the organization. The sample is discussed as the 'P-set' in the next chapter. The last section briefly discusses the validity and reliability of the thesis.

3.1 The Q methodology

The Q methodology is a way to investigate subjectivity and opinions. It looks at the similarity and dissimilarity of opinions on certain topics. Q focuses on relations between opinions and the clustering of opinions (Brown, 1993; Van Texel and De Graaf, 2005). It does not focus on a quantitative amount of people's views, so there is no need to have a large sample (McKeown and Thomas, 2013). It even functions properly with a sample of one person or a non-random sample (Brown, 1993; Van Texel and De Graaf, 2005; McKeown and Thomas, 2013). The Q methodology fits the case of investigating the views on bias well. It can be used to explore how people's views on bias cluster in practice. These clusters can be compared to the theoretical types of bias to see if they match or, if they do not match, what the clusters do represent.

Using the Q methodology requires following five steps. The steps are defining the concourse, making the Q-set, selecting the P-set, Q-sorting and factor analysis (Brown, 1993; Van Texel and De Graaf, 2005). Table 2 shows a summary of the five steps. The steps are discussed in more detail below.

Table 2. Steps of the Q methodology

Step	Explanation
Defining the concourse	Assembling statements based on discourse. They can be gathered from conversations, art, literature and more.
The Q-set	Preparing roughly thirty to sixty statements from the discourse, representative of the possible range of opinions.
The P-set	The participants in the study. Some effort should be made to include different kinds of people, but it does not matter much.
Q-sorting	The P-set is asked to rank the Q-set from 'least agreed' to 'most agreed' and distribute the statements cards on the sheet prepared by the researcher. An interview about the more extremely placed statements follows afterwards.
Factor analysis	A factor analysis with varimax rotation is performed with the results of Q-sorting. The results show groups of participants and how certain statements can help explain similarity or dissimilarity between groups of participants.

Source: author's summary

Step one: Defining the concourse

The concourse is *"the flow of communicability surrounding any topic"* (Brown, 1993, p.94). It is the relevant discourse of thoughts and ideas on a certain topic. Defining the concourse is the process of assembling statements from this discourse. These can be collected through interviews, observation by the researcher by acting as participant, popular media, literature and scientific literature (Van Texel and De Graaf, 2005). Even paintings, photographs and videos can be used as part of the concourse (Brown, 1993).

Step two: The Q-set

Step two of the process takes roughly thirty to sixty statements of the concourse and puts them together into what is called the Q-set. The Q-set is representative of the range of different opinions on a topic and is structured accordingly. The structure can be determined by theory or researchers themselves. Because of this, the Q-set is often unique to a research project, even when it is based on the same concourse. In practice, this does not present problems as the results are basically the same regardless of the Q-set (Van Texel and De Graaf, 2005; Brown, 1993).

Step three: The P-set

The P-set is the group of participants in the study. Participants are chosen to be knowledgeable on the issue (Van Texel and De Graaf, 2005). The amount of people and its representativeness are depicted as not very important by Q methodology literature, as the Q methodology would reportedly work very well under most circumstances (McKeown and Thomas, 2013; Brown, 1993; Van Texel and De Graaf, 2005). However, what is noted is that *"a conscious effort is made to ensure as much variability in the composition of the P-set as is practicable under the circumstances"* (McKeown and Thomas, 2013). For this thesis, it is deemed necessary to have a sample of fifteen or more and have different kinds of people within it.

Step four: Q-sorting

With Q-sorting, the participants of the P-set are asked to rank the statements of the Q-set. This happens through individual face-to-face sessions with the participants, but can also be done through mail. For this study, it is done in individual face-to-face sessions. The sessions are prepared by printing the statements of the Q-set as randomly numbered cards and preparing a score sheet with a distribution in which participants will have to place the cards. The sheet ranges from 'least agreed' to 'most agreed'. The breadth of the scale varies per study and often depends on the amount of statements in the Q-set. Scales can range from -3 to +3, from -4 to +4 or from -5 to +5. The distribution on the score sheet of how many statements can be placed

on the respective rankings is generally a flat normal distribution when participants are expected to be knowledgeable on the subject. Participants are asked to sort the statements accordingly (Van Texel and De Graaf, 2005; Brown, 1993).

The Q-sorting itself involves several steps. First, the participant has to read all the statements. Next, the participant is asked to sort the provided statements in three piles: statements they agree with, statements they don't agree with and statements they're neutral towards. Third, the participant is instructed to sort the statements in the score sheet. The last step involves the researcher asking the participant to elaborate on the more extremely sorted statements (Van Texel and De Graaf, 2005). In addition to the last step, the participants will also be asked about three background variables, namely the department they work in, how long they have worked at the Safety Board, and what work and study they have done in the past.

Step five: Factor-analysis

After data has been gathered in Q-sorting, the results are analyzed by using factor analysis. The factor analysis shows the similarity and dissimilarity of the participants' opinions by showing how well participants fit with different factors and loading the statements into these factors (Van Texel and De Graaf, 2005; Brown, 1993). The P-set's preferences and choices, displayed in their Q-sorted statement distribution, reveal clusters of tastes and preferences.

Luckily, most of the mechanisms of factor analysis are performed by software, which makes it relatively simple. The rotation of factors is determined first. For the Q methodology a varimax rotation is used. Second, after factorial rotation has been done, it is determined how many factors there are by looking at the Eigenvalues and more. Having an Eigenvalue higher than 1.00 determines that a factor is significant, but there are other methods to determine significance as well (Van Texel and De Graaf, 2005; McKeown and Thomas, 2013).

Lastly, the factors from the factor analysis are interpreted by the researcher. A participant's factor loading is called a defining variate if the factor loading is significant, usually meaning that the p-value is smaller than 0.01. This means that a participant fits with other participants within one specific factor (Van Texel and De Graaf, 2005). If participants who load on a factor have a strongly different opinion on a statement than participants of other factors, the corresponding statements to that factor are called distinguishing statement. These illustrate the differences between the groups of opinions, helping in the interpretation of the separate factors. When participants that load on different factors generally agree on a statement, it is called a consensus statement. Distinguishing and consensus statements help in understanding what is similar and different between extracted factors. Explanations gained from interviews

about the more extreme statements also help make more sense of what group of opinions each factor represents (Van Texel and De Graaf, 2005; Coogan and Herrington, 2011).

3.2 Gathering data at the Dutch Safety Board

For the analysis of researchers' bias views in practice, the Dutch Safety Board has accepted the request to participate in this research. Before this, I have spent three months on an internship at the Dutch Safety Board. During this internship, I have gotten to know investigators from different fields. The advantages of this are that it will be easier to contact investigators at the Safety Board and gather participants for the study. The disadvantage is that my time spent on my internship is likely to have given me some bias of my own. More details about the Q-sorting done at the Dutch Safety Board are discussed in the next chapter.

3.3 Validity and reliability

This section briefly addresses the roles of validity and reliability in this thesis and to what extent these are accounted for.

Validity

Validity measures whether research measures what it really wants to measure or not and whether the methods of measurement are accurate (Golafshani, 2003). There are two forms of validity, internal and external validity. Internal validity means that the results of the study are directly related to the independent variables of the study and are not caused by other variables. External validity concerns whether the results can be generalized to the broader population or not, outside of the investigated group, and is sustainable over other times and settings as well (Onwuegbuzie, 2000).

The internal validity is taken care of by the method used, as the Q methodology does not involve the effect of one variable on another, but the relations between participants and items and how these can be interpreted. The Q methodology also accounts for external validity, as the relations between items are valid even with a single participant, let alone twenty participants from different departments. The broader population for this would be researchers in general and the sample of investigators should cover that group very well.

Next to this, triangulation and combining methods are ways to further improve validity (Golafshani, 2003). This is also something done in this thesis, by using multiple sources and both qualitative and quantitative methods. Qualitative methods are found in the theoretical framework and the interpretation of the results. Quantitative methods are found in the preparation, execution and results of the Q methodology analysis.

Reliability

Reliability means that it is possible to replicate and repeat research (Golafshani, 2003). This thesis is written in a way to enable replication. Everything is described in full detail, for example in the next two chapters. Chapter 4 describes in full detail how the Q methodology has been applied at the Dutch Safety Board. Chapter 5 is also very thorough in describing how the results are interpreted. Throughout the study, there are also some remarks on what would be good to keep in mind for replication, for example by addressing problems with the statements in the discussion.

3.4 Data and methods summary

This chapter has explained that the Q methodology involves five steps. These are defining the concourse, assembling the Q-set, choosing the P-set, Q-sorting and factor analysis. In the chapter it was also discussed how and why the Q methodology is used to gather data at the Dutch Safety Board, and the validity and reliability of this study. The next chapter describes the research process in more detail and how it was performed in practice. It shows the research process, the choices that were made and why these choices were made.

4. Analysis: the application of Q at the Dutch Safety Board

This chapter describes how the Q methodology is applied to the case study at the Dutch Safety Board. It discusses both the choices that are made and why they are made. The structure follows the five steps of the Q methodology, namely the concourse, the Q-set, the P-set, Q-sorting and factor analysis. The next chapter discusses the results of the factor analysis and their interpretation, aided by participants' explanations.

4.1 Defining the concourse

Two sources are used to collect information for the concourse. The main source is the theoretical framework. This study seeks to explore bias and does not seek to confirm existing theoretical ideas, but the theoretical framework is used as a basis for the concourse. In order to get a representative spread of opinions for the Q-set, the statements are collected evenly across the theoretical types of bias. The types of bias are issue bias, technical bias, methods bias and publication bias. The three relevant questions are factual statements about the existence of the bias type, normative statements about the desirability of the bias type, and prescriptive statements about possible solutions for the bias types. The representativeness of statements is constructed by determining important dimensions for all three questions on each bias type. This is the first step of creating the Q-set.

The secondary input for defining the concourse is the time I have personally spent at the Dutch Safety Board during my internship and the conversations and interviews there. As most of this is confidential, this experience is only used as reference to frame the statements in such a way that they make sense to the investigators at the Dutch Safety Board.

4.2 The Q-set

To construct the Q-set, different dimensions for all three questions on each bias type are presented first. Statements are created per dimension and explanations are given how a statement is based on its dimension. This is presented in both Dutch and English in table 3a through 3d, each discussing the dimensions and questions of bias types. While the Dutch version of the statements is presented to participants during Q-sorting, the English version serves as translation for this study. It is important to note that the translation to English means that there will be some differences in nuance and meaning. For replication purposes, the statements are most accurate to this study in Dutch. Next, the Q-set is displayed in table 4 with randomly shuffled and numbered statements in both Dutch and English. Lastly, the statements for the Q-set are summarized in English in table 5, including the numbers of the shuffled Q-set.

Table 3a. Dimensions, statements and explanations for issue bias

Statements	Dimensions	Statements (Dutch)	Statements (English)	Explanations
Factual	Creating evidence	Er zijn binnen de OVV specifieke voorkeuren voor welke onderwerpen onderzocht worden	Within the Dutch Safety Board, there are specific preferences for which subjects are researched	This statement implies that issue bias exists in choosing the topic evidence will be created for
	Selecting evidence	Er zijn binnen de OVV specifieke voorkeuren voor wat voor type data wel en niet als bewijs gebruikt wordt	Within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence	This statement implies that issue bias exists in choosing what evidence is and is not looked at
	Interpreting evidence	Er zijn binnen de OVV specifieke voorkeuren voor welk type resultaten belicht wordt in onderzoek	Within the Dutch Safety Board, there are specific preferences for what type of results are highlighted in research	This implies that there is issue bias in choosing which results are deemed more important than others
Normative	Acting against bias	Het is belangrijk op te treden tegen voorkeuren voor een bepaald onderwerp	It is important to act against preferences for a certain subject	General statement to see if researchers feel the need to act against issue bias
	Issue bias as a bad thing	Het is verkeerd om je in een onderzoek slechts op een aantal onderwerpen te richten	It is wrong to focus on only a couple of subjects within a research project	This statement implies that issue bias is a bad thing
	Responsible bias	Het is beter om de tijd te nemen de onderwerpkeuze van een onderzoek ter discussie te stellen dan om er niet lang bij stil te staan	It is better to take the time to discuss the subject choice of research than to not consider it for long	This statement tests how researchers feel about using bias as a tool instead of seeing it as purely a threat
Prescriptive (good governance of evidence)	Appropriate-ness	Voorkeuren bij het kiezen van onderwerpen en soorten bewijs kunnen bestreden worden door voor elk onderzoek los te beoordelen welke onderwerpen en soorten bewijs daarvoor geschikt zijn	Preferences in choosing subjects and kinds of evidence can be tackled by determining which subjects and kinds of evidence are appropriate for every separate research project	This statement tests how researchers feel about appropriateness as solution for issue bias
	Stewardship and representation	Voorkeuren bij het kiezen van onderwerpen en soorten bewijs kunnen bestreden worden als een gezaghebbend persoon regels maakt hiervoor	Preferences in choosing subjects and kinds of evidence can be tackled if an authoritative person creates rules for this	This statement tests how researchers feel about stewardship and representation as solutions for issue bias
	Deliberation	Tegenlezen en openstaan voor kritiek kan helpen om voorkeuren bij het kiezen van onderwerpen en soorten bewijs te bestrijden	Revision and being open for criticism can help to tackle preferences of choosing subjects and kinds of evidence	This statement tests how researchers feel about deliberation as solution for issue bias

Sources: theoretical framework and personal experience at the Dutch Safety Board

Table 3b. Dimensions, statements and explanations for technical bias

Statements	Dimensions	Statements (Dutch)	Statements (English)	Explanations
Factual	Selecting evidence	Bewijs wordt binnen de OVV soms geselecteerd vanuit eigen ideeën over de bruikbaarheid ervan	Within the Dutch Safety Board, evidence is sometimes selected based on own ideas on the usability of it	This statement tests if there is technical bias in the selection of evidence
	Interpreting evidence	Conclusies in onderzoek van de OVV sluiten niet altijd aan op de bevindingen en resultaten van het onderzoek	Conclusions in the Dutch Safety Board's research do not always connect to the findings and results of research	This statements tests if there is technical bias in the interpretation of evidence
	Selection bias	Bij onderzoek wordt er binnen de OVV soms geen rekening gehouden met de representativiteit van de resultaten tegenover de onderzochte groep	With some research of the Dutch Safety Board, there is no consideration for the representativity of results towards the researched group	This statement tests if there is technical bias in the selection of samples for research
Normative	Acting against bias	Het is belangrijk hard op te treden tegen slecht uitgevoerd onderzoek	It is important to act against poorly executed research	General statement to see if researchers feel the need to act against technical bias
	Technical bias as a bad thing	Het is verkeerd om geen rekening te houden met de kwaliteit van bewijs bij het doen van onderzoek	It is wrong to not consider the quality of evidence in research	This statement implies that technical bias is a bad thing
	Pearls of wisdom	Slecht uitgevoerd onderzoek bevat niets van waarde	Poorly executed research contains nothing of value	This statement test if respondents feel there are any pearls of wisdom in technically biased research
Prescriptive (good governance of evidence)	Rigor	Alle beschikbare bewijsstukken moeten verzameld worden om slecht onderzoek tegen te gaan	All available pieces of evidence must be collected to overcome poor research	This statement tests if the researchers feel that rigorously collecting evidence is a solution for technical bias
	Quality	Om slecht onderzoek tegen te gaan moet elk soort onderzoek eigen criteria hebben	To do something against poor research, every kind of research needs to have its own criteria	This statement tests if the researchers feel that specified criteria for quality are a solution for technical bias
	Transparency and contestability	Om slecht onderzoek tegen te gaan moet het onderzoek inzichtelijk en open voor kritiek zijn voor de buitenwereld	To do something against poor research, research has to be insightful and open to criticism from the outside world	This statement tests if the researchers feel that transparent and contestable research is a solution for technical bias

Sources: theoretical framework and personal experience at the Dutch Safety Board

Table 3c. Dimensions, statements and explanations for methods bias

Statements	Dimensions	Statements (Dutch)	Statements (English)	Explanations
Factual	Transparency and contestability	Onderzoek binnen de OVV is niet inzichtelijk of open voor kritiek voor de buitenwereld	Research within the Dutch Safety Board is not insightful or open to criticism for the outside world	This statement is applicable to both technical and methods bias. It test to what degree transparency and contestability are already present. If not, this can contribute to the creation of more bias
	Common rater effects	Bevindingen in het onderzoek van de OVV worden door sociale wenselijkheid beïnvloed	Findings in the Dutch Safety Board's research are influenced by social desirability	This statement tests if social desirability is present, as this would indicate common rater effects
	Measurement context effects	De sociaal-maatschappelijke context rond een onderzoek heeft onbedoelde invloed op de bevindingen en resultaten van de OVV	The societal context surrounding research has unintended influence on the findings and results of the Dutch Safety Board	This statement test if measurement context effects are present
Normative	Acting against bias	Het is belangrijk op te treden tegen kleine meetfouten	It is important to act against small measurement errors	General statement to see if researchers feel the need to act against methods bias
	Methods bias as a bad thing	Het is verkeerd om in onderzoeken niet op juistheid van metingen te letten	It is wrong to not pay attention to the correctness of measurements in research	This statement implies that methods bias is a bad thing
	Good enough methods	Het is beter tijd te steken in het vermijden van meetfouten dan tijd te besparen door dit in mindere mate te doen	It is better to put time in avoiding measurement errors than save time by doing it to a lesser degree	This statement tests the attitude about methods bias when compared to the time solving them costs
Prescriptive	Triangulation	Het gebruiken van meerdere bronnen kan meetfouten verhelpen	The usage of multiple sources can overcome measurement errors	This statement test if triangulation is seen as a solution for methods bias
	Distance	Door metingen los van elkaar te doen kunnen deze niet onbedoeld beïnvloed worden door hun omgeving	By doing measurements separately, they cannot be unintentionally influenced by their surroundings	This statement tests if creating distance between measures is seen as solution for methods bias
	Good governance of evidence	Door het onderzoeksproces open te stellen voor kritiek kunnen meetfouten verholpen worden	By exposing the research process to criticism, measurement errors can be overcome	This statement tests if the good governance of evidence is also seen as solution for methods bias

Sources: theoretical framework and personal experience at the Dutch Safety Board

Table 3d. Dimensions, statements and explanations for publication bias

Statements	Dimensions	Statements (Dutch)	Statements (English)	Explanations
Factual	General publication bias	Er zijn binnen de OVV specifieke voorkeuren om iets niet te publiceren als er niets nieuws gevonden wordt	In the Dutch Safety Board, there are specific preferences to not publish something if nothing new is found	This statement tests if publication bias happens or not
	Media	Onderzoeksprocessen binnen de OVV worden beïnvloed door de publieke aandacht die ze krijgen in diverse media, zoals nieuwssites, kranten en televisieprogramma's	Research processes within the Dutch Safety Board are influenced by the public attention they get in a diversity of media, like news sites, papers and television programs	This statement tests if news media and influence research processes
	Policy advice	Binnen de OVV wordt onderzoek beïnvloed doordat het relevant en begrijpelijk moet zijn voor beleidsmakers	Within the Dutch Safety Board, research is influenced for having to be relevant and understandable for policymakers	This statement tests if research processes are influenced to be comprehensibly published for policy makers
Normative	Acting against bias	Het is belangrijk op te treden tegen voorkeuren van wat wel en niet gepubliceerd wordt	It is important to act against preferences of what does and does not get published	General statement to see if researchers feel the need to act against publication bias
	Publication bias as a bad thing	Het is verkeerd om een voorkeur te hebben voor wat er wel en niet gepubliceerd wordt	It is wrong to have preferences for what does and does not get published	This statement implies that publication bias is a bad thing
	Filtering studies with low power	Het is geen goede zaak dat groter onderzoek voorrang krijgt ten opzichte van kleiner onderzoek	It is not a good thing that larger research is prioritized over smaller research	This statement tests how researchers feel about preferring studies with high power over those with low
Prescriptive	Comparing results	Het vergelijken van de resultaten in onderzoeksrapporten zou helpen om te ontdekken of er voorkeuren zijn voor het publiceren van onderzoeken met sterkere resultaten	Comparing results in research reports would help to discover if there are preferences for publishing results with stronger results	This statement tests if researchers feel comparing results will help solve publication bias
	Registries	Voorkeuren voor wat wel en niet gepubliceerd wordt kunnen tegen worden gegaan door een los en publiekelijk beschikbaar register, waarin onderzoeksprocessen in detail beschreven worden	Preferences for what does and does not get published can be countered by a separate and publicly available registry, within which research processes are described in detail	This statement tests if researchers feel that registries help solve publication bias
	Policy	Met beleid en regelgeving voor publiceren kunnen voorkeuren van wat wel en niet gepubliceerd wordt tegengegaan worden	With policy and rules for publishing, preferences of what does and does not get published can be overcome	This statement tests if researchers feel that policy can help solve publication bias

Sources: theoretical framework and personal experience at the Dutch Safety Board

Table 4. Shuffled and numbered Q-set in Dutch and English (continued on next page)

Q-set statement number	Language	
	Dutch	English
1	Het is beter om de tijd te nemen de onderwerpkeuze van een onderzoek ter discussie te stellen dan om er niet lang bij stil te staan	It is better to take the time to discuss the subject choice of research than to not consider it for long
2	Bij onderzoek wordt er binnen de OVV soms geen rekening gehouden met de representativiteit van de resultaten tegenover de onderzochte groep	With some research of the Dutch Safety Board, there is no consideration for the representativity of results towards the researched group
3	Er zijn binnen de OVV specifieke voorkeuren om iets niet te publiceren als er niets nieuws gevonden wordt	In the Dutch Safety Board, there are specific preferences to not publish something if nothing new is found
4	Het is verkeerd om in onderzoeken niet op juistheid van metingen te letten	It is wrong to not pay attention to the correctness of measurements in research
5	Tegenlezen en openstaan voor kritiek kan helpen om voorkeuren bij het kiezen van onderwerpen en soorten bewijs te bestrijden	Revision and being open for criticism can help to tackle preferences of choosing subjects and kinds of evidence
6	Het is belangrijk op te treden tegen voorkeuren van wat wel en niet gepubliceerd wordt	It is important to act against preferences of what does and does not get published
7	Onderzoeksprocessen binnen de OVV worden beïnvloed door de publieke aandacht die ze krijgen in diverse media, zoals nieuwssites, kranten en televisieprogramma's	Research processes within the Dutch Safety Board are influenced by the public attention they get in a diversity of media, like news sites, papers and television programs
8	Om slecht onderzoek tegen te gaan moet elk soort onderzoek eigen criteria hebben	To do something against poor research, every kind of research needs to have its own criteria
9	Het is belangrijk op te treden tegen kleine meetfouten	It is important to act against small measurement errors
10	Door het onderzoeksproces open te stellen voor kritiek kunnen meetfouten verholpen worden	By exposing the research process to criticism, measurement errors can be overcome
11	Het gebruiken van meerdere bronnen kan meetfouten verhelpen	The usage of multiple sources can overcome measurement errors
12	Door metingen los van elkaar te doen kunnen deze niet onbedoeld beïnvloed worden door hun omgeving	By doing measurements separately, they cannot be unintentionally influenced by their surroundings
13	Het is geen goede zaak dat groter onderzoek voorrang krijgt ten opzichte van kleiner onderzoek	It is not a good thing that larger research is prioritized over smaller research
14	De sociaal-maatschappelijke context rond een onderzoek heeft onbedoelde invloed op de bevindingen en resultaten van de OVV	The societal context surrounding research has unintended influence on the findings and results of the Dutch Safety Board
15	Binnen de OVV wordt onderzoek beïnvloed doordat het relevant en begrijpelijk moet zijn voor beleidsmakers	Within the Dutch Safety Board, research is influenced for having to be relevant and understandable for policymakers
16	Er zijn binnen de OVV specifieke voorkeuren voor welke onderwerpen onderzocht worden	Within the Dutch Safety Board, there are specific preferences for which subjects are researched
17	Het is verkeerd om geen rekening te houden met de kwaliteit van bewijs bij het doen van onderzoek	It is wrong to not consider the quality of evidence in research
18	Om slecht onderzoek tegen te gaan moet het onderzoek inzichtelijk en open voor kritiek zijn voor de buitenwereld	To do something against poor research, research has to be insightful and open to criticism from the outside world

Source: statements shuffled by using https://www.random.org/lists/

Table 4 (continued).

Q-set statement number	Language	
	Dutch	English
19	Met beleid en regelgeving voor publiceren kunnen voorkeuren van wat wel en niet gepubliceerd wordt tegengegaan worden	With policy and rules for publishing, preferences of what does and does not get published can be overcome
20	Onderzoek binnen de OVV is niet inzichtelijk of open voor kritiek voor de buitenwereld	Research within the Dutch Safety Board is not insightful or open to criticism for the outside world
21	Conclusies in onderzoek van de OVV sluiten niet altijd aan op de bevindingen en resultaten van het onderzoek	Conclusions in the Dutch Safety Board's research do not always connect to the findings and results of research
22	Bewijs wordt binnen de OVV soms geselecteerd vanuit eigen ideeën over de bruikbaarheid ervan	Within the Dutch Safety Board, evidence is sometimes selected based on own ideas on the usability of it
23	Het is beter tijd te steken in het vermijden van meetfouten dan tijd te besparen door dit in mindere mate te doen	It is better to put time in avoiding measurement errors than save time by doing it to a lesser degree
24	Het is belangrijk op te treden tegen voorkeuren voor een bepaald onderwerp	It is important to act against preferences for a certain subject
25	Slecht uitgevoerd onderzoek bevat niets van waarde	Poorly executed research contains nothing of value
26	Voorkeuren voor wat wel en niet gepubliceerd wordt kunnen tegen worden gegaan door een los en publiekelijk beschikbaar register, waarin onderzoeksprocessen in detail beschreven worden	Preferences for what does and does not get published can be countered by a separate and publicly available registry, within which research processes are described in detail
27	Het is verkeerd om een voorkeur te hebben voor wat er wel en niet gepubliceerd wordt	It is wrong to have preferences for what does and does not get published
28	Voorkeuren bij het kiezen van onderwerpen en soorten bewijs kunnen bestreden worden door voor elk onderzoek los te beoordelen welke onderwerpen en soorten bewijs daarvoor geschikt zijn	Preferences in choosing subjects and kinds of evidence can be tackled by determining which subjects and kinds of evidence are appropriate for every separate research project
29	Er zijn binnen de OVV specifieke voorkeuren voor welk type resultaten belicht wordt in onderzoek	Within the Dutch Safety Board, there are specific preferences for what type of results are highlighted in research
30	Het is belangrijk hard op te treden tegen slecht uitgevoerd onderzoek	It is important to act against poorly executed research
31	Voorkeuren bij het kiezen van onderwerpen en soorten bewijs kunnen bestreden worden als een gezaghebbend persoon regels maakt hiervoor	Preferences in choosing subjects and kinds of evidence can be tackled if an authoritative person creates rules for this
32	Het vergelijken van de resultaten in onderzoeksrapporten zou helpen om te ontdekken of er voorkeuren zijn voor het publiceren van onderzoeken met sterkere resultaten	Comparing results in research reports would help to discover if there are preferences for publishing results with stronger results
33	Bevindingen in het onderzoek van de OVV worden door sociale wenselijkheid beïnvloed	Findings in the Dutch Safety Board's research are influenced by social desirability
34	Alle beschikbare bewijsstukken moeten verzameld worden om slecht onderzoek tegen te gaan	All available pieces of evidence must be collected to overcome poor research
35	Het is verkeerd om je in een onderzoek slechts op een aantal onderwerpen te richten	It is wrong to focus on only a couple of subjects within a research project
36	Er zijn binnen de OVV specifieke voorkeuren voor wat voor type data wel en niet als bewijs gebruikt wordt	Within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence

Source: statements shuffled by using https://www.random.org/lists/

Table 5. Categorized statements in English, including shuffled Q-set numbers

Statements	Issue bias	Technical bias	Methods bias	Publication bias
Factual	16. Within the Dutch Safety Board, there are specific preferences for which subjects are researched	22. Within the Dutch Safety Board, evidence is sometimes selected based on own ideas on the usability of it	20. Research within the Dutch Safety Board is not insightful or open to criticism for the outside world	3. In the Dutch Safety Board, there are specific preferences to not publish something if nothing new is found
	36. Within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence	21. Conclusions in the Dutch Safety Board's research do not always connect to the findings and results of research	33. Findings in the Dutch Safety Board's research are influenced by social desirability	7. Research processes within the Dutch Safety Board are influenced by the public attention they get in a diversity of media, like news sites, papers and television programs
	29. Within the Dutch Safety Board, there are specific preferences for what type of results are highlighted in research	2. With some research of the Dutch Safety Board, there is no consideration for the representativity of results towards the researched group	14. The societal context surrounding research has unintended influence on the findings and results of the Dutch Safety Board	15. Within the Dutch Safety Board, research is influenced for having to be relevant and understandable for policymakers
Normative	24. It is important to act against preferences for a certain subject	30. It is important to act against poorly executed research	9, It is important to act against small measurement errors	6. It is important to act against preferences of what does and does not get published
	35. It is wrong to focus on only a couple of subjects within a research project	17. It is wrong to not consider the quality of evidence in research	4. It is wrong to not pay attention to the correctness of measurements in research	27. It is wrong to have preferences for what does and does not get published
	1. It is better to take the time to discuss the subject choice of research than to not consider it for long	25. Poorly executed research contains nothing of value	23. It is better to put time in avoiding measurement errors than save time by doing it to a lesser degree	13. It is not a good thing that larger research is prioritized over smaller research
Prescriptive	28. Preferences in choosing subjects and kinds of evidence can be tackled by determining which subjects and kinds of evidence are appropriate for every separate research project	34. All available pieces of evidence must be collected to overcome poor research	11. The usage of multiple sources can overcome measurement errors	32. Comparing results in research reports would help to discover if there are preferences for publishing results with stronger results
	31. Preferences in choosing subjects and kinds of evidence can be tackled if a authoritative person creates rules for this	8. To do something against poor research, every kind of research needs to have its own criteria	12. By doing measurements separately, they cannot be unintentionally influenced by their surroundings	26. Preferences for what does and does not get published can be countered by a separate and publicly available registry, within which research processes are described in detail
	5. Revision and being open for criticism can help to tackle preferences of choosing subjects and kinds of evidence	18. To do something against poor research, research has to be insightful and open to criticism from the outside world	10. By exposing the research process to criticism, measurement errors can be overcome	19. With policy and rules for publishing, preferences of what does and does not get published can be overcome

Source: author's summary

4.3 The P-set

The P-set consists of 20 participants at the Dutch Safety Board. The P-set is selected out of 10 different sectors of the Safety Board and includes several investigators that are not tied to a cluster. Backgrounds of the participants are categorized very broadly in order to ensure anonymity. The categories and how many participants are tied to each are presented in table 6. Participants can belong to both a kind of background and a category for how long they have been employed.

Table 6. P-set backgrounds (categorized)

Background category		Number of participants in category
Study/work background	Technical background	11
	Social background	9
Years of employment	0-5 years employed	9
	5-10 years employed	5
	10+ years employed	6

Source: author

4.4 Q-sorting

The Q-sorting process was done within the building of the Dutch Safety Board. Participants were individually taken to a separate room with soundproofed walls, in which they could speak freely without worries of colleagues listening in. Participants received an explanation of what they are participating in and got the opportunity to ask questions before starting. During a session, participants were free to ask questions and make any remarks. Most remarks and questions were noted. However, limited and consistent answers were given for the sake of not guiding participants' answers.

In the first step of a session, participants were asked to tell about their education and work experiences, what sector they work in at the Dutch Safety Board and how long they have worked at the Dutch Safety Board.

Second, the actual Q-sorting took place. Participants were asked to go through statements, written on cards, individually and distribute them in piles of 'agree with statement', 'neither agree nor disagree with statement' and 'disagree with statement'. Next, participants were presented with a scoring sheet. The score sheet can be seen in Figure 1. It is in Dutch, as it was presented to the participants. The sheet ranged from 'least agreed' with the number 1 to

'most agreed' with the number 9. Participants were asked to distribute the cards in the 'agree' pile first, then the ones of the 'disagree' pile, ending with the 'neither agree nor disagree' pile. Participants were asked to explain why they placed the statements on categories 1, 2, 8 and 9. Participants were thanked for their participation and got the chance to make some more remarks and ask questions. Questions were answered to a limited degree, in order to make sure the rest of the sessions could not be influenced by participants talking about it with their colleagues. After this final opportunity for questions and remarks, the session was ended.

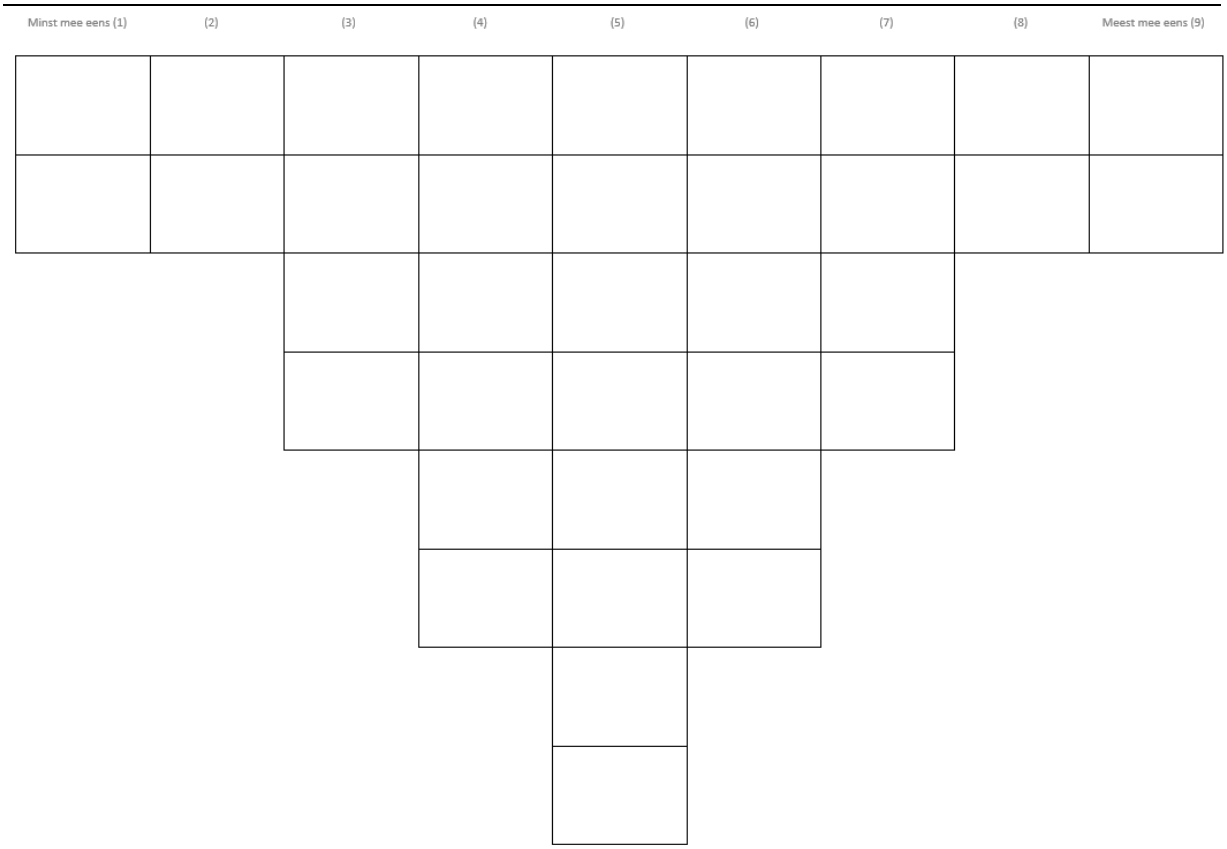

Figure 1. Q-sorting score sheet (in Dutch)

4.5 Factor analysis

For doing factor analysis for the Q methodology, the software PQMethod was used (Schmolck, 2014). Following the manual on the software's website, three steps were taken and choices were made for the factor analysis. These choices and steps are discussed in this section.

Step 1: getting started with PQMethod

The first step involves inserting the statements and raw data into PQMethod. The first step requires the user to enter all statements in the same order as during the Q-sorting sessions. The only limitation is that you can only use 60 characters per statement (Schmolck, 2014).

Second, more details about the statements and the Q-sorting session are entered into PQMethod. In this study, this indicated that there are a total of 36 statements, with the score sheet ranging from -4 to 4 and the distribution of rows for each column being 2-2-4-6-8-6-4-2-2. At this point, the distributions of all participants have been inserted into the program (Schmolck, 2014).

Step 2: unrotated factor analysis

The second step involves the first important choices in using PQMethod. For the execution of an unrotated factor analysis, the choice had to be made between a Brown Centroid factor analysis, a Horst Centroid factor analysis and PCA (Principal Components Analysis). Each choice offers something different. It helps determine, for example, how many factors are extracted for the rotated analysis. Brown Centroid factor analysis always extracts seven factors, the Horst Centroid factor analysis lets the user choose how many factors are extracted and PCA extracts the program's maximum of eight factors. For this study, PCA was chosen, because it also calculates the Eigenvalues, which is helpful for knowing how many factors to extract (Schmolck, 2014). What the statistically most correct choice would be remains not completely certain, as there is still much discussion surrounding this topic (McKeown and Thomas, 2013).

The reported Eigenvalues for this study are displayed in table 7. The table also includes the explained variance of each component in percentages and cumulative percentages. A self-made scree-plot, using the data of the Eigenvalues, is presented in figure 2. PQMethod itself does not provide users with the option of making scree-plots, but when doing a factor analysis, a scree-plot can help determine the amount of factors to extract (Williams et al., 2010).

Table 7. Eigenvalues and explained variance (percentages and cumulative percentages)

Component	Eigenvalues (%, cumul. %)	Component	Eigenvalues (%, cumul. %)	Component	Eigenvalues (%, cumul. %)	Component	Eigenvalues (%, cumul. %)
1	8.00 (40.0, 40.0)	6	0.91 (4.5, 76.3)	11	0.47 (2.3, 91.6)	16	0.13 (0.7, 98.2)
2	2.21 (11.0, 51.0)	7	0.79 (4.0, 80.3)	12	0.35 (1.7, 93.4)	17	0.13 (0.6, 98.8)
3	1.64 (8.2, 59.2)	8	0.71 (3.5, 83.8)	13	0.33 (1.6, 95.0)	18	0.12 (0.6, 99.4)
4	1.33 (6.7, 65.9)	9	0.58 (2.9, 86.7)	14	0.30 (1.5, 96.5)	19	0.08 (0.4, 99.8)
5	1.18 (5.9, 71.8)	10	0.51 (2.6, 89.3)	15	0.20 (1.0, 97.5)	20	0.05 (0.2, 100.0)

Source: PQMethod PCA output

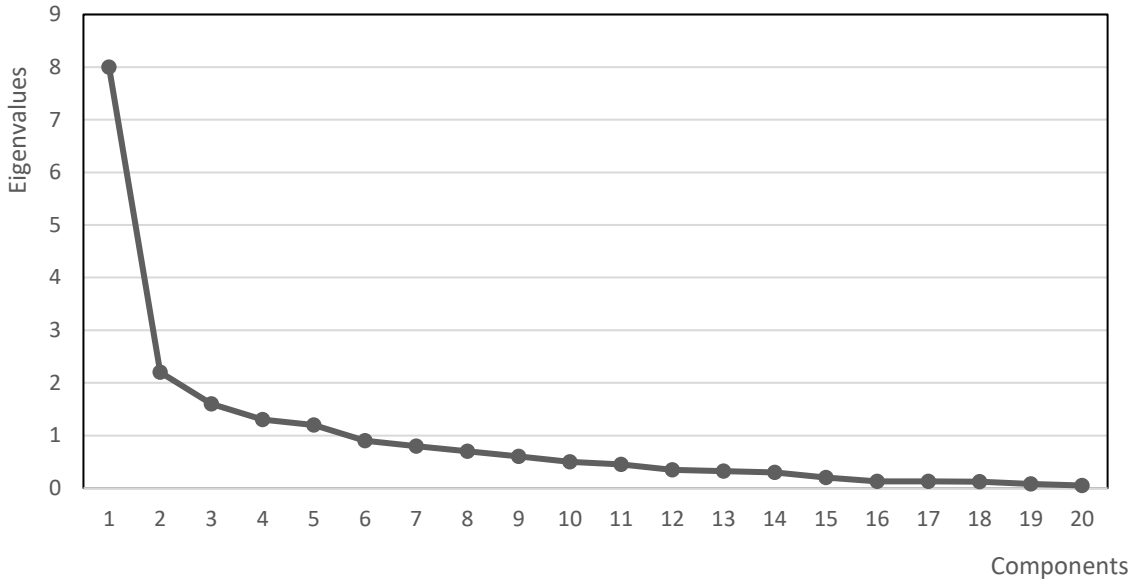

Figure 2. Scree-plot of components (x-axis) and Eigenvalues (y-axis).

Step 3: rotated factor analysis

In the last step, a rotated factor analysis is executed by the program. For this, there are two options. The user can either rotate the factors themselves or let the program do it with a varimax rotation (Schmolck, 2014). For the purposes of this Q study, a varimax rotation is used, as this is most often used with this method (McKeown and Thomas, 2013). PQMethod requires the researcher to choose how many factors should be rotated (Schmolck, 2014).

In order to choose how many factors to extract, there are many different criteria to look at. The available criteria for this study are Eigenvalues, a scree test, cumulative percentages and theory. A combination of all of these criteria are used, as this will strengthen the decision for the amount of factors to extract (Williams et al., 2010). The Eigenvalues, as shown in table 7, indicate that five factors can be extracted. When looking at the scree-plot in figure 2, the same conclusion can be drawn, as the 'elbow' is right above the sixth component. The cumulative percentage is around 60 percent with three or four components. Finally, theory suggests that four factors should be extracted, as there are four kinds of bias found within literature.

The final choice rests in a combination of the four criteria. The Eigenvalue criterium is known for extracting too many factors (Costello and Osborne, 2005). The criterium with the least amount of factors is the theory criterium, suggesting to extract four instead of five factors. Therefore, the choice is made to extract four factors.

The second part of executing the rotated factor analysis is flagging defining subjects for each extracted factor. There are two options for this. Items can be auto-flagged by the program and items can be flagged by the researcher. Auto-flagging occurs by selecting pure cases that only

load significantly on a single factor (Schmolck, 2014; McKeown and Thomas, 2013). The auto-flag option is used in this study, as this flags the P-set in exactly such a way. The resulting distribution of factors and subjects is shown in table 8. Five participants are grouped as defining for factor 1, two for factor 2, five for factor 3 and three for factor 5. At the end of the next chapter, the backgrounds of various participants are examined to see if they suggest that someone's background can predict what kind of group they belong to in the case of bias.

Table 8. Flagged defining items for the Varimax rotated factor analysis

P-set	Participant loading per factor				P-set	Participant loading per factor			
	1	2	3	4		1	2	3	4
1	0.809*	-0.132	0.280	0.112	11	0.640*	0.398	0.307	0.211
2	0.853*	0.238	0.049	0.154	12	0.711*	0.455	-0.041	0.022
3	0.385	0.280	0.668*	0.190	13	-0.112	0.058	0.822*	0.280
4	-0.270	0.287	0.028	0.780*	14	0.122	0.030	0.100	0.833*
5	0.170	0.684*	0.439	0.117	15	0.275	0.133	0.402	0.481
6	0.332	0.142	0.497	0.580	16	0.724*	-0.009	0.419	-0.054
7	0.295	0.652*	0.124	0.246	17	0.287	0.103	0.440	0.173
8	-0.004	0.860	0.103	0.143	18	0.186	0.113	0.241	0.553*
9	0.212	0.296	0.714*	0.006	19	0.410	0.303	0.493	0.186
10	0.435	0.408	0.634*	0.038	20	0.091	-0.030	0.745*	0.171

** flagged as a defining item for the factor*
Source: PQMethod analysis results

4.6 Analysis summary

This chapter showed how the Q methodology has been applied to the Dutch Safety Board in full detail. In doing so, it described five steps. Step one and two of setting up the concourse and Q-set are described first, including tables of the statements and Q-set for the analysis. Step three describes the P-set, which is the group of participants. Step four describes how the Q-sorting sessions were performed in practice at the Dutch Safety Board. The final step describes the analysis of the data of the Q-sorting, using the program PQMethod. The next chapter shows the results of the analysis and interprets the results.

5. Results and interpretation of Q's factor analysis

Three aspects of the results of the Q methodology's factor analysis are highlighted and interpreted in this chapter. The first and second aspects are the factor distribution of the statements and the distinguishing statements. These are discussed first in order to interpret factors. Third, consensus statements are discussed, looking at which statements most participants in the study agreed with and why. After the discussion of these three aspects, participant's backgrounds and some of the other factors discussed in the theoretical framework are examined to see if these can explain participants' views.

Important to remember with all these statements is that they do not represent the Dutch Safety Board in a straightforward way. The reasons for agreeing and disagreeing with statements differs from person to person and diverges in various degrees from the theoretical meanings and explanations. This is explained further in the discussion section.

5.1 Factor exploration and interpretation

Several aspects of the factors are addressed in this section. First, the factors themselves are explored, which shows how the factors are constructed and what the distinguishing statements are. The factors are discussed on what they might represent. Distinguishing statements, the statements which a respective group placed on a significantly different spot of the score sheet than the other groups during Q-sorting, are discussed to better understand the different factors.

Factor exploration

The first results show the Z-scores for all statements and factors, as can be seen in table 9. The statements with a Z-score of 1.00 or higher and -1.00 and lower fit significantly within a factor. All significant statements are shown in table 10a. Ten statements load on factor 1, twelve on factor 2, ten on factor 3 and thirteen on factor 4. However, not all statements load significantly on only one factor. Several load on different factors, namely statements 1, 5, 10, 11, 17, 20, 23, 26, 27, 31, 34, and 35. Table 10b only allows these statements to load only on one factor, in which they have the highest positive or lowest negative Z-scores. Seven statements load on factor 1, two statements on factor 2, six statements on factor 3 and ten statements on factor 4. The statements that are not mentioned in the tables do not load significantly on any factor. The most relevant distinguishing statements for each factor are summed up in table 11. Relevancy is determined by whether or not the statements were placed in categories 1, 2, 3, 7, 8 and 9. Categories 4 through 6 do not help in learning more about opinion groups, as these were used by participants for statements they did not understand or were neutral towards.

Table 9. Statements and Z-scores for each factor

	Statements	Z-scores per factor			
		1	2	3	4
1.	It is better to take the time to discuss the subject choice of research than to not consider it for long	1.54*	1.42*	1.30*	0.35
2.	With some research of the Dutch Safety Board, there is no consideration for the representativity of results towards the researched group	-0.59	0.22	-0.05	-0.60
3.	In the Dutch Safety Board, there are specific preferences to not publish something if nothing new is found	-0.33	-0.67	-0.25	-1.67*
4.	It is wrong to not pay attention to the correctness of measurements in research	0.42	-0.16	1.35*	0.52
5.	Revision and being open for criticism can help to tackle preferences of choosing subjects and kinds of evidence	2.04*	0.94	0.28	1.19*
6.	It is important to act against preferences of what does and does not get published	0.46	-0.23	-0.66	-0.64
7.	Research processes within the Dutch Safety Board are influenced by the public attention they get in a diversity of media, like news sites, papers and television programs	1.71*	0.93	0.47	-0.09
8.	To do something against poor research, every kind of research needs to have its own criteria	-0.38	-0.80	-0.51	-0.30
9.	It is important to act against small measurement errors	-0.81	0.35	-0.37	0.22
10.	By exposing the research process to criticism, measurement errors can be overcome	-0.02	1.03*	0.68	1.07*
11.	The usage of multiple sources can overcome measurement errors	0.65	1.72*	0.90	1.58*
12.	By doing measurements separately, they cannot be unintentionally influenced by their surroundings	-0.46	-0.73	0.34	-0.00
13.	It is not a good thing that larger research is prioritized over smaller research	-0.85	1.22*	-1.46*	0.06
14.	The societal context surrounding research has unintended influence on the findings and results of the Dutch Safety Board	1.33*	0.02	-0.72	0.17
15.	Within the Dutch Safety Board, research is influenced for having to be relevant and understandable for policymakers	0.05	-0.19	0.96	-0.61
16.	Within the Dutch Safety Board, there are specific preferences for which subjects are researched	1.45*	0.00	0.88	0.30
17.	It is wrong to not consider the quality of evidence in research	1.26*	1.29*	1.46*	2.23*
18.	To do something against poor research, research has to be insightful and open to criticism from the outside world	0.01	-0.60	0.86	1.16*
19.	With policy and rules for publishing, preferences of what does and does not get published can be overcome	-0.80	-0.83	-0.76	0.30
20.	Research within the Dutch Safety Board is not insightful or open to criticism for the outside world	-0.30	1.44*	-0.56	-1.72*
21.	Conclusions in the Dutch Safety Board's research do not always connect to the findings and results of research	-0.17	-0.81	-0.00	-1.46*
22.	Within the Dutch Safety Board, evidence is sometimes selected based on own ideas on the usability of it	0.94	0.20	0.25	-1.08*
23.	It is better to put time in avoiding measurement errors than save time by doing it to a lesser degree	-0.07	0.82	1.06*	1.28*
24.	It is important to act against preferences for a certain subject	-0.41	-0.10	-0.42	-0.59
25.	Poorly executed research contains nothing of value	-0.04	-2.32*	0.99	-0.22
26.	Preferences for what does and does not get published can be countered by a separate and publicly available registry, within which research processes are described in detail	-1.10*	-1.28*	-1.82*	-0.82
27.	It is wrong to have preferences for what does and does not get published	-0.69	-1.28*	-2.20*	0.59
28.	Preferences in choosing subjects and kinds of evidence can be tackled by determining which subjects and kinds of evidence are appropriate for every separate research project	0.56	-0.89	-0.08	-0.94
29.	Within the Dutch Safety Board, there are specific preferences for what type of results are highlighted in research	0.97	0.00	-0.13	0.94
30.	It is important to act against poorly executed research	-0.56	0.66	1.56*	0.82
31.	Preferences in choosing subjects and kinds of evidence can be tackled if an authoritative person creates rules for this	-1.91*	-1.83*	-1.19*	-0.22
32.	Comparing results in research reports would help to discover if there are preferences for publishing results with stronger results	0.27	0.12	-0.87	0.04
33.	Findings in the Dutch Safety Board's research are influenced by social desirability	-0.07	0.02	-0.25	-2.01*
34.	All available pieces of evidence must be collected to overcome poor research	-2.08*	1.64*	0.99	1.08*
35.	It is wrong to focus on only a couple of subjects within a research project	-2.00*	-1.32*	-2.14*	0.39
36.	Within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence	-0.02	0.00	0.12	-1.34*

** Z-score of 1.00 and higher and -1.00 and lower.*

Source: PQMethod analysis results

Table 10a. Significant statements per factor, including doubles

Factor	Significant statements (Z-scores)*
1	5 (2.04), 7 (1.71), 1 (1.54), 16 (1.45), 14 (1.33), 17 (1.26), 26 (-1.10), 31 (-1.91), 35 (-2.00), 34 (-2.08)
2	11 (1.72), 34 (1.64), 20 (1.44), 1 (1.42) 17 (1.29), 13 (1.22), 10 (1.03), 26 (-1.28), 27 (-1.28), 35 (-1.32), 31 (-1.83), 25 (-2.32)
3	30 (1.56), 17 (1.46), 4 (1.35), 1 (1.30), 23 (1.06), 31 (-1.19), 13 (-1.46), 26 (-1.82), 35 (-2.14), 27 (-2.20)
4	17 (2.23), 11 (1.58), 23 (1.28), 5 (1.20), 18 (1.16), 34 (1.08), 10 (1.07), 22 (-1.08), 36 (-1.34), 21 (-1.46), 3 (-1.67), 20 (-1.72), 33 (-2.01)

**Distribution from highest to 1 and from -1 to the lowest*
Source: PQMethod analysis results

Table 10b. Significant statements per factor, excluding doubles

Factor	Significant statements (Z-scores)*
1	5 (2.04), 7 (1.71), 1 (1.54), 16 (1.45), 14 (1.33), 31 (-1.91), 34 (-2.08)
2	11 (1.72), 25 (-2.32)
3	30 (1.56), 4 (1.35), 13 (-1.46), 26 (-1.82), 35 (-2.14), 27 (-2.20)
4	17 (2.23), 23 (1.28), 18 (1.16), 10 (1.07), 22 (-1.08), 36 (-1.34), 21 (-1.46), 3 (-1.67), 20 (-1.72), 33 (-2.01)

**Distribution from highest to 1 and from -1 to the lowest*
Source: PQMethod analysis results

Table 11. Relevant distinguishing statements per factor, compared to other factors

Factor	Distinguishing statements (score sheet placement)*	Other factor	Placements in other factors (difference from distinguishing statement)**
1	5(9), 7(9), **14(7)**, 22(7), 13(3), 34(1)	2	7(-2), 6(-3), 5(-2), 6(-1), 7(+4), 9(+8)
		3	5(-4), 6(-3), 3(-4), 5(-2), 2(-1), 7(+6)
		4	8(-1), 5(-4), 5(-2), 3(-4), 5(+2), 7(+6)
2	**20(8)**, **13(7)**, **25(1)**	1	5(-3), 3(-4), 5(+4)
		3	4(-4), 2(-5), 7(+6)
		4	1(-7), 5(-2), 5(+4)
3	30(9), 4(8), **25(7)**, **15(7)**, 14(3), **32(3)**, 13(2), **27(1)**	1	4(-5), 6(-2), 5(-2), 6(-1), 7(+4), 6(+3), 3(+1), 3(+2)
		2	6(-3), 5(-3), 1(-6), 4(-3), 5(+2), 5(+2), 7(+5), 2(+1)
		4	6(-3), 6(-2), 5(-2), 4(-3), 5(+2), 5(+2), 5(+3), 6(+5)
4	17(9), **22(3)**, **36(3)**, 3(2), **20(1)**, **33(1)**	1	7(-2), 7(+4), 5(+2), 4(+2), 5(+4), 5(+4)
		2	7(-2), 6(+3), 5(+2), 4(+2), 8(+7), 5(+4)
		3	9(0), 5(+2), 5(+2), 4(+2), 4(+3), 5(+4)

**Distribution from highest to 7 and from 3 to the lowest **Same order as distinguishing statements column*
*Notes: All statements are significant with P < 0.05**. Bold statements are significant with P < 0.01***
Source: PQMethod analysis results

Interpretation of factor 1: the importance of openness and issue bias

As is shown in table 10b, factor 1 contains statements 5 ("revision and being open for criticism can help to tackle preferences of choosing subjects and kinds of evidence"), 7 ("research processes within the Dutch Safety Board are influenced by the public attention they get in a diversity of media, like news sites, papers and television programs"), 1 ("it is better to take the time to discuss the subject choice of research than to not consider it for long"), 16 ("within the Dutch Safety Board, there are specific preferences for which subjects are researched"), 14 ("the societal context surrounding research has unintended influence on the findings and results of the Dutch Safety Board"), 31 ("Preferences in choosing subjects and kinds of evidence can be tackled if an authoritative person creates rules for this") and 34 ("All available pieces of evidence must be collected to overcome poor research"). The interpretations of statements 31 and 34 are reversed, as these were placed on the disagree side of the score sheet. Statements 5, 7, 14, and 34 are distinguishing statements, placed significantly different on the score sheet than in other factors. Of these four, participants of factor 1 agree less with statement 34 and more with statements 5, 7 and 14. Two other distinguishing statements are statement 22 ("within the Dutch Safety Board, evidence is sometimes selected based on own ideas on the usability of it"), agreed with more in this factor, and statement 13 ("it is not a good thing that larger research is prioritized over smaller research"), generally agreed with less.

This factor reflects several opinions on issue bias. Statements 1, 5, 16 and 31 mention subject and evidence choice and are based on the theory of issue bias, concerning the bias in subject choice. This and other explanations of this factor's participants show that they think subject choice is an important aspect of investigating and should be discussed openly. Next to these statements, the same participants also explained that they agree with statement 7 and 14, because the societal context does indeed have influence, for example on what you focus on in investigations, and that it is important to recognize this and know how to correctly deal with this influence (participants 1, 2, 11, 12 and 16, personal communication, September 2018). On the other end, statement 31 is seen as something that would create bias. It would oppose openness and put forth a single person's bias as the defining rule (participants 1 and 2, personal communication, September 2018). Participants interpreted statement 34 as the subjects that are inquired about and found evidence for. Participants disagreed with the statement, because you have to make choices and cannot spend an infinite amount of time trying to find all available evidence. Furthermore, patterns will show up long before you find results (participants 1, 2, 12 and 16, personal communication, September 2018). The statement shows that participants think subject choice is important for the sake of limiting your investigations to a manageable degree.

The statements reflect that openness and the experienced importance of subject choice characterize this factor. It also show signs that participants of this group find issue bias important.

These statements show that both openness and the effects of issue bias are important to the participants of this factor. Distinguishing statements 5, 7, 14 and 34 show that participants of this group find this more important than participants of other factors. Of the two other distinguishing statements, participants agree more with statement 22 than others, while most others are more neutral towards this. It supports the idea that this factor's participants have more eye for which issues you investigate. Participants agreed less than others about statement 13. One participant explained that prioritizing larger over smaller investigations is not a bad thing, because it is probably a more relevant subject. The distinguishing statements show that openness and subject choice characterize this group in comparison to the other factors.

Interpretation of factor 2: the importance of science by the book

Factor 2 includes two statements in the factor distribution, namely the agreed with statement 11 ("the usage of multiple sources can overcome measurement errors") and disagreed with statement 25 ("poorly executed research contains nothing of value"). Statement 25 also is a distinguishing statement, agreed with less than in other factors. Distinguishing statements with which participants of this group agreed more than others are statements 20 ("research within the Dutch Safety Board is not insightful or open to criticism for the outside world") and 13 ("it is not a good thing that larger research is prioritized over smaller research").

As this group is represented by only two participants and there are only two statements most significant to fit in this factor, it is difficult to interpret. Participants did not agree with statement 11 to an extreme extent, so no explanation was given about this. However, both participants really disagreed with statement 25 and explained that every investigation has data you can use, as long as you know what you are doing. They indicated that you cannot assume that you are correct and have to keep looking (participants 5 and 7, personal communication, September 2018). So far, the participants appear to find it important to look everywhere, as they agree with the use of multiple sources as a tool to overcome measurement error and indicate that you can still make use of poorly executed investigations.

The fact that statement 25 is a distinguishing statement, meaning participants disagreed with it more than in the other groups, further strengthens this interpretation. Participants of this group agreed more with statement 20 that the Safety Board is not insightful or open to criticism from outside, which, although participants did not provide an explanation, might mean they see this as an issue. About statement 13, with which the participants agree with more than others,

one of the participants explained that you can learn a lot from smaller investigations (participant 7, personal communication, September 2018). So the participants do not only think it is important to look everywhere, but also not disregard anything or anyone. This factor might just show that these are scientists by the book, not disregarding anything and keeping open to everything. The factor does not appear to relate to any type of bias.

Interpretation of factor 3: importance of the value of output and publication

Factor 3 includes the agreed with statements 30 ("it is important to act against poorly executed research") and 4 ("it is wrong to not pay attention to the correctness of measurements in research") and the disagreed with statements 13 ("it is not a good thing that larger research is prioritized over smaller research"), 26 ("preferences for what does and does not get published can be countered by a separate and publicly available registry, within which research processes are described in detail"), 35 ("It is wrong to focus on only a couple of subjects within a research project") and 27 ("it is wrong to have preferences for what does and does not get published"). Of these statements, statements 30, 4, 13 and 27 are also distinguishing statements. Other distinguishing statements are the in this factor agreed more with statements 25 ("poorly executed research contains nothing of value") and 15 ("within the Dutch Safety Board, research is influenced for having to be relevant and understandable for policymakers"), and agreed less with statements 14 ("the societal context surrounding research has unintended influence on the findings and results of the Dutch Safety Board") and 32 ("comparing results in research reports would help to discover if there are preferences for publishing results with stronger results").

Half of the statements, 13, 26 and 27, are based on the theoretical construct of publication bias, of which statement 13 and 27 are also distinguishing statements. The same goes for distinguishing statements 15 and 32. Participants explain they agree with statement 30, because poorly executed investigations could be harmful to the Safety Board's reputation after its publication (participants 3 and 13, personal communication, September 2018). One participant explained they agree with statement 4, because correct measurements are the foundation of investigations and could harm the value of conclusions (participant 9, personal communication, September 2018). Participants explained they disagree with statement 35 and think it is good to focus on only a few subjects for the sake of the added value after publication (participants 3, 9, 10 and 20, personal communication, September 2018). All statements show that participants of this factor find output and publication important. Participants find correctness and added value of investigations more important than publication bias itself. The

factor has many statements of publication bias, but participants do not see publication bias as the major issue. They do see methods and technical bias as potential issues.

As distinguishing statements 30, 4, 13 and 27 are also included in the factor, they support this interpretation. This is also supported by the fact that participants agreed more than those of other groups with statement 25 and 15. Participants 13 and 20 explain about statement 25 that in output, poorly executed investigations are deemed harmful (participants 13 and 20, personal communication, September 2018). Participant 10 explains they agree with statement 15 that being relevant to policymakers is important and adds to the added value and goals of their investigations (personal communication, September 2018). Participants agree less than others about statement 32 for the fact that they do not see the proposed solution as a valid solution for the Safety Board (participants 13 and 20, personal communication, September 2018). The distinguishing statements further support that this factor is a group that finds importance in the value of their output and what they publish.

Interpretation of factor 4: the importance of technical bias

Last is the largest group of factors, factor 4. Included in this are the agreed with statements 17 ("it is wrong to not consider the quality of evidence in research"), 23 ("it is better to put time in avoiding measurement errors than save time by doing it to a lesser degree"), 18 ("to do something against poor research, research has to be insightful and open to criticism from the outside world") and 10 ("by exposing the research process to criticism, measurement errors can be overcome"), and the disagreed with statements 22 ("within the Dutch Safety Board, evidence is sometimes selected based on own ideas on the usability of it"), 36 ("within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence"), 21 ("conclusions in the Dutch Safety Board's research do not always connect to the findings and results of research"), 3 ("in the Dutch Safety Board, there are specific preferences to not publish something if nothing new is found"), 20 ("research within the Dutch Safety Board is not insightful or open to criticism for the outside world") and 33 ("findings in the Dutch Safety Board's research are influenced by social desirability"). Of these statements, statement 17 is a relatively higher placed distinguishing statement and statements 22, 36, 3, 20 and 33 are relatively lower placed distinguishing statements. There are no other distinguishing statements than the ones loaded on the factor, so these are not discussed separately.

The factor is primarily a combination of technical bias and methods bias, which shows that participants in this factor hold much importance to this. The exception to this is statement 36. Explanation from one participant show they agree with the importance of 'evidence' and

because there are no specific preferences for this experienced in the Safety Board (participant 4, personal communication, September 2018). The most important statements to look at are statements 17, 20 and 33, as these are agreed with most. These scores indicate that participants who loaded on this factor find it important to consider the quality of evidence in investigations and feel that this is done correctly at the Safety Board. The less strongly placed statements, 22, 36 and 3 all show that these participants feel that the Safety Board acts well and correctly. Only in the explanations for statement 3 evidence and quality of investigations are not mentioned and it does not support the factor as being about technical bias. As all other statements do, this supports the idea that this factor is about technical bias and the experienced lack of technical bias problems at the Safety Board.

5.2 Consensus statements: the statements different groups agree on

There are three consensus statements participants in the Q-study generally agreed upon and do not load significantly on any of the factors. These are statements 2, 8 and 24. These were all placed in categories 4 to 6. This shows that participants generally did not have a very strong opinion on them. As during the Q-sorting sessions no questions were asked about the middle categories, there are no direct answers as to why participants agree on these three statements. Indirectly, however, some explanation can be given.

Statement 2 ("with some research of the Dutch Safety Board, there is no consideration for the representativity of results towards the researched group") has the problem that it is not completely clear in what context 'representativity' is used. Clarification of a single participant addressed that they do not have a strong opinion on the statement. They stated that, if it occurs, it is not a big problem. The one participant who completely agreed with the statement explained that often there is only one case to look at, which makes broader representativity difficult (participant 17, personal communication, September 2018).

Second, statement 8 ("To do something against poor research, every kind of research needs to have its own criteria") includes three examples of ambiguity, namely 'poor research', 'every kind of research' and 'criteria'. Participants indicated that it is not completely clear what 'poor research' means, asked what the different kinds of research are the statement refers to and what kind of criteria this would be. One participant completely agreed, indicating that it is basic knowledge to have criteria in an investigation (participants 15 and 19, personal communication, September 2018). Another participant completely disagreed, explaining that there are already criteria in place, like providing feedback to each other, that are the same for all different kinds of investigations (participant 10, personal communication, September 2018).

Lastly, statement 24 (“it is important to act against preferences for a certain subject”) also has an ambiguity problem. Questions that came up during sessions addressed the statement’s use of the word ‘subject’. It could refer to subjects chosen to address within an investigation or investigations themselves. Furthermore, participants simply could not make up their mind on this statement, as is illustrated by its mostly neutral placement.

Most participants were neutral towards the statements and this choice usually had something to do with ambiguity in what a statement is exactly referring to. Unfortunately, this does not help explain the groups in the factorial distribution. It does show that these three statements do not belong in any of the groups of subjectivity. These statements also show explanation for the rest of the statements that were not included in any of the four factors.

5.3 Statement choice and factor distribution explanations

Based on the findings and the backgrounds of participants, this section examines if backgrounds show an indication for what factor participants belong to the most. Backgrounds are examined for the type of background and the amount of years working at the Safety Board separately. Readers should be aware that this is not statistically grounded and merely meant to see if there are indications that background matters in people’s opinions on bias.

Factor 1: background examination of the openness and issue bias group

Five participants loaded on factor 1, showing they find openness and issue bias very important. The backgrounds of this group are both technical and social. The group contains participants employed at the Safety Board less than five years, more than five years and more than ten years. Participants in this category work in different kinds of departments of the Safety Board. This could indicate that openness and issue bias are not seen as more important by people of a specific background.

Factor 2: background examination of the scientists by the book group

As only two participants belong to this group, it is difficult to say much about it. It contains participants of both technical and social background. What they have in common is that they have both been employed less than five years. This could indicate that people that have worked at the Safety Board for a shorter period are more used to working by the book than longer employed longer people. It is difficult to say for sure, as the group only contains two participants.

Factor 3: background examination of the value of output and publication group

The participants that see value of output and publication as most important includes a combination of three people who have been employed at the Safety Board for less than five

years and three who have been employed for more than ten years. Also, most participants in this group do not belong to any specific department. Backgrounds are both technical and social. Similar to factor 1, not much can be said about this group. It might suggest that those without department focus more on the added value of investigations to society than department-specific investigators. People of a specific department might be more familiar with the people in their department and therefore be more ingrained in a status quo of how things are handled in a department. However, it could mean many things and this is merely speculation.

Factor 4: background examination of the technical bias group

Similar to the other factors, the group that finds technical bias important contains participants from both technical and social backgrounds, from different categories of employment time and different departments of the safety board. This shows that awareness of technical bias is not shared by only people with a certain background.

5.4 Alternate examination of backgrounds

This final section discusses backgrounds in two additional ways. Same as the previous section, this is not statistically grounded. First, as another way of investigating backgrounds, the first two subsections examine the statements that are most popular in each category and whether factors can be recognized in the most popular statements per background category. The aim of these subsections is to gain more knowledge on the effects of background variables on opinions about bias. Response differences of the two study and work background categories are examined first, then the three categories for years of employment. In the third subsection, the other possible explanations of the theoretical framework are explored to see if these correlate to the responses participants provided during the Q-sorting sessions.

Response differences in background

Of the participants, eleven have a more technical background and nine a more social one. Although these are broad categories, it is interesting to see if certain statements are prominent for people with a certain background. First, participants with a technical background are examined. After that, participants with a social background are examined. Table 12 shows an overview of which statements scored best for participants with a technical and a social background and which statements are uniquely supported by only one group. This subsection discusses categories 1 and 2 together first, then 8 and 9 and the total amount of statements. After this, the unique statements are discussed. Lastly, a brief conclusion follows on if this says anything about the different groups.

Table 12. Most relevant extremely placed statement count for each background category

Score categories	Technical background	Social background	All participants
1 and 2 (disagree)	20, 26, 27, 31, 35	26, 27, 31, 35	26, 27, 35
8 and 9 (agree)	1, 5, 7, 11, 17	10, 18, 30	1, 5, 7, 11, 17
Total	1, 17, 27, 35	26, 34, 35	1, 26, 27, 34, 35
Unique*	29, 33, 36	32	-

**A statement is placed in an extreme spot by 2 or more participants and doesn't appear in the other group*
Source: author

Generally both groups disagree with statements 26, 27, 31 and 35. There is a slight difference in opinion about statement 20 ("research within the Dutch Safety Board is not insightful or open to criticism for the outside world"). While three participants of the technical group strongly disagrees, only one participant with a social background did and the rest does not have a relatively strong opinion on the statement. There is not much difference in the two groups' opinions on statements 1, 5, 7, 10, 18 and 30. However, six participants of the group with a technical background strongly agreed with statement 17 ("it is wrong to not consider the quality of evidence in research"), while only one participant of the other group strongly agrees and the rest does not have a relatively strong opinion. The statements that were placed on extreme categories in total contain no other statements to add on to this.

Statements 29 ("within the Dutch Safety Board, there are specific preferences for what type of results are highlighted in research"), 33 ("findings in the Dutch Safety Board's research are influenced by social desirability") and 36 ("within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence") were only placed on the extreme score sheet spots by participants with a technical background, while statement 32 ("comparing results in research reports would help to discover if there are preferences for publishing results with stronger results") by participants with a social background. Although statements 32 and 29 are only placed by one person in both an extreme 'agree' and 'disagree' spot, it does show that participants have a strong opinion on them, one way or another. Participants with a technical background strongly disagreed with statements 33 and 36. The unique statements do not uncover a relation between background and theory or factors.

In total, a certain background does not appear to coincide with one of the discovered factors when comparing the unique and most placed statements with the factors found in the analysis. There does not appear to be a direct connection between statements and background categories. However, certain statements about bias are preferred by one group over the other.

Theoretically, the most extremely placed statements of participants with a technical background do not fit with a type of bias specifically. However, it is very noticeable that all four statements of this group fit in the normative category of statements. The statements of participants with a social background show no relation at all to each other.

Response differences in years of employment

Of all participants, nine have been employed at the Safety Board for less than five years, five for five to ten years and six for ten years or longer. Similar to the previous section, the categories of employment are compared to each other. The statements participants of each category disagree with using categories 1 or 2 are discussed first, then the statements participants agree with using categories 8 and 9. Next the statements that have been extremely placed the most in total are discussed and the unique statements to each participant category. Lastly, a brief conclusion follows on if this says anything about the different groups. An overview of the most relevant statements is found in table 13.

Table 13. Most relevant extremely placed statement count for different years of employment

Score categories	0-5 years of employment	5-10 years of employment	10+ years of employment	All participants
1 and 2 (disagree)	25, 26, 27, 31, 35	34	27, 35	26, 27, 31, 34, 35
8 and 9 (agree)	5, 11, 17, 34	5	1	1, 5, 7, 11, 17
Total	25, 26, 27, 34, 35	34, 35	1, 27, 35	1, 26, 27, 34, 35
Unique*	23, 25, 28	36	-	-

**A statement is placed in an extreme spot by 2 or more participants and doesn't appear in the other groups*
Source: author

Similar to the previous section, statement 35 ("it is wrong to focus on only a couple of subjects within a research project") is disliked the most across all categories of participants. Of the remaining statements, two statements stand out. First of all, most participants that have been employed for less than five years strongly dislike statement 26 ("preferences for what does and does not get published can be countered by a separate and publicly available registry, within which research processes are described in detail"). The reasons given by participants are mostly that it would not work as a solution for the Safety Board and it should be questioned if publication preferences of the organization should get countered in the first place. It is very noticeable how much more the participants of this the 0-5 years group disagree with this statement than those of the other groups. The second statement to stand out is statement 34 ("all available pieces of evidence must be collected to overcome poor research"). The opinion of the '0-5 years of employment' group completely contrasts that of the '5-10 years of employment'

group. The ‘10+ years of employment’ group does not have a relatively strong opinion on this statement. This split in opinion seems to have to do with experiences at the Safety Board. Participants who have been employed for a shorter amount of time explain that you should always strive to be as complete as possible and broaden your scope and array of evidence. On the other hand, participants who have been employed longer emphasize that collecting ‘all’ available pieces of evidence is impossible and you should be realistic and practical in how you handle investigations. They explain that it would take years to find all evidence and even then you would not know for sure if you have gathered all evidence. The statements that have been placed in extreme spots the most do not reveal other statements than the ones discussed.

Of the positive statements, statement 1(“it is better to take the time to discuss the subject choice of research than to not consider it for long”) is generally liked the most by participants of all categories. The opinions about statements 5, 11 and 17 barely differ between the groups.

The unique statements of participants who have been employed for less than five years at the Safety Board are 23 (“it is better to put time in avoiding measurement errors than save time by doing it to a lesser degree”), 25 (“poorly executed research contains nothing of value”) and 28 (“preferences in choosing subjects and kinds of evidence can be tackled by determining which subjects and kinds of evidence are appropriate for every separate research project”). Statement 36 (“within the Dutch Safety Board, there are specific preferences for what type of data is and is not used as evidence”) is uniquely placed at extreme spots of the score sheet by participants that have been employed between five and ten years at the Safety Board. Most noticeable is that five participants of the ‘0-5 years employed’ group strongly agree or disagree with statement 25, while none of participants of the other groups do. Next to this, the unique statements do not have connections to the theoretical framework or the factors.

Similar to the previous section, there does not appear to be a connection between the categories of employment with the factors or the theoretical framework. Several specific statements, like 26, 34 and 25 show some interesting connections between statements and employment, but this does not contribute much to learning about bias and what someone’s background might mean in connection to it.

Alternative explanations to bias views

Next to background variables, not many explanations of specific types of bias, as discussed in the theoretical framework, appeared in the explanations of participants. Those that did are explained in this subsection. Time pressure appears to be an important factor in participants’ opinions, as an investigation has to be relevant to society and improving the safety of society

before too much time has passed. The timing of publication is also an important factor for the impact the Safety Board has. Two of the statements, statements 1 and 23, also mention time as a factor to consider in making decisions. When looking at the factors, these are placed in issue bias and technical bias respectively. When looking at the factor distribution that shows all significant statements per factor, each factor contains at least one of the two statements. Both are significant for publication bias. It seems logical that time pressure plays an important role in publication. One participant explains about statement 34 that you cannot keep collecting information forever. There is no time to attempt collect all information. Choices have to be made in order to be relevant (participant 17, personal communication, September 2018). Several signs point at time pressure is most relevant for publication bias, but overall, there is little evidence to strongly support this.

Complexity of cases also plays a role in views on bias. For example, statement 34 of collecting all available evidence is seen by many as impossible to do with the complexity of evidence. Choices have to be made, based on valid reasons. This explanation comes closest to issue bias, just like it was presented in the theoretical framework. This statement also happens to load on factor 1 of participants who find issue bias important and is a distinguishing factor for it. This shows support for complexity being primarily linked to issue bias.

Contestation, on the other hand, does not seem to have much impact in the case of the Safety Board. Most participants explained that they have a responsibility to stay independent (personal communication, September 2018). Lastly, keeping readers interested does not seem to be the motivation to publish certain kinds of investigations over others, but to provide relevant investigations that will provide an improvement in society's safety.

5.5 Results summary

This chapter has interpreted the four factors from the Q-study and examined whether there is indication that backgrounds influence opinions on bias. The subjects characterizing what the groups think is important are 'openness and issue bias', 'science by the book', 'value of output and publication' and 'technical bias'. Background and years of employment had little influence, but time pressure may be connected to those who think publication bias is important, while complexity is likely connected to thinking issue bias is important. In the next chapter, the results of this chapter are used to answer the final sub-question and the research question.

6. Conclusion and discussion: the theory and practice of bias

The goal of this thesis is to learn more about bias in theory and practice with the following research question:

What views on the existence, desirability and solutions regarding types of bias exist at the Dutch Safety Board and how can these views be explained?

To support the research question, four sub-questions are asked. The first sub-question ("*What does existing literature teach us about views on the existence, desirability, solutions regarding different types of bias?")* was answered in the theoretical framework, which distinguishes four types of bias: issue bias, technical bias, methods bias and publication bias.

The second sub-question ("*What factors explain the differences in the views on the existence, desirability and solutions regarding types of bias?")* was answered in the final section of the theoretical framework. Although several explanations exist for the different types of bias, the most prominent explanation in all cases appears to be a person's background.

Next, the third sub-question ("*How can the Q methodology work to uncover views on the existence, desirability and solutions regarding types of bias at the Dutch Safety Board?*") was discussed in the data and methods section. It shows how the Q methodology works as a tool to research clusters of subjectivity. The analysis highlights the methodology and applies it.

Lastly, the final sub-question ("*What views on the existence, desirability and solutions regarding types of bias are uncovered from using the Q methodology at the Dutch Safety Board and how can they be explained?*") was answered in the analysis and results sections. Three main factors are uncovered and an attempt was made to explain the different groups. The next section discusses the results in detail and answers the research question.

After answering the research question, the final sections discuss and reflect on the thesis in five sections. Four sections reflect on the extent to which this thesis is also biased to a certain degree, on how well the Q methodology works as a research tool, on the Q-set of statements used in this thesis and some further reflection on plusses and minuses. In the final section, several opportunities and possibilities for future research are discussed.

6.1 Answering the research question: the theory and practice of bias

Theory

As the previous section mentioned, the theory discusses the existence, desirability and solutions regarding four types of bias. Issue bias is the biased preference for one issue over others. It might be a good thing to have issue bias, as you could gain unique knowledge over specific subjects. A possible solution is good governance of evidence, including concepts like peer feedback, and deliberative inquiry.

Technical bias is the bias in results caused by flawed research design. However, flawed research is not useless, as most research contains at least some pearls of wisdom. Some solutions to the problem of flawed research is the good governance of evidence, including concepts like transparency, and reweighting samples.

The third type of bias, methods bias, is bias in results that appears due to the methods or the application of methods used to collect the data instead of results coming forth from the data itself. Although this is not desirable, it is unavoidable to have some methods bias in research. Therefore, adopting 'good enough' methods is preferred over trying to eliminate all methods bias. Solutions to methods bias are triangulation, creating distance between measurements, statistical checks and transparency.

Lastly, publication bias is the tendency to selectively publish certain types of articles more than others. For example, studies with significant results could get prioritized over studies without significant effects. However, this might not be a problem per se, as this could also mean that larger studies that are closer to true effects and studies with higher quality get prioritized. A way to counteract the negative forms of publication bias, published studies can get compared with unpublished studies for which do and do not have significant results. Other solutions are funnel plots, public registries of both published and unpublished results and publication policies.

Practice

In practice, every factor displays a very distinct view on the role of bias and what is important in research. The statements based on characteristics of the four theoretical types of bias did not appear together in the same way as they did in theory. In the results section, some support was found for the existence of issue bias, technical bias and publication bias as separate types of bias. However, the goal of this study is not to confirm these types. On the other hand, the fact that they showed up in the case study is an interesting finding.

Factor 1 shows a group of participants that thinks openness and subject choice are very important and is concerned about issue bias in research. The theoretical idea of issue bias can

be recognized quite well here in participants' opinions. Factor 2 shows no sign of a specific kind of bias. However, it did show a group of two participants who are scientists by the book and find it important to be complete in research. Factor 3 shows signs of technical, methods and publication bias. As participants of this group find the value of the output and of what is published very important, they show they are mostly concerned with that what is published is correct. That there is publication bias to a certain degree they agree with, but they do not see this as a problem, as there are reasons only certain research is published. Research needs to have added value and relevancy for society and those who have to use the research to improve society's safety. This factor very much relates to the specific role the Dutch Safety Board holds in society. The last factor group, factor 4, was very clearly concerned with the importance of technical bias, although it also contains statements about methods bias. They think it is important to properly execute research. Technical bias and methods bias come together here for the sake of well performed research. So in this case, methods bias does not really appear as a separate type of bias from technical bias, as was also mentioned in the theoretical framework.

Explaining the views found in practice

In the case study of the Dutch Safety Board, similar concepts to the theory can be found. An explanation for why practice differs from theory is because the Safety Board has specific tasks and goals, aiming to improve safety in society. This is different than the purely scientifically-minded goals of scientific research studies. Next to this, the views on bias can be explained by the answers to the second sub-question.

The most prominent influence on people's views is someone's background. Therefore, in the Q-sorting sessions, participants were also asked about their study background, work background and how long they have been employed at the Dutch Safety Board. As was shown in the final section of the results chapter, these background variables do not explain why respondents end up belonging to different groups of subjectivity. However, participants with certain backgrounds did answer differently to very specific statements. The most interesting difference is that participants who have been employed for less than five years almost completely agree that all available pieces of evidence should be collected, while participants who have been employed for five to ten years disagreed. Although the background comparisons are not statistically grounded, this does show that background variables might have some impact on what people think about specific aspects of bias and research practices.

Next to this, not many explanations of specific types of bias, as discussed in the theoretical framework, appeared in the explanations of participants. As was discussed in the

results section, time pressure was an important factor in participants' opinions, as research has to be relevant to society and improving the safety of society before too much time has passed. It appears to be most tied to publication bias and not methods bias, as it was discussed in the theoretical framework. Another explanation of views is the complexity of evidence, which is especially connected to issue bias. Complexity means choices have to be made on what is most relevant and important to do research on. It does not necessarily lead to issue bias in research, as people can discuss their preferences, but it is shown to play an important role for issue bias.

6.2 Conclusions

The conclusion has shown that, to answer the research question, there are four distinct kinds of views on bias in the context of an organization differ from theory. The groups see the most importance in 'openness and issue bias', 'science by the book', 'value of output and publication' and 'technical bias'. Issue bias and technical bias are seen as a problem to an extent, but issue bias is also seen as unavoidable to have. It is deemed more important on how you deal with bias in practice and be open to feedback. Methods bias is connected to technical bias, instead of it being a separate type of bias. Publication bias is not seen as inherently bad, as something has to be relevant. Explanations for participants' views by examining background variables and some other possible variables are limited, but show that the role of the Dutch Safety Board as context plays an important role. Background and employment time have some influence, time pressure explains some views on publication bias and complexity views on issue bias.

The following sections discuss and reflect on several aspects of the thesis and present what could be investigated in future research.

6.3 Bias showcase: applying theory to thesis

In this thesis, researchers in general and the people of the Dutch Safety Board specifically have been connected to the biases they may or may not have. Although this research is not aimed at judging or evaluating others for having or not having bias, it does seem to pretend to have all the answers. This section shows that bias is everywhere. Even a thesis about bias will be biased in one way or another, as it is written by a human being. As many participants said during Q-sorting, it is not a bad thing if you are biased as a person, as everyone is. What matters is how you deal with the biases you have. For example, you could report the biases you have. This section will do just that for every type of bias discussed in the theoretical framework, ending with a general conclusion on the presence of bias in this thesis.

The presence of issue bias

Issue bias was present to a limited degree. While an overarching set of bias types was made in the theoretical framework, the subjects and evidence discussed here are quite broad. In-depth discussions about the presence and problems of a very specific type of bias were left out of the typology of bias of this thesis. This shows a bias to simple and more broad evidence. The type of data discussed is a mixture of qualitative, discussed in the theoretical framework, and quantitative, discussed in the analysis and results sections.

The presence of technical bias

First and foremost, it is possible that the results from using the Q methodology are not interpreted entirely correctly or that there are some flaws in using the Q methodology itself. The method is new to me personally and I do not think I understand the method to its fullest extent. Furthermore, this thesis and its design are not perfect, so it will contain some flaws. The following sections discuss some of these flaws, but also show that much of the thesis has been carefully designed in order to be as correct as possible.

The presence of methods bias

Methods bias has played the largest part out of all the types of bias in the thesis. This contains both positives and negatives. The first and possibly largest issue might have created some differences between sessions and problems with consistency. During the time I was an intern at the Safety Board, I have explained the idea behind my thesis to a couple of employees, who would later be participants for my thesis research. However, this probably has not had much of an effect. One of these previously informed participants stated before a Q-sorting session that he remembers that I explained it once, but did not remember what it was about.

For the rest, the role of methods bias in this thesis involves the Q-sorting sessions and how these were setup and performed. Problems were unexpected, but also predictable. During the sessions, questions were asked about the statements. The answers given to the respondents were broad, consistent, limited, and deliberately open and vague. However, answering questions might still have steered the interpretation of the statements and participants' opinions in a certain direction. Also, providing answers on questions has created differences between sessions. This might have influenced the results in small ways, but a small difference is still a difference. During the sessions, participants provided feedback on the statements and the problems with the statements. These problems are discussed in section 6.5. For the rest, the Q-sorting sessions went almost completely perfect. The physical environment of every conversation was identical and all took place in the building of the Dutch Safety Board in a

soundproof room. Colleagues could walk by and see participants taking part in a session, but could not hear or see what participants had to say or think.

The presence of publication bias

Although publication bias does not play a role in writing the thesis, reporting bias does. This does not appear to play a large role. The only way it might play a role is that the positive results are highlighted most, compared to what has been found to not work out. However, even this is limited, as much of this is discussed in this final discussion chapter.

Conclusion on the presence of bias

In the end, this thesis does not appear to contain much bias. However, this result might be a result of a bias of positivity towards myself, as it is unlikely that someone would want to write and think negatively about themselves. So, although this section attempted to uncover my own bias of this thesis, this evaluation of my own bias should be taken with a grain of salt.

6.4 Reflection on Q as a research tool

The Q methodology is a good tool to research subjectivity and for this purpose, it has been used as the main research tool in this thesis. However, a problem with it is that there appears to still be some discussion on the Q methodology and how to use it best. During the application of the method, I discovered that not everyone agreed on what steps should specifically be taken and why. To be most sure of the correctness of the analysis, the most agreed upon steps were followed and every step was carefully argued. There is still a lot to discover about the Q methodology and what is best. Something that is not completely clear after having performed the complete analysis is whether or not the score sheet distribution of how many statements could be fit in each category should have been flatter, for example four statements per category, instead of the current bell-shaped distribution. Articles about Q were not very clear on this, while it could lead to some different combinations of statements in the factor analysis. Overall, the Q methodology feels useful for this research. The analysis and results appear to be correct. However, some of the vagueness and discussion surrounding the methodology makes it a bit uncomfortable to use as a research tool.

6.5 Reflection on the statements

This section reflects on the statements in three ways. First, it is addressed that the answers to the statements cannot be used to evaluate the Safety Board or its employees. Second, the problems with some of the statements are discussed. Lastly, an explanation is given why, if there are such problems in the statements, I ‘muddled through’ with the Q-sorting and analysis.

First, as mentioned in the analysis, not much can be said about the Dutch Safety Board as an organization or its employees by looking at how they answered the statements. Agreeing and disagreeing with statements does not mean participants are horribly biased. Arguments given by participants show that a diversity of opinions on the statements exist at the Safety Board and that each is based on valid reasoning. Furthermore, as stated before, this thesis is not an evaluation on the Safety Board and thus nothing can and should be taken from it for other purposes than discussing bias as a theoretical concept and learning more about it.

During the Q-sorting statements, participants provided feedback on the statements and the problems with the statements. The most prominent problems mentioned with the statements are that some contain a double negative, are vague, can be interpreted in different ways, appear misinformed on certain concepts and ideas and are not really applicable to employees of the Safety Board. The double negative in statements is a poor oversight and is simply a mistake that should not have been present in the Q-set. The other problems are spread across and are very important to keep in mind when replicating this research. Although the statements were generally well understood by the participants, the feedback shows that they are still very lackluster. Luckily, even with these problems, they still show relevant results, as they group together participants who interpret and understand statements in a similar way.

Lastly, the question is why I continued the Q-sorting sessions and analysis as planned. The natural response to this is that appointments with all participants had been planned and there were no other chances. Changing the statements half-way would have been the worst solution, as this would split the group of participants in two very different halves, which would have made the entire research project worthless. As for the problems with the statements, there are less problems than with a 'normal' regression analysis. With the Q methodology, it matters less as statements are connected and compared with each other per person. So the differences between persons in their interpretation of the statements, which are likely to not be as bad as they seem to be on first glance, will not make an enormous impact on the results. Still, for repetition purposes a simpler and clearer Q-set is probably preferred.

6.6 Further reflection

There are two other aspects to reflect on. First, the representativity of the participant group. As is usual with assembling a sample group, an attempt is made to make this group somewhat representative. Although literature on the Q methodology specifically mentions that you can even use the Q methodology with a sample of a single person or a couple, as was mentioned in the data and methods section, the decision was made to involve a group of twenty people from

different departments of the Safety Board. With a larger sample, a random pick of participants will already likely do this, but with the small sample size, this could not be done. Articles on Q also mentioned that a sample should not be random, as was also mentioned in the data and methods section. However, a positive surprise during the analysis is that the participants are evenly distributed across different backgrounds and employment years as well.

Second, it can be said that the Q-sorting sessions are a good experience for both the researcher and the participants. The researcher gets to hear a lot of different viewpoints with an array of different arguments, while participants get to sort their thoughts on the statements that are presented to them. Some participants even noted they found the sorting assignment fun, as sorting the cards is like a sort of game. Furthermore, the application of the methodology on bias let the participants see for themselves which types of bias they are more and less aware of and increases awareness of the types they were less aware of before the Q-sorting session.

6.7 Future research on bias

This research has shown and discovered several different groups of opinions on bias in practice, different than in theory. It also shows people of an investigation-oriented organization recognize issue bias, technical bias and publication bias to a certain degree. However, there is still much more work that can be done regarding this topic. Replications can be performed with upgraded statements and theoretical framework to see if similar results are found about bias or not. There can be statistically grounded research on the influence on a person's background on bias, as there are hints that this could also have an effect. Bias is a complex subject and is an interesting and relevant topic to continue studying in theory and practice. Ninety percent of everything may be crud and ninety percent may be biased, but that same ninety percent also contains pearls of wisdom to continue learning more about the world.

Acknowledgment

I would like to thank the people of the Dutch Safety Board for guiding me during my internship and participating in the research for my thesis. I would also like to thank my thesis coordinator, Pieter Zwaan, for his feedback and suggestions.

References

Bauer, M.W. and Bucchi, M. (2008). *Journalism, science and society. Science communication between news and public relations.* Routledge: New York.

Bax, L. and Moons, K.G. (2011). Beyond publication bias. *Journal of Clinical Epidemiology*, 64, pp. 459-462.

Begley, C.G. and Ellis, L.M. (2012). Drug development: raise standards for preclinical cancer research. *Nature,* 483, pp. 531-533.

Bohnet, I. and Morse, G. (2016). Designing a bias-free organization. *Harvard Business Review*, 94(7/8), pp. 63-67.

Brown, S.R. (1993). A primer on Q methodology. *Operant Subjectivity*, 16(3/4), 91-138.

Brumfiel, G. (2009). Supplanting the old media? *NATURE,* 458, pp. 274-277.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flinkt, J., Robinson, E.S.J. and Munafò, M.R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience,* 14(5), pp. 365-376.

Carcasson, M. and Sprain, L. (2016). Beyond problem solving: reconceptualizing the work of public deliberation as deliberative inquiry. *Communication Theory,* 26, pp. 41-63.

Castensson, A. (April 27, 2015). *Perfect research is a myth.* Retrieved May 2017 from https://www.tidningencurie.se/en/nyheter/2015/04/27/perfect-research-is-a-myth/

Collier, D. and Mahoney, J. (1996). Insights and pitfalls. Selection bias in qualitative research. *World politics,* 49(1), pp. 56-91.

Coogan, J. and Herrington, N. (2011). Q methodology: an overview. *Research in secondary teacher education,* 1(2), pp. 24-28.

Cortes, C., Mohri, M., Riley, M. and Rostamizadeh, A. (2008). Sample selection bias correction theory. In *International conference on algorithmic learning theory* (pp. 38-53). Springer: Berlin, Heidelberg.

Costello, A.B. and Osborne, J.W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation,* 10(7).

Dutch Safety Board. (2018). *Onafhankelijk onderzoek in het publiek belang.*

Edland, A., & Svenson, O. (1993). Judgment and decision making under time pressure. In *Time pressure and stress in human judgment and decision making* (pp. 27-40). Springer, Boston, MA.

Entman, R.M. (2007). Framing bias: media in the distribution of power. *Journal of Communication,* 57, pp. 163-173.

Feinberg, M. (2007). Hidden bias to responsible bias: an approach to information systems based on Haraway's situated knowledges. *Information Research,* 12(4)

Ferretti, F., Guimarães Pereira, Â., Vértesy, D. and Hardeman, S. (2018). Research excellence indicators: time to reimagine the 'making of'?. *Science and Public Policy.*

Fudickar, R., Hottenrott, H., Lawson, C. (2016). *What's the price of consulting? Effects of public and private sector consulting on academic research.* Düsseldorf University Press: Düsseldorf.

Funtowicz, S. [SFuntowicz]. (2018, February 15). "Shortcomings in the existing mechanisms for science's quality control system, undermine trust in assessment practices around scientific excellence..., if the peer review system is in crisis, what research outcomes are evaluated as excellent?" [Tweet]. Retrieved March 2018 from https://twitter.com/SFuntowicz/status/964248869256224769

Gerber, A.S. and Malhotra, N. (2008). Publication bias in empirical sociological research: do arbitrary significance levels distort published results? *Sociological Methods Research,* 37(1), pp. 3-30.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report,* 8(4), pp. 597-606.

Harrison, D.A. and McLaughlin, M.E. (1993). Cognitive processes in self-report responses: test of item context effects in work attitude measures. *Journal of Applied Psychology*, 78(1), pp. 129-140.

Hebert, J.R., Clemow, L., Pbert, L., Ockene, I.S. and Ockene, J.K. (1995). Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *International Journal of Epidemology,* 24(2), pp. 389-398.

Hernán, M.A., Hernández-Diaz, S. and Robins, J.M. (2004). A structural approach to selection bias. *Epidemiology,* 15(5), pp. 615-625.

Holbrook, A.L., Green, M.C. and Krosnick, J.A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires. Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly,* 67, pp. 79-125.

Ioannidis, J.P.A. (2005). *Why most published research findings are false.* PLoS Medicine, 2(8), pp. 696-701.

Johnson, R.B., Onwuegbuzie, A.J. en Turner, L.A. (2007). Towards a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), pp. 112-133.

Kanuka, H. (2010). Deliberative inquiry. In: Savin-Baden, M. and Major, C.H. (2010). *New approaches to qualitative research. Wisdom and uncertainty.* Routledge: London and New York.

List randomizer. (n.d.). Retrieved from https://www.random.org/lists/

LoBiondo-Wood, G. and Haber, J. (2017). *Reliability and Validity. Nursing Research-E-Book: Methods and Critical Appraisal for Evidence-Based Practice.*

Luttrell, W. (2000). "Good enough" methods for ethnographic research. *Harvard Educational Review*, 70(4), pp. 499-523.

MacKenzie, S.B. and Podsakoff, P.M. (2012). Common method bias in marketing: causes, mechanisms, and procedural remedies. *Journal of Retailing,* 88(4), pp. 542-555.

Maede, A.W., Watson, A.M. and Kroustalis, C.M. (2007). *Assessing common methods bias in organizational research.*

McKeown, B. and Thomas, D.B. (2013). *Q Methodology*. SAGE Publications: Thousand Oaks.

Nagakawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology,* 15(6), pp. 1044-1045.

Onwuegbuzie, A. J. (2000). *Expanding the Framework of Internal and External Validity in Quantitative Research.*

Ochs, E. (2000). Socialization. *Journal of Linguistic Anthropology*, 9(1-2), pp. 230-233.

Oliver-Hoyo, M. and Allen, D. (2006). The use of triangulation methods in qualitative educational research. *Journal of College Science Teaching*, 35, pp. 42-47.

Parkhurst, J.O. (2016). Appeals to evidence for the resolution of wicked problems: the origins and mechanisms of evidentiary bias. *Policy Sciences*, 49(4), pp. 373-393.

Parkhurst, J.O. (2017). *The politics of evidence: from evidence-based policy to the good governance of evidence*. Routledge Studies in Governance and Public Policy: Abingdon, Oxon, UK.

Pauleen, D.J. and Murphy, P. (2005). In praise of cultural bias. *MIT Sloan Management Review*, 46(2), pp. 21-22.

Pawson, R. (2006). Digging for nuggets: how 'bad' research can yield 'good' evidence. *Int. J. Social Research Methodology*, 9(2), pp. 127-142.

Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R. and Rushton, L. (2006). *JAMA,* 295(6), pp. 676-680.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003). Common Method Biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology,* 88(5), pp. 879-903.

Podsakoff, P.M., MacKenzie, S.B., Podsakoff, N.P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annu. Rev. Psychol.*, 63, pp. 539-569.

Sarewitz, D., Foladori, G., Invernizzi N. and Garfinkel, M.S. (2004). Science policy in its social context. *Philosophy Today,* 48(supplement), pp. 67-83.

Schmolck, P. (2014). PQMethod [Computer software – version 2.35]. Retrieved from http://schmolck.org/qmethod/pqmanual.htm

Schriesheim, C.A. and Hill, K.D. (1981). Controlling acquiescence response bias by item reversals: the effect on questionnaire validity. *Educational and Psychological Measurement,* 41, pp. 1101-1114.

Serlin, R.C. and Lapsley, D.K. (1985). Rationality in psychological research. The good-enough principle. *American Psychologist,* 40(1), pp. 73-83.

Shaheen, N.J., Crosby, M.A., Boyzmski, E.M. and Sandler, R.S. (2000). Is there publication bias in the reporting of cancer risk in Barrett's esophagus? *Gastroenterology*

Smart, R.G. (1966). Subject selection bias in psychological research. *The Canadian Psychologist,* 7a(2), pp. 115-121

Sturgeon, T. (1958). On hand: a book. *Venture Science Fiction,* 2(2), pp. 66-69.

Tholen, B. (2017). *Virtue and responsibility in policy research and advice.* Springer.

Thompson, M. and Wildavsky, A. (1986). A cultural theory of information bias in organizations. *Journal of Management Studies*, 23(3), pp. 273-286.

Thornton, A. and Lee, P. (2000). Publication bias in meta-analysis: its causes and consequences. *Journal of Clínical Epidemology,* 53, pp. 207-216.

Turner, E.H. (2013). Publication bias, with a focus on psychiatry: causes and solutions. *CNS Drugs,* 27, pp. 457-468.

Van Exel, N.J.A. and De Graaf, G. (2005). *Q methodology: a sneak preview*. [available from www.jobvanexel.nl].

Wesselink, A., Colebatch, H. and Pearce, W. (2014). Evidence and policy: discourses, meanings and practices. *Policy Sci*, 47, pp. 339-344.

Williams, B., Onsman, A. and Brown, T. (2010). Exploratory factor analysis: a five-step guide for novices. *Journal of Emergency Primary Health Care*, 8(3).