HOGESCHOOL UTRECHT
MASTER OF INFORMATICS

MASTER'S THESIS

"Assessment criteria to reduce requirements uncertainty"

Author
Sjoert Ebben

Date
June 2018

Supervisor
Bas van Gils

Utrecht University of Applied Science
Faculty Science and Engineering
P.O. box 182
3500 AD UTRECHT
The Netherlands

Abstract

Creation of software and systems starts off with stakeholder needs, documented as requirements, most often at the early stages of the project. Requirements are created as part of a requirement engineering process and are input for the design and development process.

Large engineering projects, however, often suffer from the complexity or uncertainty of requirements, requirements that are often specified by and imposed on a contractor by the client. Uncertainty refers to the difference in the information necessary and the amount of information possessed. Research acknowledges that requirements uncertainty is one of the largest factors that have a negative impact on project performance.

This uncertainty will manifest itself at the point where the contractor receives the requirements. In the Dutch civil engineering sector, the V-model is the preferred development model, and at some point in the “V”, there is a transfer from the client to the contractor, in the “Guideline for Systems Engineering within the civil engineering sector” this is specified as a “transfer point”.

Although there are processes that can be used in the V-model to assess requirements, there is no *objective* list of criteria that should be taken into account in this process. Not having objective criteria means that the outcome of the assessment process becomes unpredictable, which could leave the contractor at risk when uncertain requirements are accepted and cause major rework in the design or development phase.

In the research for this master thesis, 66 objective criteria were distilled from scientific literature and reviewed via a Delphi study by a panel of experts within the field of systems engineering in the civil engineering area. Also from the literature review, six methods to measure uncertainty and two factors that influence how much uncertainty is acceptable were identified and reviewed by the expert panel.

The panel added extra criteria, measurement methods and factors and gave a relevance score to each item to make a distinction between items that were regarded relevant and items that were not. This resulted in validated lists of 37 requirement assessment criteria, 1 measurement method and 3 influencing factors. These lists can be used by companies in their requirements assessment process.

The results of this research contribute to the body of knowledge on project risk management and will enable contractors to reduce initial uncertainty to improve project performance.

Acknowledgements

This research is the final destination of a two-year Master of Informatics journey, a journey with many highs and some lows and many people to thank for. I cannot name all of you in person, but please be assured of my gratitude.

The journey started off with the intention of doing the postgraduate course, which is, very simplified, half the Master but without the research. The first module, Enterprise Architecture, however, made me realise there is much more to be had to satisfy my appetite for knowledge in the field of (enterprise) architecture and business-IT alignment. My enthusiasm was specifically sparked by the Enterprise Architecture lecturer, Bas van Gils, who would later become my mentor for this thesis. He should be thanked (or blamed) for my decision to go for the full monty and for his sharp reviews that brought this thesis to the next level.

Making the Master possible in practical terms (budget and time), required the support of my employer, VX Company. I'm grateful I was given the opportunity, given that at that time there was uncertainty about the role of consultants and enterprise architects within a development focussed company. A special thank you goes to my manager, Erik Zeillemaker, for finding a way to make all of it possible.

Arriving at the research question was a journey by itself. Luckily, my client Vialis offered an environment where I was inspired to find a suitable topic. They also took the time to brainstorm and give feedback on my early research proposals, for which I'm very grateful. By the end of 2017, I had two potential research questions in mind and was in doubt which one to choose. A brainstorm session with Marlies van Steenbergen helped me to make the decision, a special thanks to her for that important moment in time.

Marlies and Anand Sheombar should be thanked for guiding me back to the Delphi research method, a method that was high on my research methods itinerary half way 2017 but that somehow fell off the radar. Although using the Delphi method meant I needed an extra month of my life for this thesis, I'm convinced it added extra value to the research. Anand also deserves a special thank you for his great help during the preparation phase, his enthusiasm is contagious.

Although the ethics of the Delphi study prohibits naming the research participants in person, I am humbled by the great effort they made to fill in the rather comprehensive questionnaires. Their valuable feedback was essential to the research findings. I owe ye!

Warm thoughts go to my fellow students, who accompanied me on this journey for the past two years. We shared an experience of hard work, intensive collaboration and challenging assignments. I enjoyed working with you all, hopefully we'll have the chance to work together on great projects in the real world.

A Master study, including research, in little under two years is not something one can do without making sacrifices elsewhere. My dear friends and family, I have neglected you, but you kept supporting me and I thank you for that. I'm looking forward to spending more time with you again, enjoying the good things in life.

Finally, a special thank you goes to my wife Cathérine and my kids, Nox and Trinity. The sacrifices we made as a family with two parents doing a master are great, but we made it through with love and care for each other. Let's reclaim the quality time, the journey has ended.

Sjoert Ebben
June 2018

Table of Content

List of figures

List of tables

1 Introduction

The idea for my research originated during a large ***systems engineering*** project in the civil engineering sector: enlarging the motorway A6/A1 from Almere Havendreef to Diemen, including the construction of a reversible lane with a custom-made operating system. Figure 1 shows the demarcation of the project.

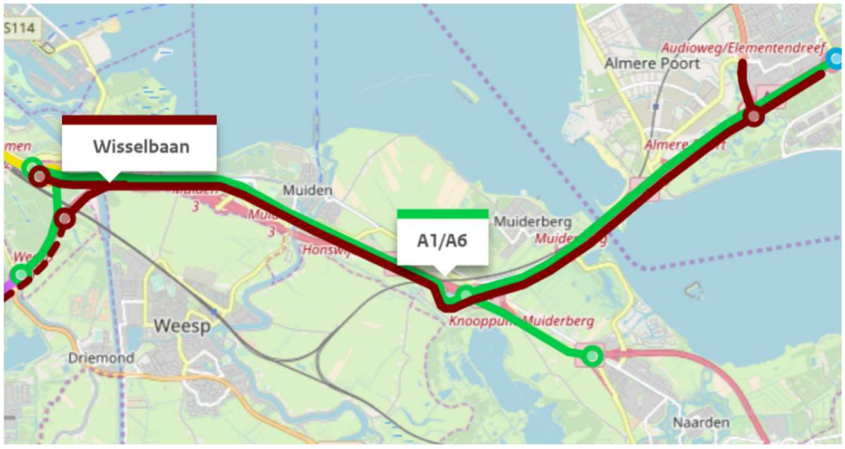

Figure 1: "Demarcation of the project"

I was involved in the project from September 2016-November 2017 in the role of senior information analyst, in between the worlds of systems and software engineering. In the project, ***requirements*** were specified by the client (Rijkswaterstaat) and transferred to the contractor (SAAOne) for development in a project that used the ***V-model*** as its development model.

Throughout the project, there were discussions and ***uncertainties*** about the requirements, resulting in rework and discussions with stakeholders about scope and interpretations of the requirements. The problem of uncertainty about the requirements surfaced in the use case documentation as the interpretation of the requirements by the design team differed from the interpretation by the client, so many use cases had to be redefined.

One of the key findings of the project evaluation in November 2017 was that for a next project, the requirements should be assessed more thoroughly before accepting them. A similar finding was reported by another infrastructure project early 2018 and although further information about that project is classified, examples of the important role of requirements in engineering project success or failure can be found in literature (Alami, 2016; Terry Bahill & Henderson, 2005).

Although requirements assessment exists as a process, in the context of the projects I was involved in it seems to be based mainly on past experience, not on objective criteria so there can be debate on what “more thoroughly” (as a finding from the evaluation mentioned above) actually means. The aim of this research is to provide objective criteria to assess requirements.

1.1 Reading guide

In chapter 2, background information is provided on the definitions and concepts mentioned in the introduction. This leads to the problem statement, the research questions and the conceptual model, after which the research context is specified.
Chapter 3 is concerned with the research method and specifies the research outline as well as the way the literature review and research method concerning the research questions are constructed.
The literature review on the research questions can be found in chapter 4, the findings of this review as well as the findings of the research rounds can be found in chapter 5. Finally, in chapter 6 the conclusions, contribution, recommendations, limitations and suggestions for future research can be found.

2 Background

In the introduction, several terms were introduced. These terms will be explained in more detail in this chapter to build up to the problem definition and the research question. The reason for explaining the terms is to help the reader to understand the background of the research and to help the reader to remain focussed on the research findings without being distracted by unfamiliar terms.

The chapter starts with information about what systems engineering is and how it is scoped as the research was done in the context of a systems engineering environment. To help the reader to understand the subject matter even more, the terms “system” and “engineering” will also be explained in more detail.

Requirements play an important role in systems engineering as they specify what needs to be build and by what constraints. Therefore, the role of requirements in systems engineering is explained in its own paragraph. In this paragraph, the path from the requirements engineering process to the deliverables of the project will be visualised. This visualisation will later be reused in a more elaborate format to identify the research topic.

All projects concerning some sort of development use a “development model” to guide the project from start to end. There are many types of development models, in this chapter the preferred development model for systems engineering projects within the context of this thesis is identified. This will provide insight in which project phase requirement assessment (and the research for this thesis) plays its role.

Then the concept of “uncertainty” is introduced. To build something in a (systems engineering) project, it should be known what to build and what the requirements and constraints are but there is often uncertainty about some element, like budget, schedule or requirements. As this thesis is about reducing requirements uncertainty, the next paragraph is dedicated to explaining in more detail what requirements uncertainty is and where it surfaces in a project. The following paragraph then identifies the negative effects of requirements uncertainty and how to deal with them in a requirements assessment process.

In dealing with requirements uncertainty in this process, however, there are some problems, as identified in the problem statement. The problems identified here lead to the research questions for this thesis, followed by the conceptual model and the identification of the research context and research contribution.

2.1 Systems engineering

The context for this research is within the area of systems engineering. The International Council on Systems Engineering (INCOSE) defines systems engineering as

> *“an interdisciplinary approach and means to enable the realization of successful systems. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, then proceeding with design synthesis and system validation”* (International Council on Systems Engineering (INCOSE), n.d.).

A more compact description of the function of systems engineering is provided by Kossiakoff et al. (2011): “*guide the engineering of complex systems*”.

As the INCOSE version includes terms like “needs” and “required”, that definition seems to be the best fit for this thesis as the research topic concerns requirements.

2.1.1 System

According to the Systems Engineering Body of Knowledge (SEBoK), the term “system” should be interpreted as an “engineered system” a *technical* or *socio-technical systems* system, created by and for people (SEBoK, 2017b).

Other definitions are:

An integrated set of elements, subsystems, or assemblies that accomplish a defined objective. These elements include products (hardware, software, firmware), processes, people, information, techniques, facilities, services, and other support elements. (Walden, Roedler, Forsberg, Hamelin, & Shortell, 2015)

A combination of interacting elements organized to achieve one or more stated purposes. (ISO/IEC, 2015)

All sources confirm that a system consists of elements that interact with each other, as illustrated in Figure 2.

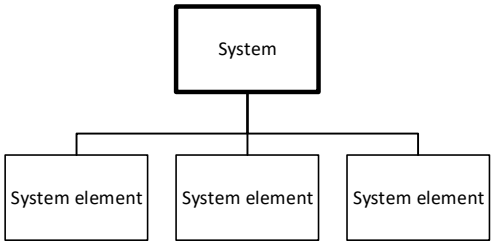

Figure 2: "Breakdown of a system structure (adapted from ISO/IEC, 2015, fig. 1)"

A system element can be at the atomic level, i.e. it cannot be further decomposed, or be representations of systems on their own (Walden et al., 2015) as illustrated in Figure 3.

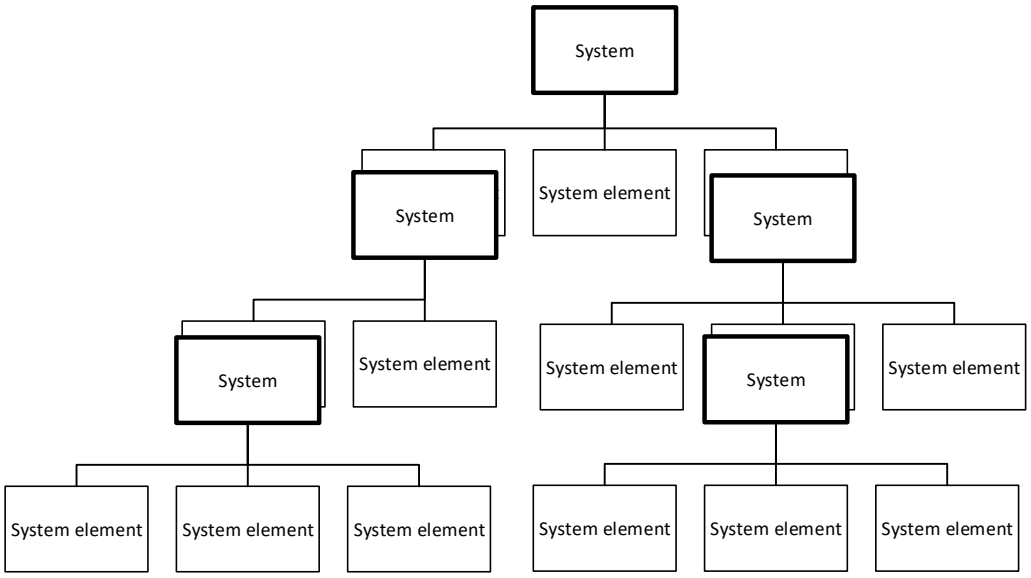

Figure 3: "A system composed of interacting system elements (adapted from ISO/IEC, 2015, fig. 2)"

All definitions provided above are suitable for this thesis, the one by Walden, Roedler, Forsberg, Hamelin, & Shortell, however, seems to be the most comprehensive one as it specifies more clearly what type of elements are involved.

2.1.2 Engineering

Many definitions of engineering can be found, they have in common that engineering is all about creation.

Walden et al, define engineering as "*the practice of creating and sustaining services, systems, devices, machines, structures, processes, and products to improve the quality of life—getting things done effectively and efficiently*" (Walden et al., 2015).

Kossiakoff et al. (2011) provide a dictionary description and define engineering as "*the application of scientific principles to practical ends; as the design, construction and operation of efficient and economical structures, equipment, and systems*".

An interesting definition is provided by Wulf (1998) as he states that engineering is "*design under constraint: by nature, by cost, by concerns of safety, reliability, environmental impact, manufacturability, maintainability, and many other such 'ilities'.*" It is unlikely engineering projects will start without requirements so that would confirm this definition. Interestingly, Wulf contradicts the definition provided by Kossiakoff et al. when it comes to science as he states that "*Engineering is not 'applied science'*". For the context of this research, however, this difference is not relevant. What is relevant, is that Wulf mentions constraints (like safety, maintainability and cost) and that makes his definition the most suitable in the context of this thesis as the mentioned constraints are very much recognizable from daily practice.

2.1.3 Scope

According to Kossiakoff et al. (2011), not everything created for people is complex enough to require systems engineering. Their definition of systems engineering is that it "guides the engineering of complex systems". According to them, the word "complex" herein restricts the concept of a system to "*systems in which the elements are diverse and have intricate relationships with one another*". This is an interesting view but not confirmed by other sources used for this research.

For the scope of systems engineering, it is important to know that systems engineering and software engineering are not the same discipline. They are not subsets of each other but rather "intertwined" (SEBoK, 2017a). However, one of the findings of a workshop "Exploring the Relationship between Systems Engineering and Software Engineering" at the 2015 Conference on Systems Engineering Research was that there are "*more similarities between systems engineers and software engineers in these regards than there are differences*" (Pyster et al., 2015). Because of the similarities of the two disciplines, some sources within the context of software engineering will also be used in this research.

2.2 Requirements in systems engineering

Creation of software and systems starts off with stakeholder needs, documented as requirements, most often at the early stages of the project. The definition of the term "requirement" is:

> "*a statement that identifies a product or process operational, functional, or design characteristic or constraint, which is unambiguous, testable or measurable, and necessary for product or process acceptability (by consumers or internal quality assurance guidelines)"* (Dick, Hull, & Jackson, 2017).

There is a difference between *product* and *process* requirements. The first are related to the properties of the system to be developed, the latter are related to how the development is done (methods, organisation). For this thesis, the scope is product requirements.

According to Walden et al. (2015), ideally "*the systems engineering process begins with a clear, complete set of user requirements and provides a disciplined approach to develop a system to meet these requirements*". However, later in their book, they go deeper into the technical processes of systems engineering and it becomes clear that creating requirements is a more iterative approach, as can be seen in Figure 4.

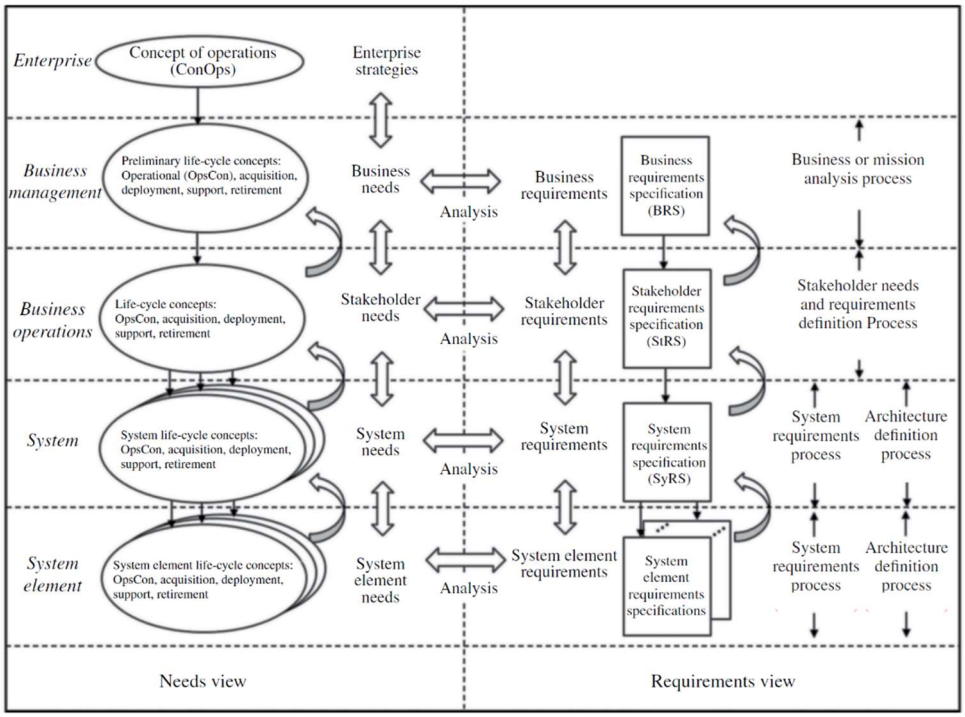

Figure 4: "Transformation of needs into requirements (Walden et al., 2015)"

The requirements engineering (or specification) process is concerned with defining "*the stakeholder requirements for a system that can provide the capabilities needed by users and other stakeholders in a defined environment*" (ISO/IEC, 2015).

For this thesis, we look at this process in a more linear and less elaborate way than shown in Figure 4 as the research scope is not the process of "engineering" requirements but rather how to assess the requirements produced by the process. Presenting the process in a more simplified view thus serves the purpose of introducing the research case and is not intended to substitute the requirements process of systems engineering.

In the process view for this thesis, requirements are engineered as part of a requirements engineering process and serve as input for the design / development process. The design / development process is concerned with creating a project deliverable (a system) based on the requirements, that is delivered to the verification/validation process. Usually, design and development are separate processes but, again, for the research case, they are combined in a single process as can be seen in Figure 5.

Figure 5: "From requirements to project verification / validation"

2.3 The preferred development model in systems engineering

Customers expect to see their needs fulfilled in the project deliverables. They also expect a high degree of predictability regarding those deliverables, the budget, lead time and quality (Collins & Baccarini, 2004). To deliver the best results within the context of the project, development teams use a development model that is best suited for this. The term "development" in the context of this research goes beyond software development and should be seen as the realization of systems, which could be software but also a bridge, a tunnel or a machine or any other engineered system.

Most projects use a structured approach to build create deliverables. This structured approach is also known as the System Development Lifecycle (SDLC) model. In software development, the "S" in SDLC refers to "Software". According to Suresh Kute & Thorat (2014), there are five development models:

1. The Linear Sequential Model (also known as "Waterfall" model)

2. The Incremental Model
3. The Prototype Model
4. The Spiral Model
5. The V-shaped Model (V-model)

Suresh Kute & Thorat do not mention the term “Agile”, but Agile is based on the idea of incremental and *iterative* development (Leau, Loo, Tham, & Tan, 2012). Interestingly, the Agile Manifesto principles (Beck et al., 2001) do not contain the term “iterative” but do mention “Deliver working software frequently”. Looking at the properties mentioned by Suresh Kute & Thorat on their incremental model (iterative in nature; focuses on the delivery of an operational product with each increment) we could share Agile under the Incremental model.

There seem to be plenty of (nearly religious) discussions on agile versus more traditional models (like Waterval and V-model) but research shows that there are many similarities (e.g. Palmquist, Lapham, Garcia-Miller, Chick, & Ozkaya, 2013) and that the “right” model depends on the organisation (e.g. Balaji & Murugaiyan, 2012). Each model has its own characteristics in dealing with requirements. Stable requirements is one of the project characteristics for choosing the V-model is (Balaji & Obaidy, 2016).

Projects concerning the development of complex systems often use a V-model approach (Komoto & Tomiyama, 2010; Scheithauer & Forsberg, 2013; Seyedhosseini & Keyghobadi, 2014). V-model stands for “Verification and Validation model” and is a sequential process, just like the well-known waterfall model: each phase must be completed before the next phase begins (Suresh Kute & Thorat, 2014).

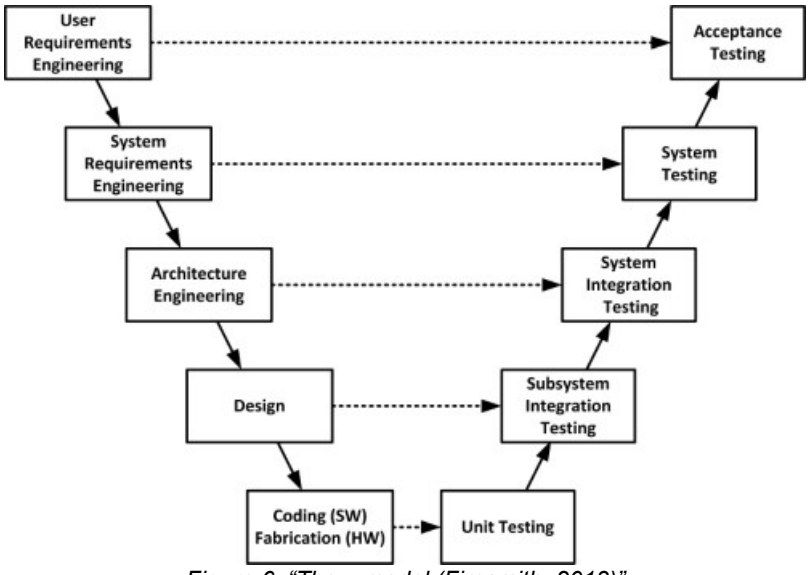

Figure 6: “The v-model (Firesmith, 2013)”

In the V-Model, there is a link between each phase of the development life cycle and its associated testing phase in order to “*improve efficiency and effectiveness of software development and reflect the relationship between test activities and development activities*” (Mathur & Malik, 2010).

The model was first introduced in the late 1980’s but has evolved over time. Various adaptations and improvements of the V-model exist (e.g. Ghanim, 2015; Mathur & Malik, 2010; Scheithauer & Forsberg, 2013; Yadav, 2012), even ones that combine agile practices (e.g. Anitha, Savio, & Mani, 2013; Mateen, Tabassum, & Rehan, 2017; Mc Hugh, Cawley, McCaffcry, Richardson, & Wang, 2013).

In the V-model, requirements engineering is at the very start of the “V”, as can be seen in Figure 6. There can be two phases concerning requirements:

1. **User requirements engineering**. *Define the stakeholder requirements for a system that can provide the capabilities needed by users and other stakeholders in a defined environment*. (Walden et al., 2015)

2. **System requirements engineering**. *Transform the stakeholder, user oriented view of desired capabilities into a technical view of a solution that meets the operational needs of the user*. (Walden et al., 2015)

However, in some V-model views, there is only one phase concerning requirements, being followed by a specifications phase (e.g. Balaji & Murugaiyan, 2012; Mathur & Malik, 2010).

In the Dutch civil engineering sector, the V-model is the preferred development model, as can be derived from the "Guideline for Systems Engineering within the civil engineering sector" (Rijkswaterstaat et al., 2013). As this thesis is within this domain, the V-model and its constraints on dealing with requirements will be the scope.

At some point in the "V", there is a transfer from the client to the contractor, in the "Guideline for Systems Engineering within the civil engineering sector" (Rijkswaterstaat et al., 2013) this is specified as a "*transfer point*" (see Figure 7).

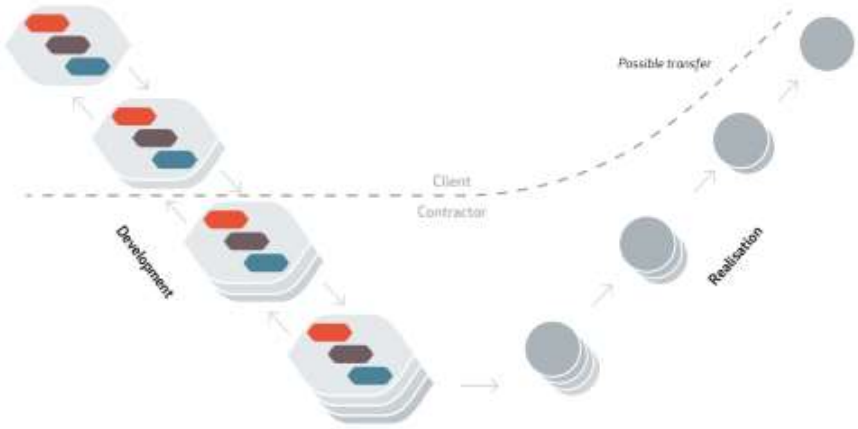

Figure 7: "Transfer point (Rijkswaterstaat et al., 2013)"

Within the context of this thesis, we assume a transfer point between the requirement engineering process and the design/development process. The client has engineered the requirements and is now transferring them to the contractor to start the design / development process.

2.4 Uncertainty in systems engineering

According to Thunnissen (2003), uncertainty "plays a critical role in the analysis for a wide and diverse set of fields from economics to engineering", its definition being "*something not definitely known or knowable*". The concept of uncertainty can be found in various fields, with different classifications.

For systems engineering, Thunnissen classifies uncertainty as show in Figure 8. He uses two classifications: the first a more theoretical one, the second a more practical one from a risk perspective.

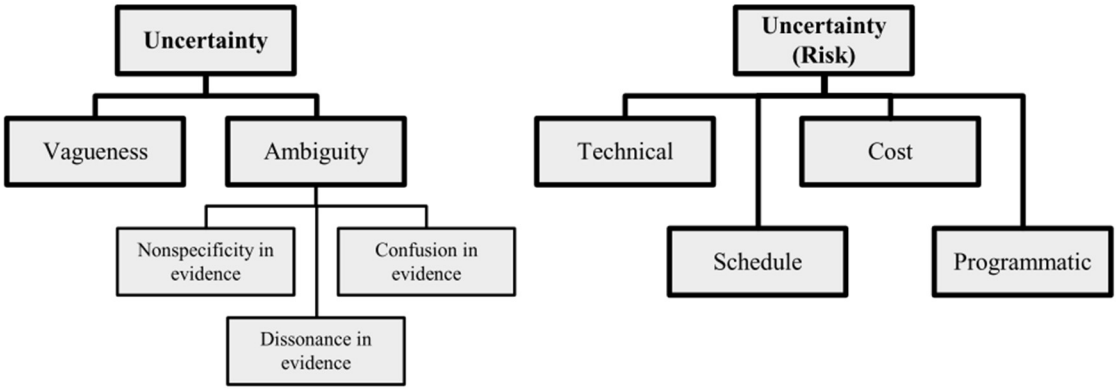

Figure 8: "Uncertainty classification in systems engineering (Daniel P Thunnissen, 2003)"

The classification is at an abstract level. The classification with the risk perspective identifies areas in which uncertainty manifest itself. All these areas are influenced by requirements, for example:

1. The technical solution depends on what the client's requirements of the system are.
2. The cost depends on the number of the requirements and how difficult these will be to implement.
3. The schedule depends on the number of requirements and how difficult these will be to implement.
4. The programmatic area (e.g., resources, contractual) depends on the priority of the requirements, their impact (lifesaving equipment or a commercial website) and how much effort is required to implement the requirements (size of the project).

Uncertainty in systems engineering can be caused by various factors, like tight schedule, budget constraints, unknown technology but also by *uncertain requirements*.

2.5 Requirements uncertainty

According to Nidumolu (1996), requirements uncertainty refers to "*the difference in the information necessary to identify user requirements and the amount of information possessed by the developers*". Stephenson, Attwood & McDermid (2011) classify requirements uncertainty as "*the phenomenon in which the requirement as stated is believed by the requirements reader not to be the requirement that is intended by the requirements writer*". In other words: there is a knowledge gap between (the people involved in) the requirements engineering and the design / development process.

Nidumolu (1996) identifies three dimensions of requirements uncertainty:

1. *Requirements instability*. The extent of changes in user requirements over the course of the project;
2. *Requirements diversity*. The extent to which users differ among themselves in their requirements;
3. *Requirements analysability*. The extent to which the process for converting user needs to a set of requirements specifications can be reduced to mechanical steps or objective procedures.

The second and third dimension manifest themselves at the requirements engineering process. Assuming not all uncertainty is reduced to zero before the aforementioned "transfer point" (see paragraph 2.3), at the transfer point we could speak about "*initial requirements uncertainty*"[1] for the contractor: requirements uncertainty that exists before the design / development process. This is illustrated in Figure 9.

The first dimension manifests itself during the design and development process in a project and is also referred to as "requirement volatility" (Nurmuliani, Zowghi, & Powell, 2004). Assuming that at the transfer point some measures will have been taken by the contractor to reduce uncertainty but that uncertainty cannot be reduced to zero, we could speak about "*residual requirements uncertainty*"[1] for the contractor: requirements uncertainty that remains after the assessment or that emerges during the design / development phase. This is illustrated in Figure 9.

[1] The terms "initial requirement uncertainty" and "residual requirements uncertainty" have been created for this thesis and are inspired by the use of the terms "initial performance risk" and "residual performance risk" in the article "Standardization, requirements uncertainty and software project performance" (Nidumolu, 1996). Walker et al. (2003) and Nolan et al. (2011) also use the term "residual uncertainty".

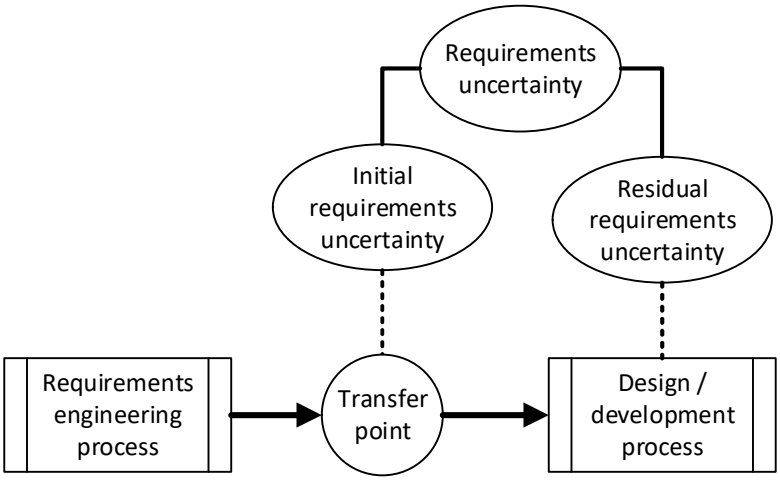

Figure 9: "Manifestations of requirements uncertainty in a project"

Most research on the concept of requirements uncertainty is not very recent, searching for papers on the concept of requirements uncertainty in papers from 2013 onwards only returns results concerning how to manage the uncertainty (e.g. Michalik, Keutel, & Mellis, 2014; Salay, Chechik, Horkoff, & Di Sandro, 2013). A reason for this lack of current research was not found, a possible explanation could be a need for more applied science but that should be subject to future research.

2.6 Negative effects of requirements uncertainty and how to deal with them

Large projects often suffer from the *uncertainty* of requirements (Ebert & De Man, 2005; Han & Huang, 2007; Schmidt, Lyytinen, & Mark Keil, 2001). Han & Huang (2007) even conclude that the "*requirement risk dimension is the principal factor affecting the project performance*". According to research by Nidumolu (1996) and later confirmed by Na, Li, Simpson, & Kim (2004) requirements uncertainty plays an important role on project performance. Although their research is focused on software development projects, due to the similarities with systems engineering (SEBoK, 2017a) we assume the same impact of requirements uncertainty on systems engineering. As identified in research by Ebert & De Man (2005), requirements uncertainty leads to changing requirement and insufficient functionality, which leads to project delays.

An uncertain requirement can also cause a ripple effect in a set of requirements, as can be concluded from a thesis on "Propagating and mitigating uncertainty in the design of complex multidisciplinary systems" (Daniel Pierre Thunnissen, 2005). Thunsissen mentions that requirements influence other requirements or design variables and provides an example of a requirement on the orbit of a spacecraft, which sets requirements for propulsion as well. Dealing with requirements uncertainty looks like a necessity.

To deal with the negative effects of requirements uncertainty on project performance during the design and development process (like rework), resources (people, budget) are required. These resources, however, are often scarce, as identified in an exploratory study of 30 leading firms on portfolio management (Cooper, Edgett, & Kleinschmidt, 2006). This means that requirements need to be assessed and their initial uncertainty reduced as much as possible/feasible, before accepting them for the design and development process, to limit the risks on project performance.

In a transfer from client to contractor, there is usually a process to assess requirements, like the "System Requirements Review" (SRR) from the MIL-STD-15218 standard (Department of Defense, 1985) which is referenced by Rijkswaterstaat for tunnel projects (e.g. Rijkswaterstaat, 2017). The SRR is "*A formal, system-level review conducted to ensure that system requirements have been completely and properly identified and that a mutual understanding between the government and contractors exists*" (Defense Acquisition University, 2017).

So, what we see in literature is that a requirement assessment process is concerned with what to do with a requirement:

1. Accept it for the design/development process? If the requirement is fully understood and the contractor certain they can design and build the deliverable, the requirement can be accepted.
2. Refuse it and send the requirement back to the requirement engineering process? If the requirement is not understood or leaves too much uncertainty, the requirement should be improved or perhaps even dropped.
3. Take mitigating actions to reduce the requirements uncertainty? This could be in the form of workshops or prototyping, taking up resources.

We can visualise this as follows:

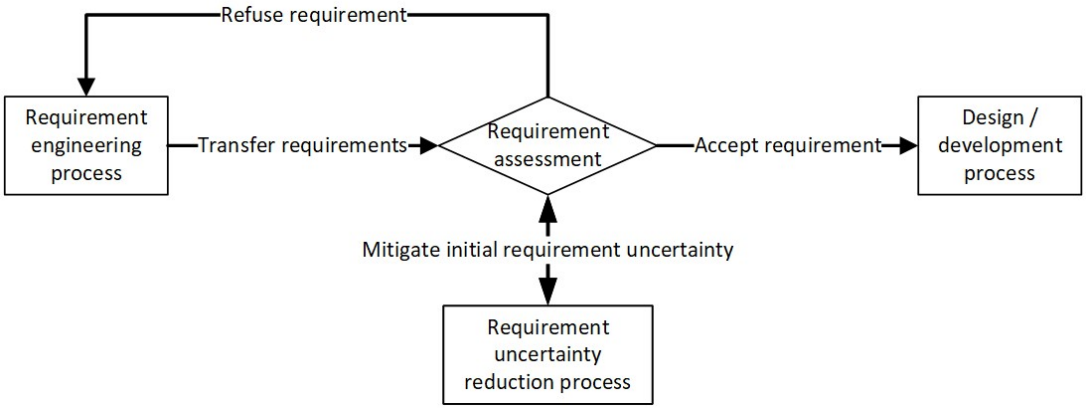

Figure 10: "Requirement assessment process"

The goal is to reduce uncertainty to an acceptable level, complete certainty is not likely to be achievable nor affordable (as was mentioned before on the scarce resources). The assumption is that the acceptable level of requirements uncertainty also varies. It would be likely, for example, that acceptable uncertainty of a requirement concerning an insulin pump would be much lower than acceptable uncertainty of a requirement concerning a customer contact form.

2.7 Problem statement

Although there is a process that can be used in the V-model to assess requirements, there is no *objective* list of criteria that should be considered in this process. The decision if a requirement should be accepted or not seems to be based solely on *subjective* criteria like experience, gut feeling or commercial opportunism. Not having objective criteria means that the outcome of the assessment process becomes unpredictable, which could leave the contractor at risk when uncertain requirements are accepted and cause major rework in the design or development phase. Some examples have been given in the introduction of this thesis, but further company specific examples are not allowed to be shared. Some famous examples can be found in the journal articles "Requirements development, verification, and validation exhibited in famous failures" (Terry Bahill & Henderson, 2005) and "Why do Projects Fail?" (Alami, 2016).

In the assessment process, the outcome might be that mitigating measures are required to reduce the requirements uncertainty. Again, it's not clear on the basis of what criteria this decision is taken. Also, before taking mitigating actions to reduce requirements uncertainty , a question like the one asked by Regnell, Svensson and Wnuk (2008) is relevant: "*What level of uncertainty and degree of approximation can we tolerate*"? Some research (eg. Feather, Cornford, Hicks, Kiper, & Menzies, 2008; Nolan, Abrahão, Clements, & Pickard, 2011) mention that, due to the scarce resources, there must be some trade-off between the uncertainty reduction and the mitigating effort required to achieving this reduction. However, no general acceptable uncertainty level seems to exist.

In order to determine what the level of uncertainty is before the assessment, or to measure the effect of reducing uncertainty, an understanding on how to actually measure uncertainty is important. However, there is no standard measurement method or a standard scale for uncertainty. If uncertainty cannot be quantified, there is a risk that too much or too little effort will be spend in reducing uncertainty.

2.8 Research question

The research objective is to produce a list of validated criteria for the requirement assessment that will help to reduce requirements uncertainty and thereby reduce the negative impact of requirements uncertainty on project performance.

The research question is:

> *RQ1: Which objective assessment criteria are relevant for determining if a requirement should be accepted into the design and development phase?*

To measure the effect of reducing requirements uncertainty and to make sure not too much or too little effort is put into the process, the following sub question will also be researched:

> *RQ1.1: Which methods are relevant to measure initial requirements uncertainty?*

As some research indicate that there must be some trade-off between the uncertainty reduction and the mitigating effort required to achieving this reduction (see chapter 2.7), this leads to the following sub question:

> *RQ1.2: Which factors that influence how much initial requirements uncertainty is acceptable are relevant?*

This sub question is not about determining an exact level of uncertainty, this would be subject to future research.

One of the findings in the initial literature review related to RQ1 is that requirements are volatile (Nurmuliani et al., 2004), meaning that they can (or will) change during a project: the aforementioned "residual uncertainty" (see paragraph 2.4). This is recognized as a risk factor for projects (Schmidt et al., 2001). So even if there are objective criteria in the requirement assessment that help reduce initial requirements uncertainty, residual requirements uncertainty could also have an impact on a project. Dealing with this type of uncertainty, however, is outside the boundaries of this research.

2.9 Conceptual model

Based on RQ1, RQ1.1 and RQ1.2, a conceptual model can be constructed, see Figure 11. In the conceptual model, the concepts from the research questions return:

- Assessment criteria[2]. These are the criteria to be used in the requirements assessment to reduce the initial requirements uncertainty.
- Requirements uncertainty measurement methods. Methods to measure the level of uncertainty, to help understand the current level or to make a decision what level of uncertainty is acceptable.
- Acceptable uncertainty influencing factors. These factors influence how much initial requirements uncertainty is acceptable for the design / development phase.

The initial requirements uncertainty is the uncertainty that exists at the transfer point and that can be reduced by the requirements assessment process. Requirements uncertainty has a negative impact on project performance. Theoretically, the uncertainty could be zero and, in that case, would not have a negative impact, but zero uncertainty is highly unlikely (see paragraph 2.6).

[2] The term "acceptance criteria" was also considered but this term is already in use for another concept: a set of conditions that is required to be met before deliverables are accepted (Project Management Institute, 2013, p. 526). Requirements are used to define what needs to be build, acceptance criteria define when the deliverables based on the requirements is good enough.

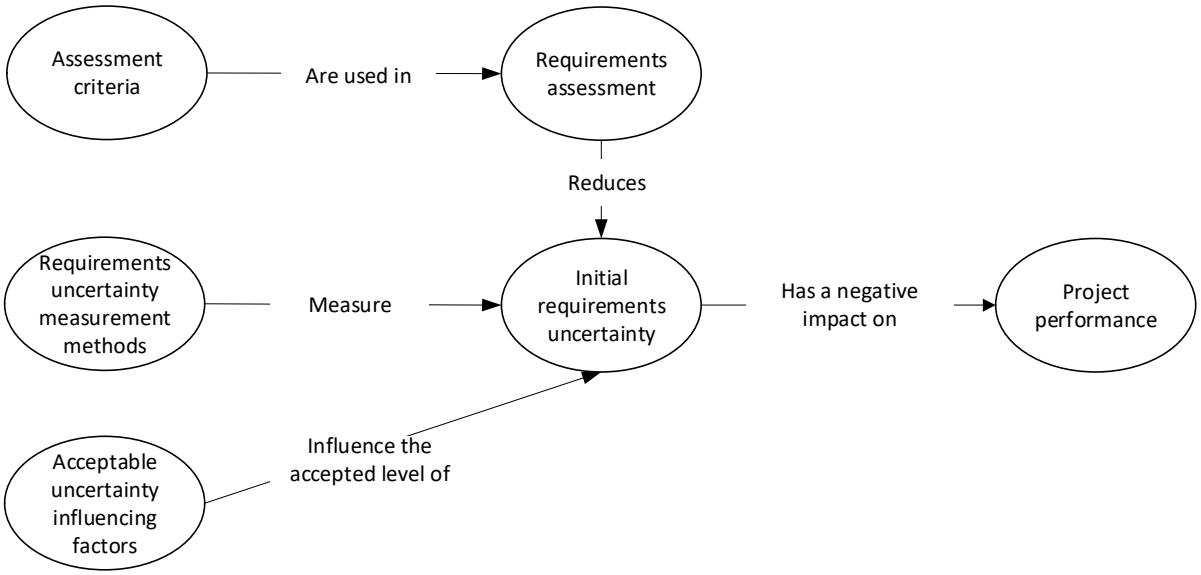

Figure 11: "Conceptual model"

2.10 Research context

This research will be done within a systems engineering context in the civil engineering sector in The Netherlands for projects that use the V-model project methodology.

3 Research method

This chapter describes the research method that is used for this research. First, the outline of the research process is defined as well as the rationale of the choices made herein. The main stages of the research are then explained in more detail in the following paragraphs.

3.1 Research outline

The research is about finding criteria for the requirement assessment, gaining insight in the factors that determine what uncertainty is acceptable and how uncertainty can be measured.

The research follows the process as outlined below:

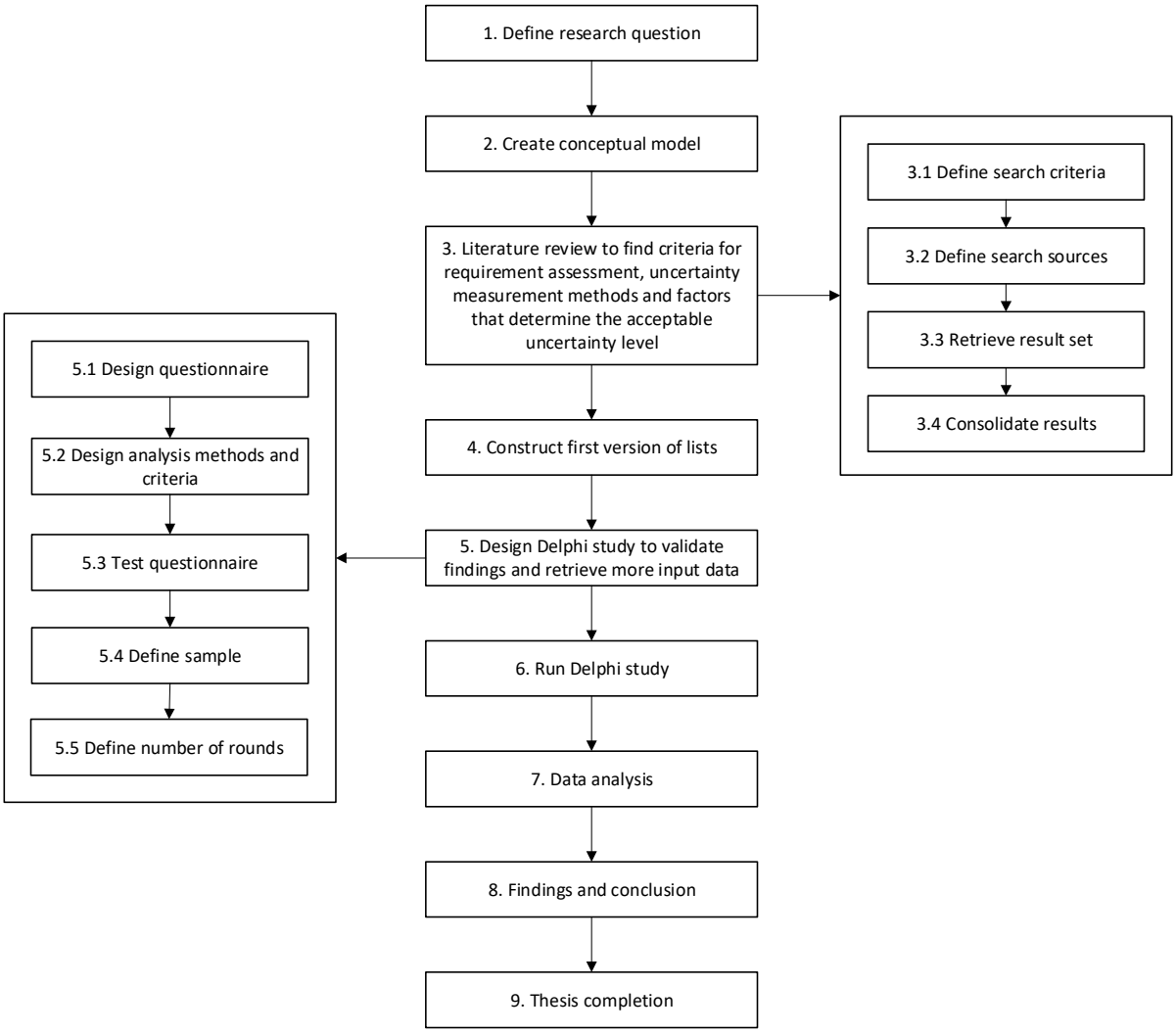

Figure 12: research process design

A literature study followed by a Delphi study are the core of the research. In the literature study, initial information will be found after which a panel of experts in the field of systems engineering will be asked to validate the findings.

3.2 Literature study

The literature study is used to find data on the research questions and to construct an initial list of criteria, factors and measurement types.

3.2.1 Source selection

To find relevant literature for the research, Google Scholar (GS) was used. GS is a freely accessible search engine for scientific papers and contains information from many, though likely not all, repositories. The main reasons for using GS are for this research are:

- Available to everybody from any location with internet access, making the search results transparent and repeatable.
- Contains sources from a wide variety of repositories, making it a single point of access to articles instead of having to use a variety of search engines

A recent literature study, however, warns that GS can be “easily manipulated and its indexing quality still remains a challenge” (Halevi, Moed, & Bar-Ilan, 2017). Therefore, as a control measure, some queries for this research were executed in the AIS eLibrary and the ACM Digital Library as well as these sources contain “research papers and journal articles relevant to the information systems academic community” (Elsevier Inc., n.d.) and a “database focused exclusively on the field of computing” (Association for Computing Machinery, n.d.). The result set of the queries was compared against the result set from GS to verify if GS provided sufficient representative search results.

Table 1: "Sample of query results with the AIS eLibrary and the ACM Digital Library"

ID of query used for the literature review (see paragraph 3.2.5)	# results in AIS with this query	# results in ACM with this query	Findings of the comparison against GS
RQ1_2	0	5	No additional sources to include compared to the results of the query in GS
RQ1_4	1	0	No additional sources to include compared to the results of the query in GS
RQ1_8	87	44	No additional sources to include compared to the results of the query in GS

Based on the samples mentioned in Table 1, it was concluded that GS provided sufficient representative search results.

3.2.2 Literature selection process

A literature review was done on papers searched with GS, no year restrictions, sorted by relevance. Input for the search were various queries that were constructed using pre-defined search terms.

The following criteria were used for selecting articles for review:

1. Patents and citations were discarded.
2. Only scientific articles were considered for review (also as part of books). Books where all chapters are written by the same authors and presentations were discarded as these are not independently reviewed or contain research findings.
3. Non-English language articles were discarded
4. Articles that were not fully accessible online, publicly or via HU institute, were discarded. Although some repositories (like the IEEE Xplore Digital Library) offer paid access to articles that might be relevant, it is not achievable to purchase them for research, they might even turn out to be irrelevant.

The result set for each query was exported to Excel by using the citation tool “Publish or Perish” (Harzing, 2018). This tool was limited to maximum 1000 results per query to avoid Google Scholar blocking queries due to excessive usage of its service. Even with this maximum, no more than 980 results per query could be retrieved.

Processing the results of each query consisted of two rounds:

1. Review of the result set of GS. When there were more than 100 search results, only the first 100 results were included for review. When the query consisted of more than one OR clause, an extra 100 results were reviewed with a maximum of 500.

2. A review of the result set in Publish or Perish where results were ordered by number of citations (ranked high to low). Papers that were not within the first one hundred results of round 1 but with an h-index (Hirsch, 2005) higher than the h-index of the whole result set, were reviewed.

Each review consisted of a maximum of three stages:

1. Scrutinizing titles and summaries in the result page. Articles whose title did not indicate relevance, whose summary was not relevant to the topic or did not contain the search terms were discarded.
2. Initial analysis of the article content. The abstract, conceptual model (if included), discussion and conclusion in the article were scanned first. Non-relevant articles were discarded.
3. Full analysis of the article content. If the initial analysis looked promising, the whole article was reviewed. Non-relevant articles were discarded.

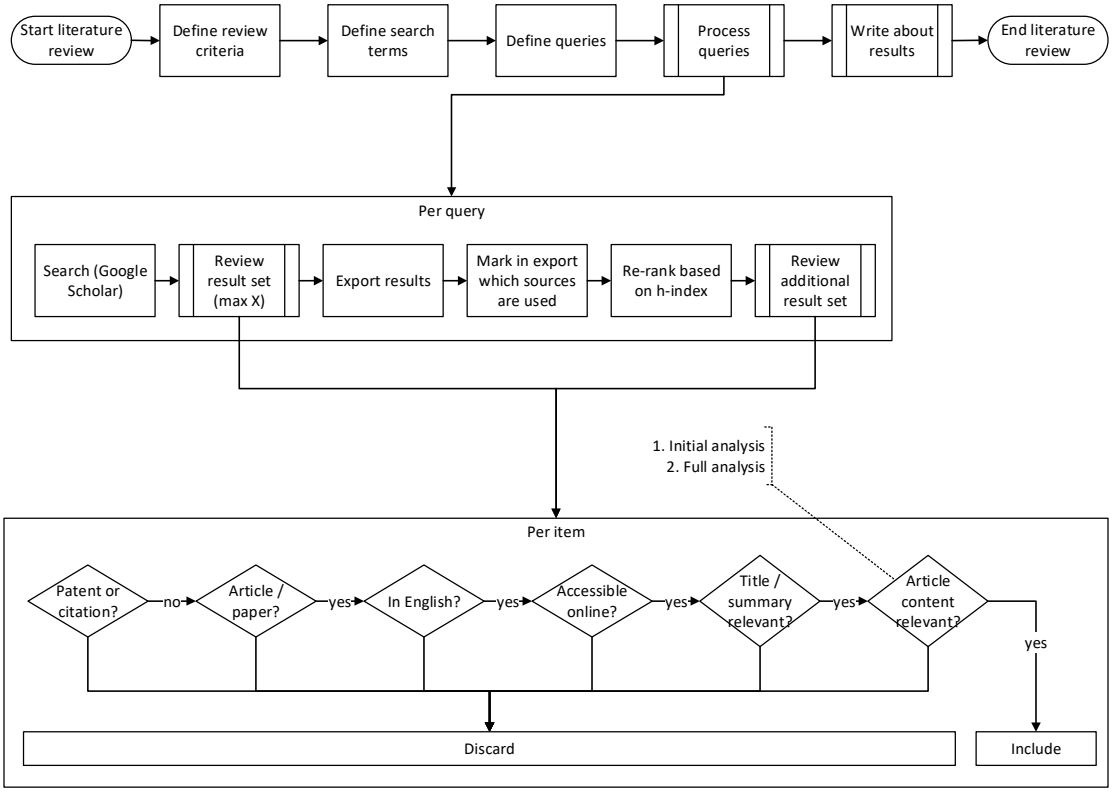

Figure 13: "Literature research process"

Additional sources can be include based on the literature review:

1. When the article refers another article that can provide additional information, that article can be included
2. When information found in an article requires additional information that cannot be found in the result set, a search will be done for additional information.

3.2.3 Indication of article relevance

Not all articles are relevant to the research. In the review stages, relevance was determined by looking at some article properties or content. Only when all criteria were met, the article was marked as relevant:

Table 2: "Determining article relevance"

Property	Criteria
Title	• Contains one or more of the search terms, or • Topic is in the area of systems or software engineering
Summary* **/ abstract**	• Contains one or more of the exact search terms, or

Property	Criteria
	• Combined search terms (like “requirements uncertainty”) are not separated by a comma in the text, or • Indicates coverage of research question • Is not about requirements uncertainty (and related processes) within project phases before or after the assessment • Is not about proving the effect of requirements uncertainty on project performance. This is already covered in the introduction and used as starting point for this research. • Is not about uncertainty in agile projects (this research has the V-model as its context)
discussion / conclusion	• Indicate coverage of the research question, and • Is not indicating the article is about requirements uncertainty (and related processes) within project phases before or after the assessment, and • Is not indicating the article is about proving the effect of requirements uncertainty on project performance. This is already covered in the introduction and used as starting point for this research.
Conceptual model	• Contains one or more of the search terms, or • Covers (parts of) the research question
Full text	• Provides answers to the research question

* Summary of the item in the result set in GS

3.2.4 Search terms

To find relevant articles with GS, queries with appropriate search terms needed to be constructed. The search terms were collected in a mind map format and grouped per research question. Additional search terms were added when:

- The query did not supply relevant search results
- A new relevant term was found in the literature review

This resulted in the search term map as shown in Figure 14.

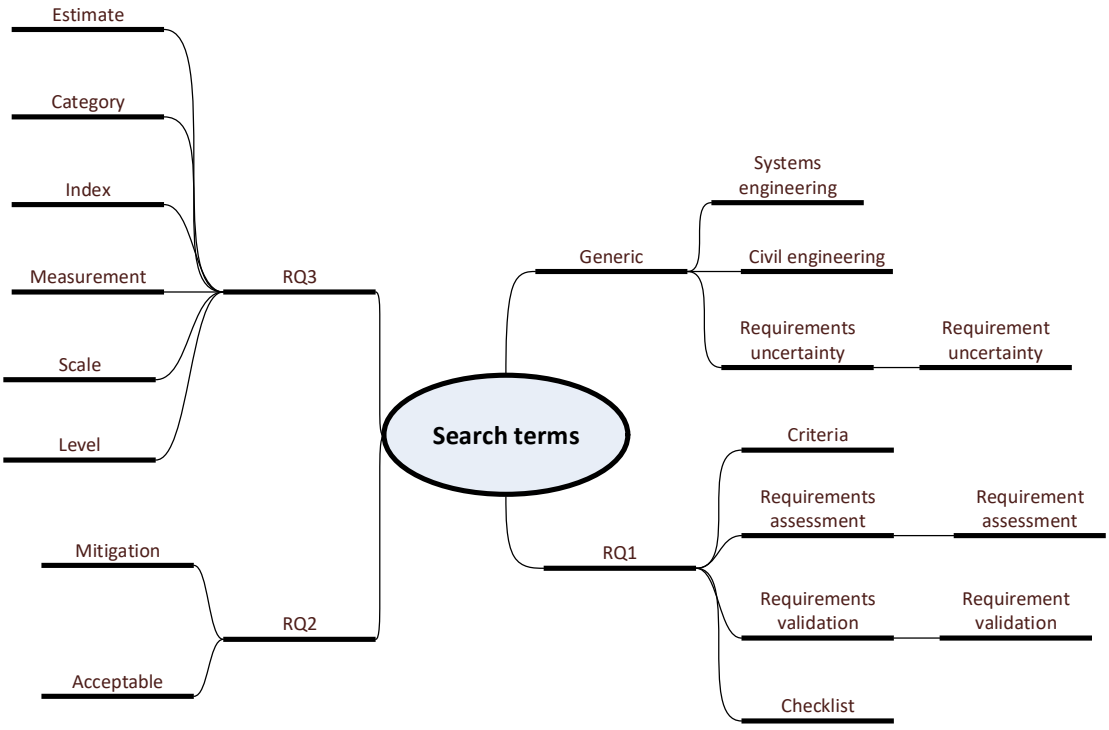

Figure 14: "Search terms"

3.2.5 Search queries and results

For RQ1, the following search queries were used:

Table 3: "Search queries RQ1"

#	Search queries	Search results	Included after round 1	Included after round 2
RQ1_1	criteria "systems engineering" "requirement assessment" OR "requirements assessment"	434 (100)	1	1
RQ1_2	"civil engineering" "systems engineering" "requirement uncertainty" OR "requirements uncertainty"	26	0	0
RQ1_3	"systems engineering" "requirement uncertainty" OR "requirements uncertainty"	437 (100)	1	1
RQ1_4	"systems engineering" "civil engineering" "requirements validation" OR "requirement validation"	84	1	1
RQ1_5	"systems engineering" "requirements validation" OR "requirement validation"	2780 (100)	1	1
RQ1_6	"systems engineering" "requirements validation checklist" OR "requirement validation checklist"	6	1	1
RQ1_7	"systems engineering" "requirements validation criteria" OR "requirement validation criteria"	15	0	0
RQ1_8	"requirements uncertainty" OR "requirement uncertainty"	2540 (100)	2	2
	Totals	**6296 (531)**	**7**	**7**

Searching for relevant articles with query RQ1_1 only retrieved one useable source. Search queries with the search term "civil engineering" did not retrieve relevant results, the term was therefore omitted in later queries.

For RQ1.1, the following search terms are used:

Table 4: "Search queries RQ1.1"

#	Search queries	Search results	Included after round 1	Included after round 2
RQ1.1_1	"requirements uncertainty" category OR index OR measurement OR scale OR level OR estimate	2060 (500)	6	6
RQ1.1_2	"requirement uncertainty" category OR index OR measurement OR scale OR level OR estimate	603 (500)	0	0
	Totals	**2663 (1000)**	**6**	**6**

For RQ1.2, the following search queries were used:

Table 5: "Search queries RQ1.2"

#	Search queries	Search results	Included after round 1	Included after round 2
RQ1.2_1	mitigation acceptable "systems engineering" "requirements uncertainty" OR "requirement uncertainty"	119 (100)	0	0

#	Search queries	Search results	Included after round 1	Included after round 2
RQ1.2_2	mitigation "systems engineering" "requirements uncertainty" OR "requirement uncertainty"	215 (100)	1	1
RQ1.2_3	acceptable "systems engineering" "requirements uncertainty" OR "requirement uncertainty"	191 (100)	1	1
RQ1.2_4	"systems engineering" "requirements uncertainty" OR "requirement uncertainty"	438 (100)	0	0
	Totals	**955 (400)**	**2**	**2**

3.2.6 Content analysis

It is expected that the literature study reveals assessment criteria but that these appear in various shapes and formats and can fall into different groups. To interpret the information found, practices from an inductive “qualitative content analysis” approach will be used (Bryman & Bell, 2015; Elo & Kyngäs, 2008). In this approach, themes or categories are distilled from the articles by marking pieces of text that contain important information and then finding common concepts within these markings. This is an iterative process; the list of themes or categories will be constantly revised.

The findings will be processed as follows:

- **Listing the criteria found.** Assessment criteria found in the literature review will be identified using C_n, where $_n$ is a sequence number starting at 1.
- **Rephrasing**. The assessment criteria will be rephrased into a statement that can be answered by “yes” or “no”. This will make the criteria uniform and suitable for use in a checklist.
- **Identifying categories**. It is expected that the literature study reveals assessment criteria that can fall into a set of categories. Qualitative content analysis will be used for this.
- **Adding remarks**. The remarks are used to clarify the processing of the item or to give further explanation on the item
- **Removing duplicates**. Duplicate items will be merged into one. Merged assessment criteria will be identified by CP_n where $_n$ is a sequence number also starting at 1.

A similar approach will be used for measurement methods and influencing factors.

3.3 Delphi study

The Delphi study is an iterative multistage process designed to combine opinion into group consensus (Hasson, Keeney, & McKenna, 2000). For this research, a standard design is used (Brady, 2015):

1. A first round with a questionnaire developed from the literature
2. A second round to allow participants to provide feedback on all responses from Round 1.
3. Finally, a third round to reach a final consensus on a given topic

The method is used to validate and expand the findings from the literature study by a panel of experts in the field of systems engineering who will be asked to:

1. Give feedback on the criteria found
2. Give feedback on the factors that influence how much uncertainty they accept
3. Give feedback on how uncertainty can be measured.

For each topic they will be asked to:

1. Specify if the item (criterion/factor/measurement type) is relevant or not
2. Specify the importance of the item
3. Add additional items, if they have any

The Delphi study combines both qualitative and quantitative research components:

- Quantitative: providing statistical data on how the items are scored on relevance.
- Qualitative: adding additional items and using content analysis to guide the expert panel for the next round and finally to improve the criteria, measurement methods and influencing factors.

3.3.1 Rationale

Based on initial review remarks and a review of papers, it was concluded that the Delphi method was the most suitable method for this research. Okoli & Pawlowski (2004), for example, created a comparison between the traditional survey and the Delphi method. Some of their most relevant findings on the Delphi method are:

- The sample size can be much smaller
- The number of individual responses is much higher
- The method can employ further construct validation
- It results in richer data

De Bruin & Rosemann (2007) also gathered various motivations for using the Delphi method. The finding that "anonymity leads to more creative outcomes and adds richness to data" would be especially interesting in the context of this research. Finally, Skulmoski, Hartman, & Krahn (2007) conclude that the Delphi method would be a suitable candidate for research that focusses on problems, opportunities, solutions and forecasts as it is:

- Well suited when there is incomplete knowledge about phenomena
- Not just a quantitative method, but works very well in qualitative research

The use of a traditional survey was considered to gather input on the criteria, factors and measurements found in literature. However, there a few challenges regarding using a traditional survey for this research:

- The group of experts on the subject (performing requirements analysis in the field of systems engineering in the civil engineering sector in The Netherlands) is limited but presumably still large enough to make it difficult to get the required sample size. Based on an estimated population (N) of 500 and a precision (e) of 5%, calculating the required sample size (n) based on a proportion formula (Israel, 1992) indicates that a sample size of 222 experts would be required:

$$n = \frac{N}{1 + N(e)^2} = \frac{500}{1 + 500(.05)^2} = 222$$

  Similar results were given by using online sample calculators (e.g. Australian Bureau of Statistics, 2018). As these experts work at companies that are also competitors, getting access to the required sample was not possible for this research.
- There is not much information on the topic of requirement assessment criteria, consensus from experts is sought but in a traditional survey there is no interaction between the experts.
- Input from the experts might lead to new insights that cannot be verified in a single survey so more than one round would be required to validate new input.

Due to the strengths of the Delphi method and the limitations of the traditional survey in the context of this research, the Delphi method surfaced as the preferred research method.

3.3.2 Sample

In the Delphi study, questionnaires will be sent to a sample of professionals working in the field of systems engineering in the civil engineering area. Participants are sought in the context of Vialis (part of the VolkerWessels group) projects as this domain is accessible by the researcher but also representative for the civil engineering sector in The Netherlands. Participants are sought with a role that concerns requirements assessments:

- Project/integration/design managers

- (Lead) engineers/architects
- (Lead) test engineers

Only professionals working for Vialis (including contractors) will be included, Rijkswaterstaat employees are excluded as they are on the requirements engineering side of the process.

A group of thirteen professionals were asked to participate, eleven agreed. In Table 6, an overview of the sample with their roles and years of experience is provided. Years of experience is categorised to make sure the individual participant cannot be traced (also see paragraph 3.3.8 for ethical considerations).

Table 6: "Sample for the Delphi study"

ID	Role	Years of relevant experience
1	Systems architect	2-5 years
2	Systems architect	10+ years
3	Systems architect	10+ years
4	Systems architect	10+ years
5	Systems architect	10+ years
6	Lead test engineer	10+ years
7	Systems architect	10+ years
8	Design manager	10+ years
9	Systems architect	1-2 years
10	Systems architect	10+ years
11	Design manager	10+ years

This seems a rather small group, but because it is very homogeneous (within the narrow field of systems engineering in the civil engineering area in The Netherlands), a group size between ten and fifteen should yield sufficient results (Okoli & Pawlowski, 2004; Skulmoski et al., 2007).

3.3.3 Consensus

The goal of the Delphi rounds is to obtain consensus on the opinion of experts (Hasson et al., 2000). As found in literature by Hasson et al. (2000), there are various views on what defines the level of consensus required, varying from 51% to 80% to defining consensus as the stability of responses through a series of rounds. For this thesis, a level of larger than 75% will be used, the rationale being that for the criteria, measurement methods and influencing factors to be useful within a corporate environment, a large majority needs to agree on their relevance,

For this thesis, consensus is determined by calculating the variance of the Relevance score per item. Variance is an indication of the spread of scores; if they are far apart, the variance will be high and vice versa. The variance is easy to calculate on the data from the questionnaires and therefor suitable to use for this thesis.

A variance equal to or lower than 0.9 will be considered to indicate consensus, this corresponds to a percentage of 77.5% or more using a maximum variance of 4[3] using the following formula:

$$ConsensusLevel = 100 - \big((100 \div 4) \times 0.9\big) = 77.5\%$$

This percentage falls within the range of levels from previous research, as mentioned above.

[3] The maximum variance with eleven participants giving a score from 1 to 5 is 3.966. This is rounded to 4 for the calculation of the percentage.

3.3.4 Design of the rounds

In Delphi study, the participants will be presented the results from the literature study and are asked to:

- validate them (Likert scale and open question)
- suggest additional items (open questions)

Each round is expected to run for a maximum of two weeks. In the first week, the participants fill in the questionnaire, the second week is used for data analysis and preparing the next round.

The Delphi study will run for a maximum number of three rounds. This limit is introduced for the following reasons:

- There is a time constraint of three months maximum on this research, including literature review, analysis and writing.
- It is expected the sample will reach consensus within three rounds
- There is a limited amount of time resources for this research can be claimed
- Research indicates that two to three rounds are preferred, as found in literature by Hasson et al. (2000)

The rounds are designed as follows:

Round 1
In this round, the initial list of criteria, factors and measurements found in the literature review will be presented to the expert panel. The participants evaluate them and provide their own.

Round 2
In this round, the results of round 1, including the new criteria, factors and measurement provided by the participants will be presented as well as statistical information. The participants are asked to evaluate the results and to review the new items. The new items will be marked as such to make them recognisable. No new items can be added in this round.

Round 3
In this round, the results of the review of round 2 will be presented. Items on which consensus has been reached in Round 2, will not be presented again. Comments from the first two rounds will be used, where applicable, to update the criteria, measurement methods and influencing factors. The participants will be asked to do a final review to reach consensus on the remaining items.

The results of round 3 will be the input for the final analysis of this research.

3.3.5 Questionnaire design

The questionnaire is split in four parts:

1. Participant identification
2. The assessment criteria
3. Measurements of uncertainty
4. The factors influencing how much uncertainty is acceptable

The questionnaire is rather complex and various types of input are asked on each item. This level of complexity is not supported by online questionnaire tools like Google Forms or SurveyMonkey, so an alternative had to be found. There are advantages and disadvantages of using Excel:

Disadvantages:
- Complex to design
- Needs to be secured so no change of formatting is allowed
- Input checks are more difficult to implement
- Hard to combine the results

Advantages:

- Allows complex design that meets research goals
- Does not require online access
- Can be stored for archiving in folders
- No additional costs

The fact that Excel allows a complex design is the decisive reason to use it.

With the questionnaire, an instruction will be provided to make sure the participants fill in the questionnaire completely with the right kind of data. For each research question, background information will be provided to help the participants understand the research. This information is created based on the findings for each research question in the literature review.

In the questionnaire, the participants will be presented with:

- The ID of the item
- Title or short description
- Any category, label or type found (as mentioned in paragraph 3.2.6)
- The research remarks

The participants will be asked to provide feedback on the items by providing:

- Relevance. The participant can indicate the relevance of the item found. This is measured on a scale from 1-5:
  - 1=Not relevant
  - 2=Slightly relevant
  - 3=Fairly relevant
  - 4=Relevant
  - 5=Highly relevant
- Any input on the category, label or type
- Comments. This input has a maximum of 500 characters per item to keep the input manageable within the boundaries of the research.

Asking for comments serves two purposes:

1. A means to influence the other participants as they do not know who is participating;
2. A source for new research insights.

In Round 2 and 3, all participant input from the previous round will be part of the questionnaire so all participants can see what the opinion of the others is regarding an item.

In the first round, participants are allowed to add new items. These will then be presented in the round 2 for review by the other participants. In the last round, participants will not be asked to provide comments.

During the rounds, the questionnaire design might be changed to accommodate new types of questions that arise from the results of the rounds.

3.3.6 Data analysis

Statistical analysis using SPSS (IBM, 2017) will be used to identify the relevance of the criteria, measurement methods and influencing factors. Variables can be found in Appendix H , Table 37, Table 38 and Table 39.

Statistical analysis will be done using the following values:

- **Median**. The median is used for ranking the items on relevance as the relevance takes the shape of Likert-type item (Boone & Boone, 2012). The advantage of the median is that it will give a better representation of the overall relevance when the input is significantly skewed.

Items with a median of at least 4 will be regarded as relevant in the context of the research question. Instead of median, a mean could also be used, even if the relevance takes the shape of a Likert-type item but for analysis the median is the better option. See paragraph 3.3.6.1 for the rationale.

- **Variance**. Variance is used to determine if consensus has been reached. See paragraph 3.3.3 for details.
- **Range**. The range is used to indicate the difference between the highest and lowest relevance score and will be used to give the participants an idea of how aligned they are in their opinions.
- **Minimum and Maximum.** These values will be used go get an idea of the item is scored on relevance and are used for calculating the range.

Missing values will not be recoded and are excluded from the analysis. For the statistical analysis, only the relevance score is important.

Content analysis on the comments will be used to identify general findings, findings that are not related to the criterion, measurement method or influencing factor itself. Remarks from the rounds could lead to new questions or background information required to fill out the questionnaire.

3.3.6.1 Mean versus Median controversy

Using a mean for Likert-type items is a controversial topic in statistics as the data is regarded as an ordinal measure (Clason & Dormody, 1994; Sullivan & Artino Jr, 2013). However, as Sullivan & Artino (2013) conclude after research, even with Likert-type items it is possible to use parametric tests, which was also concluded before by Norman (2010). On the other hand, when looking at the research of Boone & Boone (2012), the questions in the questionnaire are Likert-type items and the "Central Tendency" should therefore be analysed by the median, not the mean.

A mean (average) is a more common concept than the median but as research seems to indicate that the mean is allowed but not preferred, for the statistical analysis of data from Delphi rounds the median will be used.

3.3.7 Test round

The first version of the questionnaire and the instruction will be reviewed by a test engineer who is not part of the sample. The goal of this review is to make sure that:

- The questionnaire can be understood by the participants
- The results can be properly analysed
- The instruction is easy to understand

3.3.8 Ethical considerations

Participation in the study will be anonymous to everyone but the researcher. All participants will be given the following information regarding confidentiality:

- The organisations where the participants are active, have been asked for possible suitable candidates but have not been informed which candidates actually participate.
- Input will not be traceable to the individual participant. Data will be anonymized, a list where participant names are linked to returned questionnaires will be held outside the publication of the research.
- The participants are approached individually and not as a group.
- Should a review committee request insight in participation details to evaluate the scientific integrity of the research, the participants approval will be asked before exposing their names.

The participants will be asked not to discuss the research with others during the Delphi research so they "*can present and react to ideas unbiased by the identities and pressures of others*" (Hasson et al., 2000).

4 Literature Review

This chapter contains the results of a literature study on the topics concerning the research questions. The selection process and criteria used to retrieve the relevant literature are covered in detail in paragraph 3.2.

Information on the background topics (systems engineering, requirements, V-model and requirements uncertainty) of this research has already been covered in paragraph 2 and will not be further reviewed here.

4.1 Requirements assessment criteria

This paragraph includes a literature review on RQ1. A total of seven relevant articles were found to identify assessment criteria. A brief summary of each article is provided as well as the number of criteria found in the article. The articles were scanned for possible categories as well (see paragraph 3.2.6). In chapter 5, the overall findings of the literature review will be discussed.

4.1.1 Ebert & De Man

In the requirement assessment, the requirements are evaluated by looking at various criteria. Ebert & De Man (2005) start their article with the quote "*Requirements are uncertain*" (Ebert & De Man, 2005, p. 553) and identify root causes and "early project symptoms" of requirements uncertainty that lead to symptoms of uncertainty.

One of their findings is that requirements should be evaluated from an external *perspective*, with a business case view, and name four processes they selected to reduce the amount of uncertainty in the project early in the process. This cannot be mapped directly onto the requirement review process as their goal is to *prevent* uncertainty, which is relevant *before* the assessment takes place. However, the symptoms they mention can be rephrased into assessment criteria.

A total of four criteria were found in the article, see Appendix D Table 26. The mentioned "perspective" is an indication of a label that can be applied (see paragraph 3.2.6). We will include "Perspective" as a label and will try to apply this concept for the results from the other articles. If none can be found, this label will be discarded.

4.1.2 Kamalrudin & Sidek

In the context of software engineering, Kamalrudin & Sidek (2015) state that requirements validation "*ensures the correctness, completeness and consistency of a requirement*" (Kamalrudin & Sidek, 2015, p. 1). They call these "types of requirement quality" and provide definitions for them. From this list of definitions, assessment criteria can be distilled. Not all definitions they describe are requirement assessment criteria as in some cases, the item concerns another deliverable (like "program", SRS" or "requirement's document").

A limitation of this article for this research is that it that its context is software engineering and not systems engineering. As there are similarities between these fields, the criteria mentioned in the article are added initially but might be discarded when coding the total result set when they are too far apart from the other criteria found or too specific software related.

A total of eleven criteria were found in the article, see Appendix D Table 27; it was possible to apply the label "Perspective" to the criteria. The mentioned "types of requirement quality" is an indication of a label that can be applied (see paragraph 3.2.6).

4.1.3 Scott et al

In a paper by Scott et al (2004) on a grammar to describe requirements more precise. The background of their research is that "Project Failures are often attributed to poor requirements". To reduce the risk, they define the need to look at each requirement's quality and present a list of "high quality properties" of requirements they gathered from previous research.

The grammar they come up with is highly relevant for the requirement engineering process but could also be of use in a requirement parser to check incoming requirements. They have created a prototype

of such a parser to prove the validity of the grammar and propose other tools. For the research question, the tool is not relevant as this research is about the requirement assessment criteria, not about the tool that uses these criteria. However, the properties they mention can be used in the requirement assessment and need little rephrasing.

A total of fifteen criteria were found in the article, see Appendix D Table 28; it was possible to apply the label “Perspective” to the criteria. The mentioned “high quality properties” are an indication of a label that can be applied (see paragraph 3.2.6). At this point, a common type of labelling seems to emerge as the previous articles contain similar labels. For the research, we will now call this label “Meta aspect” and update the tables containing the review results (see Appendix D). The meta aspect describes the characteristics of a criterion in a single term.

4.1.4 Grenn, Sarkani and Mazzuchi

Grenn, Sarkani and Mazzuchi (2014) introduce the “Requirements Entropy Framework” (REF) “*for measuring requirements trends and estimating engineering effort during system development*” (Grenn et al., 2014, p. 462). Based on a literature review, they have selected fourteen requirement quality attributes to use in the REF, plus an additional “zero quality” attribute. These quality attributes are indicators of requirement quality. When a requirement has these quality attributes, it has the desired “end state” in the REF.

Usage of the REF itself is outside the boundaries of the research question, but the mentioned quality attributes are relevant and look at requirements mostly from a quality perspective. The format of the attributes is in line with the required formatting for this research so no rephrasing had to be done.

A total of fourteen criteria were found in the article, see Appendix D Table 29. It was possible to apply the label “Perspective” to the criteria.

4.1.5 Hirshorn

When looking for checklists in relation to requirement validation, relevant information can be found. A technical report from NASA, “*Expanded Guidance for NASA Systems Engineering. Volume 2: Crosscutting Topics, Special Topics, and Appendices*” (Hirshorn, 2017), contains a comprehensive “Requirements Validation Checklist” that includes a large number of requirement quality aspects.

The drawback of this report is that is has not been cited yet and the checklist does not refer to other literature which makes it less reliable from a scientific point of view. However, the criteria mentioned resemble the earlier findings, can be rephrased easily and are therefore included.

A total of 45 criteria were found in the article, see Appendix D Table 30. It was possible to apply the label “Perspective” to the criteria.

4.1.6 Stephenson, Attwood and McDermid

Another checklist is provided by Stephenson, Attwood and McDermid (2011). They specify ten uncertainties in four areas and claim that this list contains “*a number of different aspects that might not ordinarily be considered*” (Stephenson et al., 2011, p. 6). They use this list to identify uncertainties of requirements for use in a process that mitigates the uncertainties in the design process instead of going back to the requirements engineering process to improve the requirement and manage the uncertainties there.

This could be a valuable approach to deal with residual uncertainties but that is outside the scope of this research as we’re interested in finding criteria to assess the requirements. Their statement that the list they provide contains aspects that are not ordinarily considered, cannot be confirmed as the other articles reviewed mention similar criteria.

The criteria they mention, are not in the form of statements like in the other articles but are further explained in the column “Form”. Information in this column is used for rephrasing the items into the format for this thesis and to put them in the right perspective.

A total of ten criteria were found in the article, see Appendix D Table 31. It was possible to apply the label “Perspective” to the criteria.

4.1.7 Sadia, Beg & Faisal

Sadia, Beg & Faisal (2014) look at requirements from a risk perspective and propose in their paper a method to "*implement inspection technique for identifying the key requirement risk factors responsible in achieving successful outcome*" (Sadia et al., 2014, p. 13). They do so in the field of software development, which has similarities to the field of systems engineering but is not the same. Another drawback of their list is that they do not give a detailed description of the risks nor specify how they got to the items in the list or where in other literature the list refers to. The risks they mention, however, resemble criteria already found and are therefore included in the list of criteria as this list will be processed and then reviewed by experts.

One of the risks they include is "Volatile requirements", which means that requirements can (and will) change according to Nurmuliani et al (2004). As mentioned in paragraph 2.8, volatility is outside the scope of this research so "Volatile requirements" will not be included as a criterion. They also specify "Impossible requirements" and "Requirement not attainable". As these are synonyms, only "Impossible requirements" will be included.

The eight criteria from this article are listed in Appendix D Table 32. It was possible to apply the label "Perspective" to the criteria. The meta aspect of the criteria were partially based on a risk taxonomy from the article. The challenge with that exercise was that three levels were specified.

4.1.8 Conclusions

In total, 107 assessment criteria were identified from the literature review. The article by Hirshorn contains the largest number of criteria by far but, as mentioned with the review, is more a technical report or guide than an article.

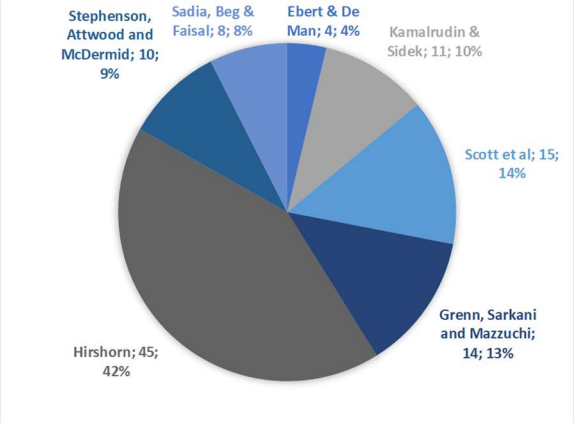

Figure 15: "Number of criteria per article"

Most of the articles are not (exclusively) related to the field of systems engineering and not at all within the sector of civil engineering. A possible reason for the lack of articles on this topic in this sector could be that systems engineering within the sector of civil engineering is relatively new, as also mentioned in the "Guideline for Systems Engineering within the civil engineering sector": "*Since the publication of the Guideline for Systems Engineering version 2 in 2009, a lot has happened in the field of Systems Engineering (SE) within the civil engineering sector. The support base has broadened, due to organisations realising that SE helps to make projects manageable.*" (Rijkswaterstaat et al., 2013, p. 3)

Only the articles by Hirshorn and Grenn, Sarkani and Mazzuchi are directly systems engineering related, the others are within the software engineering domain. However, as the fields of systems engineering and software engineering are very similar (see paragraph 2.1.3), the resulting criteria were marked as relevant to enter the Delphi study. They will be validated in the Delphi study for their relevance within the field of systems engineering. Expanding the search for criteria to other domains might reveal new criteria, but that should be the subject of future research.

From the review, two types of labelling were identified:

1. Perspective. A view on the requirement (for example: "Quality")
2. Meta aspect. Describes the characteristics of a criterion in a single term

Detailed findings and discussions on the literature review can be found in paragraph 5.1.1.

4.2 Measuring requirements uncertainty

This paragraph includes a literature review on RQ1.1. A total of six relevant articles were found to identify measurement methods. A brief summary of each article is provided. In chapter 5, the overall findings of the literature review will be discussed.

4.2.1 Gemino, Saur and Reich

Gemino, Saur and Reich (2007) recognise requirements uncertainty as one of the factors that are a project risk. They use the term “requirement definition” in a table of variables but actually mean “requirements uncertainty” as they explain in the paragraph on the variables they use. Each variable they use has its own values for measurement but for requirements they use a 7-point Likert scale with the following values:

1. Very Certain
2. Certain
3. Somewhat Certain
4. Average
5. Somewhat Uncertain
6. Uncertain
7. Very Uncertain

One of their findings is that “reducing the level of uncertainty will increase the likelihood of performance”, which could justify spending time reducing uncertainty before the project starts. Requirements uncertainty is identified as the most important risk, as can be seen in Figure 16.

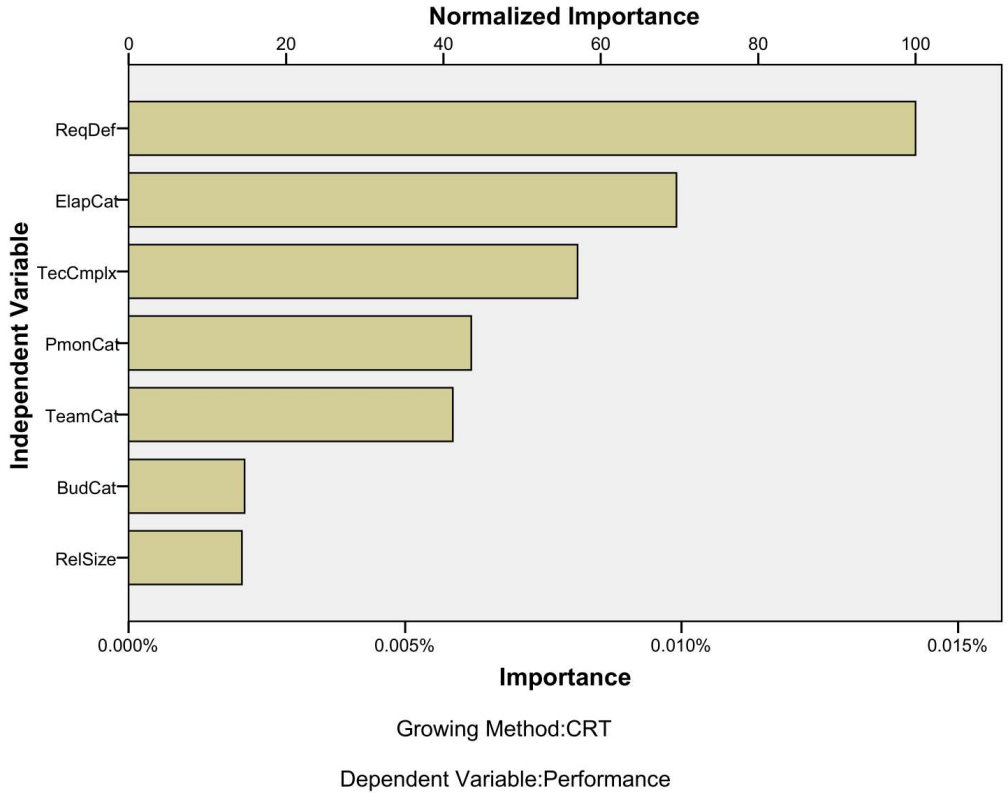

Figure 16: Importance of requirements uncertainty (ReqDef) as a risk factor in projects (Gemino et al., 2007)

The scale items they use to measure uncertainty are not an exact science and depend on the interpretation of the requirement reviewers. No further explanation was given on how to determine when to score uncertainty with a specific value.

4.2.2 Mellis, Loebbecke & Baskerville

In the field of software development, Mellis, Loebbecke & Baskerville (2010) also look at project performance based on different variables, one of them being requirements uncertainty. To test some

of their hypothesis, they use a “degree of requirements uncertainty” and divide into four categories for their samples:

1. Highest
2. Higher
3. Lower
4. Lowest

The same remark on what these values mean can be made as with the review on Gemino, Saur and Reich (paragraph 4.2.1): no further explanation was given on how to determine when to score uncertainty with a specific value. Again, no exact science but more based on the person performing the measurement.

4.2.3 Toth

Also in the area of software development, Toth (2008) looks at estimating techniques and influencing factors on projects. He identifies requirements as one of the key influencing factors, with uncertainty as one of its attributes, see Figure 17.

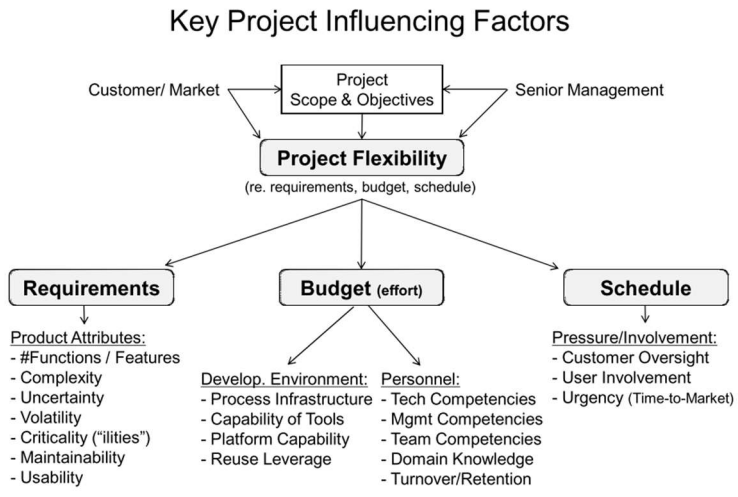

Figure 17: "Key project influencing factors (Toth, 2008)"

Requirements uncertainty is then measured with two values:

1. Low uncertainty. This is used for unambiguous, precise, and logically complete specifications
2. High uncertainty. This is used for high-level, vague and incomplete requirements.

The meaning of the values is clarified to some extend but leave room for interpretation. A measurement with just two values is not as elaborate as the ones in the articles reviewed before but that could also be an advantage. Like the methods before, measuring uncertainty using this method is not an exact science.

4.2.4 Keutel, Michalik and Mellis

Keutel, Michalik and Mellis (2011) did research on different requirements uncertainty situations and how these should be dealt with. They observed several dimensions of requirements uncertainty but focussed on just three of them as these three “*only describe the characteristics of the requirement itself and not external constraints*” (Keutel et al., 2011, p. 78). The three dimensions are:

1. Number of alternatives. “Describes the amount of possible different specifications of a single requirement”.
2. Diversity. “Describes the extent to which the future users’ needs differ regarding a single requirement”.
3. Complexity. “Describes the difficulty of understanding, specifying and communicating the requirement”.

The dimensions are rated on the following scale:

1. High
2. Low

For each dimension, they provide examples on when to rate with High or Low, for example: *We rate complexity also as high, if the question cannot be expressed in one simple sentence but needs more detailed explanation (difficulty to understand and communicate).* (Keutel et al., 2011, p. 78)

They visualise these dimensions and ratings in a three-dimensional model, as can be seen in Figure 18.

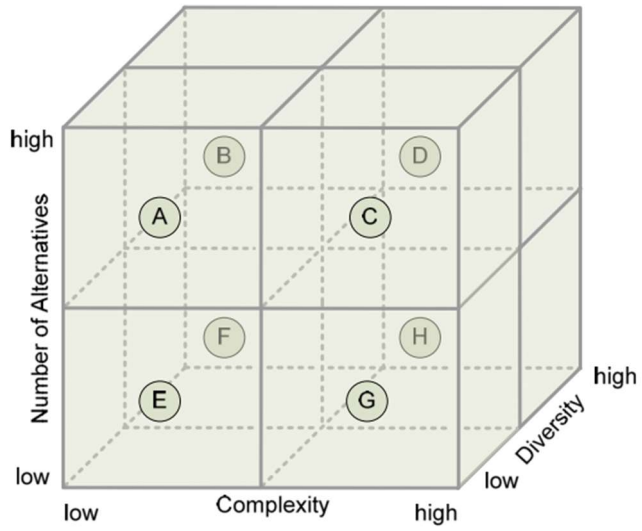

Figure 18: "Requirements Uncertainty Situation Types (Keutel et al., 2011)"

Although there might be more dimensions, the choice of limiting them to just three options with a scale of two, creates an easy to understand model. The examples given for when to use High or Low give some directions but what the boundaries of these values might need consensus on a project or company level.

Measuring uncertainty using this method compared to the previous reviewed ones is potentially more exact as number of alternatives can be measured and the boundaries of the dimension configured, for example: up to 5 = Low, more than 5 = high. Unfortunately, Diversity and Complexity seem much more difficult to measure precisely so it's up to the person performing the measurement again and not an exact science.

4.2.5 Nolan, Abrahão, Clements, & Pickard (1)

A similar approach is taken by Nolan et al. (2011), but their interpretation of requirements uncertainty is one of volatility: unexpected requirements change that can cause impact through rework. Although this is about residual requirements uncertainty and not initial requirements uncertainty (see paragraph 2.5) and more a risk tool, the method might still be useful for initial requirements uncertainty and is therefore included.

The authors use a technique with standard risk principles assigning *probability* (of requirement change) and *impact* values to each requirement and plotting the result in a matrix, as shown in Figure 19.

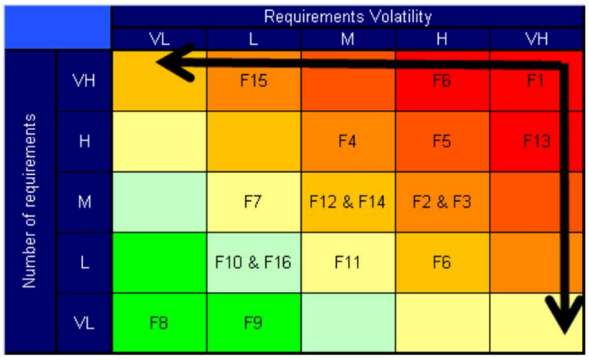

Figure 19: "Matrix with plotted values for probability and impact (Nolan et al., 2011)*"*

In a 3x3 matrix, the following values are used:

- L=Low
- M=Medium
- H=High

The same methodology can also be used in a 5x5 variant with the following values:

- VL=Very low
- L=Low
- M=Medium
- H=High
- VH=Very high

The method used is very straightforward but assigning values for probability and impact is, again, not an exact science.

4.2.6 Nolan, Abrahão, Clements, & Pickard (2)

In the same paper, Nolan et al. (2011) also take a more elaborate approach to predict requirements uncertainty and add four attributes to requirements that together are used in a formula to calculate a "Risk Index" (RI). As well as the method reviewed in paragraph 4.2.5, this method is also about residual requirements uncertainty and not initial requirements uncertainty (see paragraph 2.5) and more a risk tool. However, the method might still be useful for initial requirements uncertainty and is therefore included.

The RI is calculated with the following attributes:

- Volatility (V). This is the probability the requirement will change through the course of the project. Volatility is measured by the following values with their correspondent weight and percentage of volatility:
  - Very high volatility = 9; >90% volatility.
  - High volatility = 7; 70%-90% volatility.
  - Medium volatility = 5; 50%-70% volatility.
  - Low volatility = 3; 30%-50% volatility.
  - Very low volatility = 1; 10%-30% volatility;
- Impact (I). This is the degree that a change in the requirement will negatively affect a development program. Impact is measured by the following values with their correspondent weight and effect (on cost and schedule):
  - High impact = 9; >20% cost, >4 weeks slip.
  - Medium impact = 3; 10%-20% cost, 1-4 weeks slip.
  - Low impact = 1; <10% cost, <1 week slip.
- Precedence (P). This indicates the experience of the company with a similar requirement in other projects. Precedence is measured by the following values with their correspondent weight and definition:
  - Low precedence = 9; No experience of concept or environment. Historically volatile.

    - Medium precedence = 3; Some experience in related environments. Some historic volatility;
    - High precedence = 1; Concept already in service. Low historic volatility.
- Time Criticality (TC). This indicates the priority in when the requirement is needed. Time Criticality is measured by the following values with their correspondent weight and definition:
    - Urgent = 9; Needed now, this project phase.
    - Next = 3; Needed soon, next project phase.
    - Delay = 1; Needed later, two or more project phases from now.

The formula they use to calculate RI is:

$$RI = \frac{V * I * P * TC}{729}$$

To understand what the risk of the individual requirement is in the set of requirements, the “Proportional Risk Index” (PRI) is used. This is calculated as follows:

$$PRI = \frac{RI}{\sum RI}$$

When plotting the PRI’s in a chart, it can be identified to project management which requirements should be focussed on first.

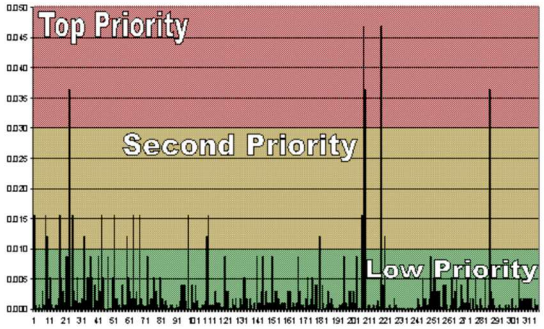

Figure 20: "Identifying priority on requirements based on their PRI (Nolan et al., 2011)"

The method is quite elaborate and will be time consuming when used for a large set of requirements. Assigning values for the attributes is, again, not an exact science but this method seems to be the most exact of them all.

4.2.7 Conclusions

A total of six measurement methods were found, ranging from simple to very complex. Although the concept of requirements uncertainty and its impact on project performance is confirmed, measuring uncertainty seems to be more an art than a science as none of the methods is (fully) based on objective parameters.

An interesting finding is that all articles were published roughly within the same time frame (from 2007 to 2011). Also, only six methods is a rather small set compared to the more than 2600 search results from the query (see paragraph 3.2.5), especially because it is recognised that requirements uncertainty has a negative effect on project performance. Perhaps the influence of controllers demanding a financial justification of requirements uncertainty reduction efforts is relatively new and therefore the topic has escaped the attention of the scientific community. This would be subject to further research as no apparent reason was found in this research.

To make any of the methods useful within the context of an organisation, some guidelines will need to be created on how to fill in values for the parameters and it might be good to let a group of people do the measurement to reach consensus on the outcome. This also would be subject to future research.

Detailed findings and discussions on the literature review can be found in paragraph 5.1.2.

4.3 Factors that influence acceptable requirements uncertainty

This paragraph includes a literature review on RQ1.2. A total of two relevant articles were found to identify influencing factors. A brief summary of each article is provided. In chapter 5, the overall findings of the literature review will be discussed.

4.3.1 Nolan et al

In a paper by Nolan et al (2011), a theoretical cost-benefit trade-off (cut-off point) between risk and mitigation costs is introduced in a model that shows a “sweet spot”: the point where the business achieves maximum return-on-investment, as can be seen in Figure 21.

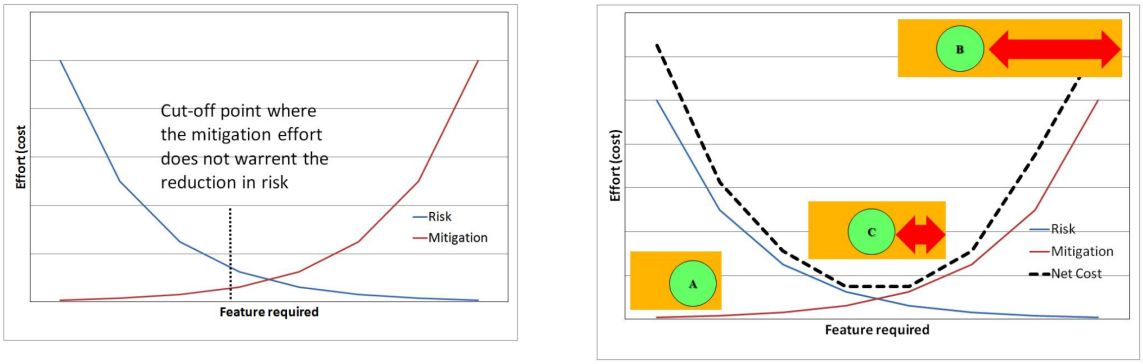

Figure 21: “Cost-benefit cut-off point or ‘sweet spot’ (Nolan et al., 2011)*"*

Taking additional mitigating actions to reduce the risk does is not justified because of the actual increase in costs.

The article was written in the context of Rolls-Royce, where the cut-off point is determined at a point where mitigation effort is less 1/3rd the expected cost of the risk of the uncertainty. The main point of view is Risk, not Uncertainty itself.

Because not much information was found based on the search query specified in paragraph 3.2.5, an additional query was done with search terms that were found in the article by Nolan et al. ("systems engineering" risk cost benefit) to check if an additional source could confirm the cost/benefit view. This returned a relevant paper by Faber & Stewart (2003) on risk assessment for civil engineering facilities. In this paper, the authors present a model that defines areas of risk acceptation on an abstract level. In this model, the middle part represents the “As Low As Reasonably Possible” (ALARP) or “Tolerable” region, in which risk reduction is impracticable or too costly, see Figure 22.

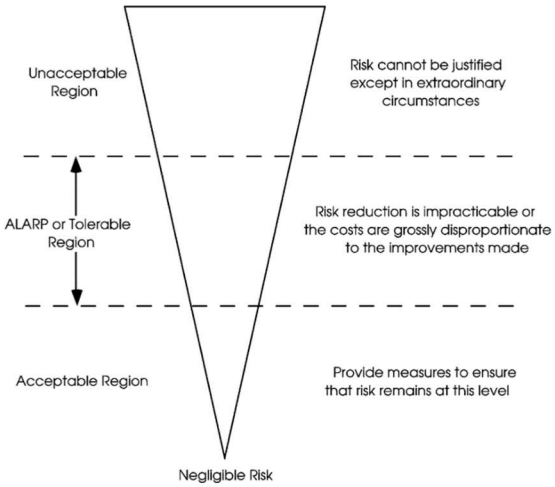

Figure 22: "Risk acceptation regions (Faber & Stewart, 2003)"

Although the model in the paper by Faber & Stewart cannot be mapped one-on-one to the model by Nolan, the key concept of cost-benefit is confirmed.

4.3.2 Walton

Walton (2002) looks at uncertainty from a portfolio point of view in his thesis “Managing uncertainty in space systems conceptual design using portfolio theory”. Requirements uncertainty was identified as the highest source of uncertainty according to the interviewees in his research, modern portfolio theory was applied to manage the uncertainty.

Portfolio theory is concerned with recommending investment strategies that balance the needs of an individual investor to achieve the maximum return for as little uncertainty as possible. In his research, Walton identifies the level of “risk aversion” of decision makers as an important factor in defining the right strategy to deal with uncertainty. There are three levels specified:

1. Low risk aversion
2. Moderate risk aversion
3. High risk aversion

Depending on the aversion appetite of a company, a lower or higher degree of uncertainty is acceptable. Figure 23 is an example of the value/uncertainty trade with the three levels. Although the scale values are specific to the case study by Walton, it becomes clear that with a high risk aversion, the cost for the strategy of dealing with uncertainty are higher than for a strategy with a low risk aversion. In that sense, the risk aversion view is related to the cost/benefit view but adds a layer on top of it to determine what cost/benefit ratio is acceptable.

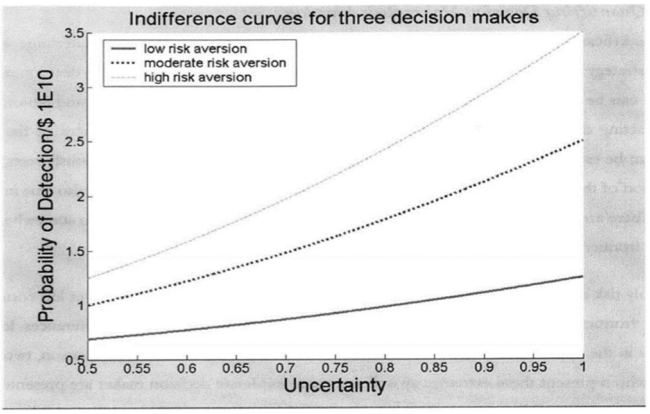

Figure 23: "Risk aversion types (Walton, 2002)"

Because not much information was found based on the search query specified in paragraph 3.2.5, additional queries were done with search terms that were found in the article by Walton to check if an additional source could confirm the risk aversion view.

A query with the search terms “risk”, “averse” and "requirements uncertainty" returned a recent article by Li, Harman, Wu & Zhang (2017). In their article, they state that for a risk-averse decision maker, risk reduction is more valuable than the gained utility so he “*will not choose optimal-yet-risky solutions and would accept the guaranteed robust-yet-suboptimal solutions*” (Li et al., 2017, p. 590). Risk-loving decision makers, on the other hand, would rather go for the “optimal-yet-risky solutions”. A high risk aversion means accepting less requirements uncertainty and accepting more cost to mitigate the uncertainty. This confirms the findings by Walton.

4.3.3 Conclusions

Not much information on influencing factors was found in literature. The two factors found (cost-benefit and risk aversion), however, seem to be concepts that are well known in economics (e.g. Boardman, Greenberg, Vining, & Weimer, 2017; Dow & da Costa Werlang, 1992; O’Donoghue & Somerville, 2018). When searching for “cost benefit” in combination with “systems engineering”, many results are returned but not specifically in relation to requirements uncertainty.

Perhaps the concept of “acceptable requirements uncertainty” is not well recognised or a different set of search terms would have produced more results. Factors like “past experience” or “knowledge level” could be valid as well, although these could be sub factors influencing the factor risk aversion. In the Delphi study, the participants will be allowed to add factors, maybe this will reveal factors that come from experience in the field rather than from scientific studies.

Detailed findings and discussions on the literature review can be found in paragraph 5.1.3.

5 Findings and discussions

In this chapter, the findings of the literature review and the Delphi study will be discussed. The results of the literature review are input for the Delphi study.

5.1 Literature review

This paragraph contains the findings on the literature review that was done in chapter 4, in the same order.

5.1.1 Requirements assessment criteria

In the literature review, a total of 107 assessment criteria were found. After coding them to eliminate synonyms, 66 criteria remained (the full list can be found in Appendix E . This is more than the nine characteristics specified in the NEN-ISO/IEC/IEEE 29148:2011 norm on Requirements engineering (ISO/IEC/IEEE, 2011) but some of the items found in literature match the characteristics, an example is provided in Table 7.

Table 7: "Example similarity between NEN-ISO/IEC/IEEE 29148:2011 and literature findings"

From literature	From NEN-ISO/IEC/IEEE 29148:2011 en
CP_{16}: The requirement does not contain ambiguities and can be understood the same by all readers	Unambiguous. The requirement is stated in such a way so that it can be interpreted in only one way. The requirement is stated simply and is easy to understand.

In the literature review, two types of categories were identified:

- **Perspective**. Criteria found for RQ1 can be grouped by their *perspective*: a view on the requirement (for example: “Quality”). These perspectives are only used for easy filtering or grouping purposes in the Delphi study and are not a part of the research question. However, they might trigger the expert panel to rank criteria higher or lower based on the perspective but this is not explicitly researched.
- **Meta aspect**: criteria can fall within the same aspect group, for example “consistency”. The meta aspect describes a property/attribute for the criteria in a single term. The meta aspects are only used for easy filtering or grouping purposes in the Delphi study and is not a part of the research question. However, they might trigger the expert panel to rank criteria higher or lower based on the meta aspect, but this is not explicitly.

From the list, three perspectives were identified:

1. **Business**. Criteria with this perspective haven an influence on the business case, cost/benefit or the stakeholders.
2. **Quality**. Criteria with this perspective are concerned with how well the requirement is written, concerning format, language, style, syntax, required sections and information.
3. **Risk**. Criteria with this perspective are concerned with the consequences of the requirement for the project, like conflicting information, constraints, assumptions. There is a thin line between the Quality and Risk perspective, but when a requirement with a high quality (as described above) could still cause a problem for the project, it gets a Risk perspective

The number of perspectives was intentionally kept low to be useful and were kept close to the domains of the intended participant roles:

- Project/integration managers look at requirements more from a business and project risk point of view
- (Lead) engineers/architects look at requirements from a technical feasibility point of view, which is considered a risk perspective in this research as low feasibility is risk to the project.
- (Lead) test engineers are often qualified in the area of requirements engineering and look at quality and risk.

Many of the criteria seem to be related to either Quality (23) or Business (13), which could indicate that a better requirement engineering process leads to less uncertainty as this process is responsible for creating well defined, good quality requirements that match the business case. The Delphi rounds will reveal what the final list of perspectives is, this might support the finding.
An initial number of 49 different meta aspects were found in the literature review but this was brought down to 18 after coding the initial assessment criteria in which the meta aspects were coded as well to eliminate synonyms. The resulting list is:

1. **Administration** (3 criteria)
2. **Completeness** (5 criteria)
3. **Consistency** (20 criteria)
4. **Correctness** (2 criteria)
5. **Cost/Benefit** (3 criteria)
6. **Design independency** (3 criteria)
7. **Feasibility** (1 criterion)
8. **Interfaces** (3 criteria)
9. **Maintainability** (1 criterion)
10. **Necessity** (4 criteria)
11. **Performance** (2 criteria)
12. **Precision** (1 criterion)
13. **Reliability** (4 criteria)
14. **Stakeholder support** (1 criterion)
15. **Testability** (1 criterion)
16. **Traceability** (4 criteria)
17. **Validity** (1 criterion)
18. **Verification** (7 criteria)

The number of criteria for Consistency seems rather high compared to the number of criteria for the other meta aspects. In the Delphi study, it will be validated by an expert panel if the list of meta aspects and the assignment of criteria to them is valid (within the context of the sample).

5.1.2 Measuring requirements uncertainty

In the literature review, three different types of measurements methods were found in six articles:

1. One dimensional measurement. The articles reviewed in paragraph 4.2.1, 4.2.2 and 4.2.3 use a single value to indicate requirements uncertainty:
    a. M_1: Very Certain, Certain, Somewhat Certain, Average, Somewhat Uncertain, Uncertain, Very Uncertain (Gemino et al., 2007)
    b. M_2: Highest, Higher, Lower, Lowest (Mellis et al., 2010)
    c. M_3: Low uncertainty, High uncertainty (Toth, 2008)
2. Multi-dimensional measurement. The articles reviewed in paragraph 4.2.4 and 4.2.5 use two attributes with a value to indicate requirements uncertainty:
    a. M_4: Number of alternatives, Diversity, Complexity | High, Low (Keutel et al., 2011)
    b. M_5: Probability, Impact | Very Low, Low, Medium, High, Very High (Nolan et al., 2011)
3. Complex measurement. The article reviewed in paragraph 4.2.6 uses a formula with four parameters to indicate requirements uncertainty:
    a. M_6: Volatility (Very high, High, Medium, Low, Very low), Impact (High, Medium, Low), Precedence (Low, Medium, High) and Time Criticality (Urgent, Next, Delay) (Nolan et al., 2011)

All measurement methods are based on estimates and interpretations by the persons reviewing the requirement and depend on consensus on what the values mean (a value of “Somewhat Uncertain” could have a different meaning for person A than for person B). The meaning might be project or company specific, the value chosen by the reviewer highly dependent on his or her experience or knowledge.

A group process to combine estimates might be introduced to make the outcome less dependent on the individual reviewer and make the estimate more reliable. An example of such a process is found in Agile development models: *planning poker*, a method where a group of experts each give points (on a

Fibonacci scale) to a user story (a functional increment) (Agile Alliance, 2018). There are studies that seem to confirm that planning poker produce better estimates than individual expert estimates (Mahnič & Hovelja, 2012; Moløkken-Østvold, Haugen, & Benestad, 2008). The planning poker process could be adapted to suit the requirements uncertainty measurement estimate by making the reviewers discuss their individual estimates and reach consensus.

5.1.3 Factors that influence acceptable requirements uncertainty

The literature review reveals two factors that influence what level of requirements uncertainty is acceptable:

1. F_1: Cost/benefit. If the (estimated) cost of trying to reduce the requirements uncertainty is higher than the benefit, no further effort should be made to reduce the uncertainty, unless the risk is unacceptable.
2. F_2: Level of risk aversion. The higher the level of risk aversion, the less uncertainty will be tolerated. Risk aversion could be a property of the reviewer, the project or even the organisation.

Even though the search queries initially returned a large number of search results, very little relevant information was found and only two articles made it through the selection process. It also proved to be hard to find articles that confirmed the findings from the two articles reviewed. Therefore, the reliability of the results is low and should be evaluated in the Delphi study. This might also lead to additional factors.

5.2 Delphi rounds

This paragraph contains the findings of the Delphi rounds.

5.2.1 Round 1

A response of 100% was received for Round 1. In SPSS (IBM, 2017), the data was split to compare groups based on the IDs of the items. Items with a median of 4 or 5 are considered to be “relevant” in the light of the research question. The participant comments were analysed to find common findings (not specific to the individual items).

5.2.1.1 Assessment criteria

Fourteen new criteria were added to the list by the participants. On 59 criteria, at least one participant wanted to change the perspective. On 21 criteria, at least one participant wanted to change the meta aspect. At this point it is too early to draw conclusions on these changes as these kinds of changes can be expected in a Delphi study.

Table 8: "Statistical data on assessment criteria Round 1"

Relevance score (median) [4]	# criteria
3	21
3.5	3
4	31
5	11
	66

Range[5]	# criteria
1 point	2
2 points	10
3 points	17
4 points	37
	66

Variance[6]	# criteria
> 0.9	46
=< 0.9	20
	66

More detailed statistics can be found in Appendix I .

[4] Only criteria from the literature review

[5] Only criteria from the literature review, new criteria received only one score (by the participant who added the criterion).

[6] Only applicable to items from the literature review

Although 42 out of 66 criteria have a median of 4 or higher, which means they are relevant, there is a high variance on 46 criteria and also a high range on 37 of them. This is interesting to send back into Round 2 to see if opinions will be changed by reviewing each other's comments.

From the review comments and general comments sent with the questionnaires, a few common findings emerge (see Appendix J for the content analysis data). A general finding, according to the comments with returned questionnaires, is that a list of 66 criteria is regarded as too large to be practical:

"*The collection of criteria could possibly be reduced to about 20+ items*"[7]

"*There are quite many criteria. In practice, you'll never work with 60+ at the same time*"[8]

The goal of this research is to produce a list of relevant criteria and the results of round are that this list is smaller than the initial list (see Table 8). Direct evidence for a preferred maximum of requirements assessment criteria was not found but in the field of "multi-criteria decision making" (MCDM)[9], research seems to indicate a maximum number of ten criteria (González-Araya, Rangel, Lins, & Gomes, 2002; Nijkamp & van Delft, 1977).

Another general comment is that the "Perspective" classification might not add value to the list of criteria.

"I noticed that that most perspectives have a 1-to-1 relation to a meta aspect." [10]

Regarding the criteria, the difference between the meta aspect and perspective is not really clear.[11]

The NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011) only uses a meta aspect (like "Unambiguous"). However, in a paper by Wholin & Aurum (2005) on software requirements with the research question "*Is any perspective more important than others when deciding what requirements to include in a specific project or release and is it likely to change over time?*" , thirteen criteria were identified from three different perspectives that influence the decision-making process, "business" being one of them. Although their research is about the decision-making process on what requirements to include in a project and not on requirements uncertainty, their findings might indicate that Perspective is a useful label for assessment criteria as well.

From the comments it can also be concluded that it is important to interact with the client about how well the requirement is understood. It is also important to understand the rationale of the requirement to make sure the right solution will be implemented:

"Is my interpretation of your requirement correct?"

"How do you assess this criterion without having interaction with the client?"

This finding is confirmed by the activity "Analyze and maintain stakeholder requirements" from the NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011): "*It is important to continue to perform requirements negotiation during the analysis and allocation of requirements, because conflicts will*

[7] Translated from Dutch. Original quote: "de verzameling criteria wellicht tot zo'n 20+ verschillende ingedikt zou moeten kunnen worden"

[8] Translated from Dutch. Original quote: "Het zijn wel heel veel criteria. In de praktijk zul je er nooit 60+ naast elkaar gaan hanteren."

[9] Multiple criteria decision making (MCDM) refers to making decisions in the presence of multiple, usually conflicting, criteria (Xu & Yang, 2001)

[10] Translated from Dutch. Original quote: "het viel me op dat de meeste meta-aspecten aan een specifiek perspectief gekoppeld zijn"

[11] Translated from Dutch. Original quote: ""Ten aanzien van de criteria is ook het onderscheid ''perspecfief'' een ''meta data'' niet echt helder.

occur. […] In most cases, it is necessary to consult with the stakeholder(s) to reach a consensus on an appropriate trade-off."

Some of the criteria seem to belong to the business domain. The business case or added value of the functionality is what counts, this is outside the scope of the contractor:

"Client should do his prioritising"

"Actually, our client is responsible for describing all necessary functions in the requirements of the contract"

"If there is no business need, a requirement is simply not valid"

Another common remark is that it is important that there is room in the design phase to find suitable solutions, the client should be more concerned with *what* is needed (the black box):

"Business experts are not qualified to define technical requirements"

"Specifying behaviour and output is important. But at a level which is in line with the development phase of the project. Design freedom must stay in place."

Again, the NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011) confirms this finding: "*Implementation Free. The requirement, while addressing what is necessary and sufficient in the system, avoids placing unnecessary constraints on the architectural design.*"

One participant commented on several criteria that they should be allocated to a higher level than the individual requirement:

This criterion seems to look at the quality of the tender process or the project strategy and not the requirement itself.

This is confirmed by the NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011) as it specifies that "*There are certain characteristics that need to be considered for the set of stakeholder, system, and system element requirements rather than for any individual requirement*". As the comments came from one participant only, Round 2 will be analyzed for similar comments from other participants.

Although the research goal is to provide an objective list of validated criteria, some comments indicate that the skills of the professional who deals with the requirements are important as well when looking at the requirements:

"A question of professional capabilities to avoid the pitfall"

"It is the professional who needs to be critical and who should acquire clarification"

5.2.1.2 Measurement methods

One new measurement method was added to the list by the participants, see Table 10. Only one measurement method found in literature scored a median of 4 or higher. On most items, however, no consensus was reached yet, see Table 9.

Table 9: "Statistical data on measurement methods Round 1"

Relevance score (median)	# methods
1	0
2	1
3	4
4	1
5	0
	6

Range[12]	# methods
1 point	0
2 points	1
3 points	2
4 points	3
	6

Variance[13]	# methods
> 0.9	4
=< 0.9	2
	6

Table 10: "Measurement methods added by the participants"

ID	Description	Type
M7	Business uncertainty of requirement quality (0% - 100%) Technical uncertainty of requirement quality (0% - 100%)	multi-dimensional

More detailed statistics can be found in Appendix K .

From the review comments, a few common findings emerge (see Appendix L for the content analysis data). Although the measurement methods should provide a means of measuring / quantifying uncertainty, some participants confirm that quantifying remains difficult, even with a measurement method:

As many business needs are hard to quantify but still relevant, the measurements need to be in the same terms.

Only with very experienced reviewers this can be useful. Otherwise the results will vary a lot and are probably less useful.

One participant mentioned that the concept of requirements uncertainty is not recognised:

Whole issue of requirement uncertainty not really recognized. A requirement is contract for us. So no uncertainty!!

In Round 3, the concept of requirements uncertainty will be explained again to make clear that the assessment criteria are not about refusing to deliver required functionality but rather to make sure there is certainty about what to build.

Most methods were regarded as either too complex or not elaborate enough (sometimes even both, as was the case with M_3):

Two is not enough to make a good distinction

Simplicity is often good. […]. Therefore, less is more and change management is crucial.

This measurement seems overly complex on first glance.

I think this method can be useful, however also relatively complex.

The overall finding after reviewing the comments is that there seems to be a need for a method that is not too hard but not too simple but there does not seem to be an agreement on the number of variables required.

[12] Only measurement methods from the literature review, new methods received only one score (by the participant who added the method).

[13] Only applicable to items from the literature review

5.2.1.3 Influencing factors

Five new influencing factors were added to the list by the participants, see Table 12. All influencing factors found in literature scored a median of 4 or higher.

Table 11: "Statistical data on influencing factors Round 1"

Relevance score (median)[14]	# factors
1	0
2	0
3	0
4	1
5	1
	2

Range[15]	# factors
1 point	0
2 points	1
3 points	1
4 points	0
	2

Variance[16]	# factors
> 0.9	0
=< 0.9	2
	2

Table 12: "Influencing factors added by the participants"

ID	Description
F_3	Past experience
F_4	Project/tender strategy
F_5	Law and regulations
F_6	Organization
F_7	design phase of project

More detailed statistics can be found in Appendix M .

Both influencing factors scored a median of 4 or higher. What is interesting, is that five new factors were identified, much more than was found in literature. In the literature review conclusion, this was already suspected. Perhaps this indicates a gap between scientific research and daily practice or the fact that little research has been done in the specific field of systems engineering and/or the civil engineering sector.

From the review comments, a few common findings emerge (see Appendix N for the content analysis data). Experience seems like an important factor that influences the acceptable uncertainty:

> *If it is known how to deal with a requirement from past projects, more uncertainty is probably allowed.*

> *How familiar the project organization, colleagues, etc are with the rules and context of the environment in which the implementation should operate defines how certain (how elaborate) the requirements need to be*

A cost/benefit and risk relation are also mentioned:

> *Very often risk aversion is solely focused on cost and very often leads to suboptimal benefits.*

> *Cost/benefit is part of the risk assessment as risks will be measured in cost/benefits and are therefore less its own factor*

[14] Only influencing factors from the literature review

[15] Only influencing factors from the literature review, new influencing factors received only one score (by the participant who added the factor).

[16] Only applicable to items from the literature review

This could indicate that F_1 and F_2 are not separate factors, but that conclusion cannot be drawn from just two comments. The observation that cost-benefit and risk have a relation can be found in literature on risk management (e.g. Boehm, 1991; Merkhofer, 2012) but an exact analysis is outside the scope of this thesis.

5.2.2 Round 2

A response of 100% was received for Round 2. In SPSS (IBM, 2017), the data was split to compare groups based on the IDs of the items. Items with a median of 4 or 5 are considered to be “relevant” in the light of the research question. The participant comments were analysed to find common findings (not related to the individual criteria).

In Round 2, the participants were presented with the results from Round 1 by:

- a research summary;
- scores and comments from Round 1 as part of the questionnaire for Round 2 (without revealing the names of the participants).

5.2.2.1 Assessment criteria

From the list of assessment criteria, eight were identified by the participants as duplicates of other criteria, these are removed from the tables below. Of the 72 remaining criteria, the participants reached consensus on 29 that these were relevant (median 4.0 or higher). They also reached consensus on the non-relevance of thirteen of the 72 criteria. On 30 criteria, no consensus was reached yet, these will be included in Round 3.

No new criteria were allowed to be added.

Table 13: "Statistical data on assessment criteria Round 2"

Relevance score (median)	# criteria
2.0	4
3.0	24
3.5	1
4.0	34
4.5	1
5.0	8
	72

Range	# criteria
1 point	6
2 points	24
3 points	22
4 points	20
	72

Variance	# criteria
> 0.9	30
=< 0.9	42
	72

Per criterion, the Perspectives and Meta aspects had to be chosen from a list of possible Perspectives and Meta aspects consisting of the original ones from the research plus the ones added for that criterion in Round 1 by the participants. The remaining perspectives and meta aspects can be found in Table 14 and Table 15.

Table 14: "Remaining perspectives after Round 2"

Perspective	# Criteria
Business	22
Quality	38
Requirements technique	1
Risk	11
	72

With all criteria, the perspective was determined by a majority vote. A notable finding is that the perspective “Requirements technique” was chosen for one criterion only. This could indicate that this perspective is not a valid one or part of one of the other perspectives. According to one participant (in the comments), the perspective “Quality” is a better match. This will be input for the questionnaire of Round 3.

Table 15: "Remaining meta aspects after Round 2"

Meta aspect	# Criteria
Acceptance	1
Administration	2
Business value	1
Clearness	1
Completeness	16
Consistency	8
Correctness	4
Cost/Benefit	4
Dependency	1
Design independency	3
Feasibility	2
Maintainability	1
Necessity	4
Reliability	2
Safety	1
Stakeholder support	1
Traceability	6
Unambiguity	6
Validity	1
Verification	7
	72

The meta aspect was also determined by a majority vote, it is notable that the meta aspect "Completeness" was assigned the most (22%). More detailed statistics can be found in Appendix O .

Content analysis was performed again on the comments (see Appendix P) and some new general findings were found or findings from Round 1 confirmed. In Round 1, a finding by one participant was that some criteria were about the set instead of the individual requirement. In Round 2, similar remarks were made:

Most of the quality criteria for requirements also apply to the entire set of requirements

As discussed for Round 1, this is confirmed by the NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011). The criteria concerned will be marked as set requirements in the final result. These criteria have in common that they have similar statements, like "*There are requirements concerning […]*" or "*[…] have been specified*".

Although related to just one criterion (CP_{66}), the participants had a strong opinion on assumptions in requirements:

Requirements should never be based on assumptions but on facts or calculations

Moving assumptions into validated requirements or remove them altogether should be your aim during the requirements review

The NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011) does not confirm that assumptions should not be there. A literature scan on Google Scholar also did not produce evidence that assumptions are not allowed.

The benefits of a Delphi study (see paragraph 3.3.1) become clearly visible:

- The number of individual responses is high
- Rich responses
- A movement towards consensus (measured using Variance)

Some participants seem to stick with their initial scores and responses from Round 1, but others have changed their opinion based on comments of other participants. As one participant puts it:

> *I was surprised by the number of comments given by the participants and that has certainly made me realise that there are other / better views to look at the criteria*[17]

5.2.2.2 Measurement methods

An interesting finding of Round 2 is that none of the measurement methods found in literature were regarded as relevant, although for two of them there was no consensus yet, so they might get a higher score in Round 3. No new measurement methods were allowed to be added to the list.

Table 16: "Statistical data on measurement methods Round 2"

Relevance score (median)	# methods
1.0	0
2.0	1
3.0	5
4.0	1
5.0	0
	7

Range	# methods
1 point	0
2 points	3
3 points	4
4 points	0
	7

Variance	# methods
> 0.9	2
=< 0.9	5
	7

More detailed statistics can be found in Appendix Q .

One participant stated that some measurement methods (M_5, M_6) concern remaining uncertainty and are not useful for measuring initial uncertainty:

> *This measurement is designed for measuring uncertainty during a project and not at the initial start of a project, thus about volatility.*

As these measurement methods concerned Probability and Volatility, they indeed indicate future change, but this might not be a problem for determining current uncertainty: if there is a great chance of future change, it is likely there will be uncertainty now. As no consensus was reached yet on the two methods, they will return in Round 3. Content analysis did not reveal new general comments and insights.

5.2.2.3 Influencing factors

Consensus was found on most items, of which two were regarded as relevant and three as not relevant. The two items without consensus will return in Round 3. Content analysis did not reveal new general comments and insights. No new influencing factors were allowed to be added to the list.

Table 17: "Statistical data on influencing factors Round 2"

Relevance score (median)	# methods
1.0	0
2.0	0
3.0	2
3.5	1
4.0	4
5.0	0
	7

Range	# methods
1 point	0
2 points	3
3 points	4
4 points	0
	7

Variance	# methods
> 0.9	2
=< 0.9	5
	7

More detailed statistics can be found in Appendix R .

[17] Original quote (Dutch): “Ik was verrast door de hoeveelheid comments gegeven door de deelnemers en die heeft me zeker doen inzien dat er andere / betere views zijn om naar de criteria te kijken.”

5.2.3 Round 3

A response of 100% was received for Round 3. In SPSS (IBM, 2017), the data was split to compare groups based on the IDs of the items. Items with a median of 4 or 5 are considered to be “relevant” in the light of the research question.

In Round 3, the participants were presented with the results from Round 2 by:

- a research summary;
- scores and comments from Round 2 as part of the questionnaire for Round 3 (without revealing the names of the participants).

The participants were only asked to score on relevance, not to provide comments on their score as Round 3 is the final round so no influencing of the other participants is required. No new items were allowed to be added.

5.2.3.1 Assessment criteria

Of the thirty assessment criteria presented in Round 3, on eight was reached consensus that they are relevant (median 4 or higher), on five was consensus that they are not relevant. That means that on seventeen criteria, no consensus was reached.

Table 18: "Statistical data on assessment criteria Round 3"

Relevance score (median)	# criteria
1.0	0
2.0	3
3.0	10
4.0	17
5.0	0
	30

Range	# criteria
1 point	0
2 points	5
3 points	19
4 points	6
	30

Variance	# criteria
> 0.9	17
=< 0.9	13
	30

More detailed statistics for Round 3 can be found in Appendix S .

With the results of Round 3 known, on a total number of 55 criteria consensus was reached (76%), on 17 there was no consensus (24%). Out of the 55 criteria with consensus, 37 (67%) were marked as relevant (median 4 or higher), 18 (33%) were marked as not relevant. The list of relevant assessment criteria can be found in Table 19.

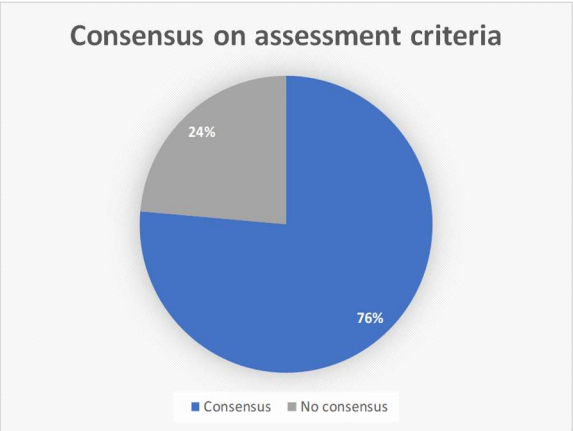

Figure 24: "Consensus on assessment criteria"

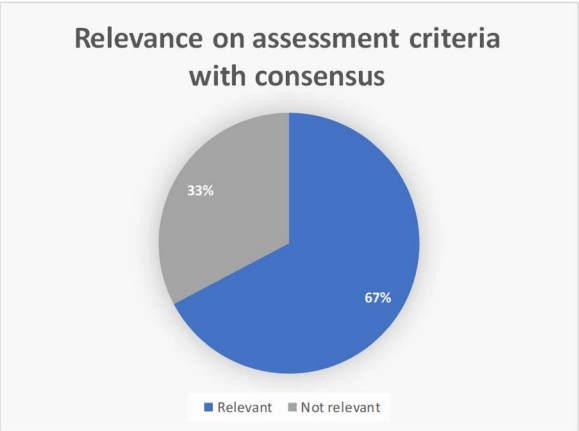

Figure 25: "Relevance for assessment criteria with consensus"

Table 19: "Relevant assessment criteria"

ID	Criterion	Perspective	Meta aspect
CP_1	The requirement reflects the needs of the stakeholders	Business	Business value
CP_7	The requirement can be met within the cost and schedule constraints of the program.	Business	Cost/Benefit
CP_8	All described functions are necessary and together sufficient to meet mission and system goals and objectives	Business	Necessity
CP_9	The requirement is bidirectionally traceable to a higher-level requirement or mission or system-of-interest scope (i.e., need(s), goals, objectives, constraints, or concept of operations)	Quality	Traceability
CP_{12}	The requirement is valid	Business	Validity
CP_{15}	The requirement specifies required behaviour and output for all possible states under all possible constraints	Quality	Completeness
CP_{16}	The requirement does not contain ambiguities and can be understood the same by all readers	Quality	Unambiguity
CP_{17}	Parts of the requirement do not conflict with each other	Risk	Consistency
CP_{18}	All parts of the requirement have the right information (accurate to the right level)	Quality	Completeness
CP_{21}	The requirement links to the needs statements of the stakeholders	Business	Traceability
CP_{23}	The requirement is written in such a way that enables cost effective verification that the system fulfils it	Quality	Cost/Benefit
CP_{24}	The requirement's numeric quantities are accurate and specified to an appropriate level.	Quality	Correctness
CP_{26}	The requirement is concise and simple	Quality	Unambiguity
CP_{27}	The requirement does not contain multiple statements	Quality	Unambiguity
CP_{28}	The requirement and its information is correct	Quality	Correctness
CP_{30}	The requirement is free of unverifiable terms (e.g., flexible, easy, sufficient, safe, ad hoc, adequate, accommodate, user-friendly, usable, when required, if required, appropriate, fast, portable, light-weight, small, large, maximize, minimize, sufficient, robust, quickly, easily, clearly, other “ly” words, other “ize” words)	Quality	Unambiguity
CP_{34}	The positioning of the requirement (section heading, informative context) does not conflicts with its contents	Quality	Consistency
CP_{36}	The requirement does not contain placeholders	Quality	Completeness
CP_{37}	The requirement does not contradict another requirement	Risk	Consistency
CP_{38}	The requirement uses consistent terminology	Quality	Consistency
CP_{39}	The requirement uses appropriate terminology	Quality	Unambiguity
CP_{40}	The requirement does not conflict with other requirements	Risk	Consistency
CP_{41}	The requirement does not have unexpected dependencies with other requirements	Quality	Consistency
CP_{42}	The requirement is (technically) feasible / realistic	Risk	Feasibility
CP_{44}	For the requirement, a procedure can be found to verify that the system meets the requirement	Quality	Verification
CP_{47}	There are no missing requirements	Business	Completeness
CP_{50}	All required performance specifications and margins are listed	Quality	Completeness
CP_{52}	The requirement can be verified in the system	Quality	Verification
CP_{53}	The requirement can be verified at the correct level in the system	Risk	Verification
CP_{54}	All external interfaces are clearly defined	Business	Completeness
CP_{56}	All interfaces are necessary, sufficient, and consistent with each other	Risk	Completeness
CP_{57}	The requirements concerning maintainability of the system have been specified in a measurable, verifiable manner	Business	Maintainability
CP_{58}	There are clearly defined, measurable, and verifiable reliability requirements specified	Business	Reliability
CP_{62}	The accomplishment of the requirement can be measured, and its compliancy verified.	Quality	Verification
CP_{63}	The criteria for verification can be stated	Quality	Verification

ID	Criterion	Perspective	Meta aspect
CP_{65}	In the requirement, don't care" conditions values are explicitly stated (Correct identification of "don't care" values may improve a design's portability.)	Quality	Completeness
CP_{68}	Provide rationale for every requirement that does not have an obvious reason for why the requirement exists.	Business	Necessity

A finding in Round 2 was that the perspective "Requirements technique" was only used for one criterion so in Round 3 the participants were asked to choose the perspective again for this item (CP_{48}). A majority with the narrowest of margins (6 against 5) again voted for "Requirements technique". However, as no consensus was reached on the item (variance of 1.455), this perspective will not appear in the list of relevant assessment criteria. In Table 20, the remaining perspectives and their counts can be found.

Table 20: "Remaining perspectives on relevant assessment criteria after Round 3"

Perspective	# Criteria
Business	10
Quality	21
Risk	6
	37

An interesting finding is that most criteria have a perspective related to business (10) or quality (21). The quality of requirements is highly dependent on the requirements engineering process, the business perspective is also largely related to the requirements engineering phase as it is there that business representatives are asked to indicate business relevance. This could indicate that reducing requirements uncertainty can be achieved not by only assessing requirements at the transfer point but also by strengthening the requirements engineering capability of the client. Other research mentions that improving requirements engineering leads to business improvements (e.g. Kalinowski et al., 2015; Sommerville & Ransom, 2005) so this could confirm the hypothesis.

Of the 20 meta aspects remaining after Round 2, 13 remain after Round 3, the details can be found in Table 21. The meta aspect "Completeness" is used the most followed by "Consistency", both being related to the quality of the requirement, which is in line with the findings on the perspective.

Table 21: "Remaining meta aspects on relevant assessment criteria after Round 3"

Meta aspect	# Criteria
Business value	1
Completeness	8
Consistency	6
Correctness	2
Cost/Benefit	2
Feasibility	1
Maintainability	1
Necessity	2
Stakeholder support	1
Traceability	2
Unambiguity	5
Validity	1
Verification	5
	37

5.2.3.2 Measurement methods

On both measurement methods presented in Round 3, consensus was reached but neither was considered to be relevant.

Table 22: "Statistical data on measurement methods Round 3"

Relevance score (median)	# methods
1.0	0
2.0	1
3.0	1
4.0	0
5.0	0
	2

Range	# methods
1 point	1
2 points	1
3 points	0
4 points	0
	2

Variance	# methods
> 0.9	0
=< 0.9	2
	2

More detailed statistics can be found in Appendix T .

With the results of Round 3 known, only one relevant measurement method remains, a method that was added by the participants in Round 2 (see Table 23). All other measurement methods were regarded as not relevant. This could indicate a gap between scientific research and daily practice.

The remaining method still relies on estimates of the persons who provide the input values and the weighing mechanism is not specific. Further research would need to be done to make the method more specific.

Table 23: "Relevant measurement methods"

ID	Method	Type	Description
M7	Business uncertainty of requirement quality (0% - 100%) Technical uncertainty of requirement quality (0% - 100%)	multi-dimensional	Score the uncertainty of a requirement quality for all quality aspects that are relevant from the business perspective. All quality aspects OK is 0%, none is 100%. Score the uncertainty of a requirement quality for all quality aspects that are relevant from a technical perspective. When combining to one uncertainty value for each requirement, the uncertainty from the business perspective should weigh higher than that of the technical perspective.

5.2.3.3 Influencing factors

On both measurement methods presented in Round 3, consensus was reached and one of them was considered to be relevant.

Table 24: "Statistical data on measurement methods Round 3"

Relevance score (median)	# methods
1.0	0
2.0	0
3.0	0
4.0	2
5.0	0
	2

Range	# methods
1 point	0
2 points	1
3 points	0
4 points	1
	2

Variance	# methods
> 0.9	0
=< 0.9	2
	2

More detailed statistics can be found in Appendix U .

With the results of Round 3 known, three relevant influencing factors remain of which one was added by the participants, see Table 25. Three influencing factors were considered not relevant, on one no consensus was reached.

Table 25: "Relevant influencing factors"

ID	Description
F_1	Cost/benefit
F_2	Level of risk aversion
F_4	Project/tender strategy

These factors seem to be in the strategy or project management realm, more on a team than on an individual level. The factors are not exact science but could be used at the start of the requirements assessment process to create awareness or to define a strategy concerning requirements uncertainty.

6 Conclusion and Recommendations

6.1 Answering the research questions

The rationale for the research in this thesis is that requirements uncertainty has a negative effect on project performance. Although there is a process that can be used to assess requirements, there is no *objective* list of criteria that should be considered in this process. The decision if a requirement should be accepted or not seems to be based solely on *subjective* criteria like experience, gut feeling or commercial opportunism. Not having objective criteria means that the outcome of the assessment process becomes unpredictable, which could leave the contractor at risk when uncertain requirements are accepted and cause major rework in the design or development phase.

In the requirements assessment process, a decision is made what to do with a requirement: accept it, mitigate the uncertainties or send it back to the requirement engineering process. But what level of uncertainty is acceptable, is not a generic value and could be influenced by specific factors. Some kind of measurement method would be required to determine the acceptability level but also to measure the initial uncertainty of a requirement and how much reduction is achieved. An overview of measurement methods and influencing factors would be required.

In this thesis, research was done on the topics described above based on a main research question and two sub questions. This research consisted of a literature review and a Delphi study based on the results of the literature review.

Main research question (RQ1): Which objective assessment criteria are relevant for determining if a requirement should be accepted into the design and development phase?

In the literature review, a list of 107 criteria were found. This list was compressed to 66 criteria by taking out duplicates and then presented to the participants of the Delphi study for validation. The participants added another 14 criteria but many of these were found to be duplicate as well so a list of 72 remained. Consensus was sought on which criteria were relevant by having the participants score the criteria on relevance with a value from 1 (not relevant) to 5 (highly relevant). After three rounds in the Delphi study, on 76% of the criteria (55 of 72), consensus was reached. On 37 criteria consensus was reached that these are relevant (a median of 4 or higher). This list can serve as input for the requirements assessment process.

Two labels were identified for the assessment criteria:

1. Perspective. A view on the requirement (for example: “Quality”)
2. Meta aspect. Describes the characteristics of a criterion in a single term

An interesting finding is that most criteria were assigned a perspective related to business (10) or quality (21), both with a strong relation to the requirements engineering process. This could indicate that reducing requirements uncertainty can be achieved not only by assessing requirements at the transfer point but also by strengthening the requirements engineering capability of the client.

Sub question 1 (RQ1.1): Which methods are relevant to measure initial requirements uncertainty?

In the literature review, a set of six measurement methods were found. These were validated in the Delphi study, in which one new method was added by the participants.

None of the methods found were based on exact scores but relied on estimations by the users, which could limit their usefulness in practice.

After Round 3 of the Delphi study, only one relevant measurement method remains, the method that was added by the participants in Round 2; all other measurement methods were regarded as not relevant. This could indicate a gap between scientific research and practice or the fact that the methods found in literature were not specific to the field of systems engineering and/or the civil engineering sector.

Sub question 2 (RQ1.2): Which factors that influence how much initial requirements uncertainty is acceptable are relevant?

In the literature review, two influencing factors were found but in the Delphi study another five were added by the participants. The fact that so many factors were found in the Delphi study, seems to indicate a gap between scientific research and daily practice. Perhaps not much research has been done in the specific field of systems engineering and/or the civil engineering sector.

The two factors found in literature as well as one added by the participants were marked as relevant in the Delphi study. These factors seem to be in the strategy or project management realm, more on a team than on an individual level. The factors are not exact science but could be used at the start of the requirements assessment process to create awareness or to define a strategy concerning requirements uncertainty.

6.2 Contribution

This thesis is intended to make both a scientific and a practical contribution. It has brought together previous research on requirements uncertainty criteria into one overview and combined this with the concepts of uncertainty measurement and acceptable uncertainty, a combination that was not found in existing literature. This resulted in a list of validated criteria, which is not only suitable for practical use but is also new information for scientific research.

The research adds value to the body of knowledge on requirements uncertainty / project risk management as the literature review shows that very little research on requirements assessment criteria has been done. The measurement methods and influencing factors that were added during the Delphi rounds are valuable to the body of knowledge as well as they have not been identified in literature before.

The research outcome adds value to the requirements review process that helps to start the design and development phase with less uncertainty and therefore helps to improve project performance, or at least reduce the risk on project performance. This is of great value to companies in the civil engineering sector as requirements review is a common process in this sector but lacks objective assessment criteria. Companies that operate in other sectors, perhaps even using other project methods than the V-model, might also benefit. This, of course, would need further research to validate this assumption.

6.3 Conclusions

Some conclusions can be drawn from the research, based on the findings described in chapter 5.

Conclusion 1: there seems to be a gap between scientific research and daily practice
In the Delphi study, the measurement methods found in literature were all considered to be non-relevant, the only relevant method remaining was the one added by the participants. A similar trend was observed with the influencing factors as a literature search only resulted in two factors while the participants added another five from their practical experience.

Maybe scientific literature is too much focussed on theoretical concepts, creating a distance between the science and the daily practice. On the other hand, in the world of civil engineering, systems engineering and its accompanying methods are gaining ground but the industry might still need to get used to formal methods.

The fact that there seems to be a gap between research and practice, however, does open up opportunities for future (applied) research at companies in the field of systems engineering and/or the civil engineering sector.

Conclusion 2: a Delphi study is highly suitable to get a quality response from a group of experts
A Delphi study was chosen for this research to get rich responses and a high response rate, which indeed happened: a response rate of 100% and many elaborate comments on the criteria. The goal of a Delphi study is to reach consensus, which was clearly observed from Round 2 onwards, resulting in an extra thirteen assessment criteria on which consensus was reached in Round 3.

The benefits of a Delphi study were observed during the research for this thesis:

- The number of individual responses is high (100%)
- Rich responses
- A movement towards consensus (measured using Variance)

The participants spent a serious amount of time filling in the questionnaires (up to 2.5 hours per round for some), not only because of the lengthy questionnaire but also because they felt involved. This conclusion is also a recommendation to use the Delphi method more often.

Conclusion 3: measuring uncertainty is more an art than a science
All measurement methods found in literature and the one added by the participants in the Delphi study dependent on some form of estimates, which are not unambiguous. Measuring uncertainty is therefore not an exact science but depends on the persons performing the measurement.

The impact of this conclusion is that it could still be hard to measure the effects of reducing requirements uncertainty or to define exactly what level of uncertainty is acceptable.

A group process to combine estimates might be introduced to make the outcome less dependent on the individual reviewer and make the estimate more reliable. An example of such a process is found in Agile development models: *planning poker*, a method where a group of experts each give points (on a Fibonacci scale) to a user story (a functional increment).

6.4 Recommendations

From the research findings, some important recommendations can be distilled.

Recommendation 1: use a subset of the validated assessment criteria
At the end of the Delphi rounds, consensus on relevance was reached on 37 criteria, validated by the research participants. This number of criteria is, according to the research findings, too large to handle within a project. However, the list of criteria can serve as input for a project or organisation to choose criteria that suit their needs most.

Recommendation 2: also look at the set of requirements, not just the individual ones
A finding in the Delphi study, and confirmed by the NEN-ISO/IEC/IEEE 29148:2011 norm (ISO/IEC/IEEE, 2011), is that looking at the set of requirements is also important for the assessment. There are relevant criteria that can easily be applied to the set or that only concern the set (like “There are no missing requirements”).

When using the list of assessment criteria, a subset could be applied to the whole list and another subset to the individual requirements, in line with recommendation 1.

Recommendation 3: interact with the client
An important finding of the research is that many of the assessment criteria seem to be related to the Quality and Business perspectives, which could indicate that reducing requirements uncertainty can also be achieved by strengthening the requirements engineering capability of the client. This is usually outside the sphere of influence of the contractor. The need for interaction with the client to gain certainty, however, is an important finding of the research and is a recommendation that accompanies the use of the assessment criteria.

6.5 Limitations

This research is constrained to a systems engineering context in the civil engineering sector in The Netherlands for projects that use the V-model project methodology. The assessment criteria, measurement methods and influencing factors were validated by a sample in the realm of Vialis, further validation by samples from similar companies would be required to increase validity.

As the sample group mainly consisted of participants in the role of systems architect, there is a risk that the study was looked at mainly through technical glasses, a sample with more evenly distributed roles might have resulted in a different outcome although some of the participants currently in the role of systems architect have had project lead roles as well.

As the study was done specifically in the field of systems engineering within the sector of civil engineering, the result is not applicable per se to other domains, like software engineering or health sciences.

An important limitation of the literature review is that it was constrained to scientific literature, accessible to the researcher. Books, blogs, other non-scientific literature and scientific papers behind pay walls were excluded but might contain valuable information. Some ISO norms were used to check findings in the Delphi study but were not used in the literature review.

6.6 Future research

The research for this thesis has revealed 37 relevant assessment criteria that were validated by a sample group within a specific context. Expanding the search for acceptance criteria to other domains and/or validating them by a sample group in different fields or sectors might reveal new criteria or make the list of criteria more generic. A logical first step, however, would be to take the same research in a similar domain but within the context of another company similar to Vialis to strengthen the validity.

To make requirements uncertainty measurement methods useful within the context of an organisation, more research is needed. A case study within an organisation might be a good starting point to test if a delta between initial and residual requirements uncertainty can be measured.

Another valuable research would be a study on the effects of reducing requirements uncertainty: is the project performance really better after performing an assessment using the criteria found in this thesis? If uncertainty can be properly measured, an existing project could be taken as example case. The delta between the initial uncertainty and the residual uncertainty (after assessing the requirements using the assessment criteria) could give an indication of the effects on project performance. A case study where the list of criteria is applied on a new project and the project performance is compared against previous and similar projects, would also be an interesting research.

A final recommendation for future research would be to check if the list of assessment criteria is useful for reducing residual requirements uncertainty as well. As this thesis is only concerned with initial requirements uncertainty, new criteria could emerge, and existing criteria evaluated differently.

7 References

Agile Alliance. (2018). Planning poker. Retrieved May 3, 2018, from https://www.agilealliance.org/glossary/poker/

Alami, A. (2016). Why do Projects Fail? *PM World Journal*, *5*, 1–9.

Anitha, P. C., Savio, D., & Mani, V. S. (2013). Managing requirements volatility while “Scrumming” within the V-Model. *2013 3rd International Workshop on Empirical Requirements Engineering, EmpiRE 2013 - Proceedings*, 17–23. https://doi.org/10.1109/EmpiRE.2013.6615211

Association for Computing Machinery. (n.d.). ACM Digital Library. Retrieved April 4, 2018, from https://dl.acm.org/

Australian Bureau of Statistics. (2018). Sample size calculator. Retrieved May 14, 2018, from http://www.nss.gov.au/nss/home.nsf/pages/Sample+size+calculator

Balaji, S., & Murugaiyan, M. S. (2012). Waterfall vs. V-Model vs. Agile: A comparative study on SDLC. *International Journal of Information Technology and Business Management*, *2*(1), 26–30.

Balaji, S., & Obaidy, M. Al. (2016). Project characteristics used for methodology selection to develop the software project. In *Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on* (pp. 3570–3573).

Beck, K., Beedle, M., Bennekum, A. van, Cockburn, A., Cunningham, W., Fowler, M., … Thomas, D. (2001). Principles behind the Agile manifesto. Retrieved May 10, 2018, from http://agilemanifesto.org/principles.html

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2017). *Cost-benefit analysis: concepts and practice*. Cambridge University Press.

Boehm, B. W. (1991). Software risk management: principles and practices. *IEEE Software*, *8*(1), 32–41.

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of Extension*, *50*(2), 1–5.

Brady, S. R. (2015). Utilizing and adapting the Delphi method for use in qualitative research. *International Journal of Qualitative Methods*, *14*(5), 1609406915621381.

Bryman, A., & Bell, E. (2015). *Business Research Methods* (Fourth). Oxford: Oxford University Press.

Clason, D. L., & Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education*, *35*, 4.

Collins, A., & Baccarini, D. (2004). Project success—a survey. *Journal of Construction Research*, *5*(2), 211–231.

Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (2006). Portfolio Management for New Product Development.

de Bruin, T., & Rosemann, M. (2007). Using the Delphi technique to identify BPM capability areas. *ACIS 2007 Proceedings*, 42.

Defense Acquisition University. (2017). System Requirements Review (SRR). Retrieved June 30, 2018, from https://www.dau.mil/acquipedia/Pages/ArticleDetails.aspx?aid=7edfe47b-6259-424e-8edb-af965f01a621

Department of Defense. (1985). *Military Standard: Technical Reviews and Audits for Systems, Equipment and Computer Software (MIL-STD-1521B)*. Department of Defense. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a285777.pdf

Dick, J., Hull, E., & Jackson, K. (2017). *Requirements Engineering* (Fourth). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-61073-3

Dow, J., & da Costa Werlang, S. R. (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica: Journal of the Econometric Society*, 197–204.

Ebert, C., & De Man, J. (2005). Requirements uncertainty: influencing factors and concrete improvements. In *Proceedings of the 27th international conference on Software engineering* (pp. 553–560).

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115.

Elsevier Inc. (n.d.). AIS eLibrary. Retrieved April 3, 2018, from http://aisel.aisnet.org/

Faber, M. H., & Stewart, M. G. (2003). Risk assessment for civil engineering facilities: critical overview and discussion. *Reliability Engineering & System Safety*, *80*(2), 173–184.

Feather, M. S., Cornford, S. L., Hicks, K. A., Kiper, J. D., & Menzies, T. (2008). A broad, quantitative model for making early requirements decisions. *IEEE Software*, *25*(2).

Firesmith, D. (2013). Using V Models for Testing. Retrieved February 23, 2018, from https://insights.sei.cmu.edu/sei_blog/2013/11/using-v-models-for-testing.html

Gemino, A., Sauer, C., & Reich, B. (2007). Estimating Risk in Information Technology Projects. *AMCIS*

2007 Proceedings, 106.
Ghanim, Y. (2015). The Adapted V-Model, A Practical Approach to Agile Testing. *Social Media and Publicity*, 83.
González-Araya, M. C., Rangel, L. A. D., Lins, M. P. E., & Gomes, L. F. A. M. (2002). Building the additive utility functions for CAD-UFRJ evaluation staff criteria. *Annals of Operations Research*, *116*(1–4), 271–288.
Grenn, M. W., Sarkani, S., & Mazzuchi, T. (2014). The Requirements Entropy Framework in Systems Engineering TT -. *Systems Engineering TA -*, *17*(4), 462–478.
Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature. *Journal of Informetrics*, *11*(3), 823–834. https://doi.org/https://doi.org/10.1016/j.joi.2017.06.005
Han, W.-M., & Huang, S.-J. (2007). An empirical analysis of risk components and performance on software projects. *Journal of Systems and Software*, *80*(1), 42–50.
Harzing, A.-W. (2018). Publish or Perish. Harzing. Retrieved from https://harzing.com/resources/publish-or-perish
Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569.
Hirshorn, S. R. (2017). *Expanded Guidance for NASA Systems Engineering. Volume 2: Crosscutting Topics, Special Topics, and Appendices*.
IBM. (2017). IBM SPSS 64bit. IBM.
International Council on Systems Engineering (INCOSE). (n.d.). What is Systems Engineering. Retrieved May 8, 2018, from https://www.incose.org/AboutSE/WhatIsSE
ISO/IEC. (2015). ISO/IEC/IEEE International Standard - Systems and software engineering -- System life cycle processes. *ISO/IEC/IEEE 15288 First Edition 2015-05-15*, 1–118. https://doi.org/10.1109/IEEESTD.2015.7106435
ISO/IEC/IEEE. (2011). NEN ISO/IEC/IEEE International Standard - Systems and software engineering -- Life cycle processes --Requirements engineering. *NEN-ISO/IEC/IEEE 29148:2011 En*, 1–94. https://doi.org/10.1109/IEEESTD.2011.6146379
Israel, G. D. (1992). *Determining sample size*. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS Gainesville.
Kalinowski, M., Sp'\inola, R. O., Conte, T., Prikladnicki, R., Méndez Fernández, D., & Wagner, S. (2015). Towards building knowledge on causes of critical requirements engineering problems.
Kamalrudin, M., & Sidek, S. (2015). A review on software requirements validation and consistency management. *International Journal of Software Engineering and Its Application*, *9*(10).
Keutel, M., Michalik, B., & Mellis, W. (2011). Coping with Requirements Uncertainty: An Interpretive Case Study.
Komoto, H., & Tomiyama, T. (2010). A system architecting tool for mechatronic systems design. *CIRP Annals-Manufacturing Technology*, *59*(1), 171–174.
Kossiakoff, A., Sweet, W. N., Seymour, S. J., & Biemer, S. M. (2011). *Systems engineering principles and practice* (Vol. 83). John Wiley & Sons.
Leau, Y. B., Loo, W. K., Tham, W. Y., & Tan, S. F. (2012). Software development life cycle AGILE vs traditional approaches. In *International Conference on Information and Network Technology* (Vol. 37, pp. 162–167).
Li, L., Harman, M., Wu, F., & Zhang, Y. (2017). The value of exact analysis in requirements selection. *IEEE Transactions on Software Engineering*, *43*(6), 580–596.
Mahnič, V., & Hovelja, T. (2012). On using planning poker for estimating user stories. *Journal of Systems and Software*, *85*(9), 2086–2095.
Mateen, A., Tabassum, M., & Rehan, A. (2017). Combining agile with traditional V model for enhancement of maturity in software Development. *arXiv Preprint arXiv:1702.00126*.
Mathur, S., & Malik, S. (2010). Advancements in the V-Model. *International Journal of Computer Applications*, *1*(12), 29–34.
Mc Hugh, M., Cawley, O., McCaffcry, F., Richardson, I., & Wang, X. (2013). An agile V-model for medical device software development to overcome the challenges with plan-driven software development lifecycles. *2013 5th International Workshop on Software Engineering in Health Care, SEHC 2013 - Proceedings*, 12–19. https://doi.org/10.1109/SEHC.2013.6602471
Mellis, W., Loebbecke, C., & Baskerville, R. (2010). Moderating effects of requirements uncertainty on flexible software development techniques. In *International Research Workshop on IT Project*

Management.
Merkhofer, M. W. (2012). *Decision Science and Social Risk Management: A Comparative Evaluation of cost-benefit analysis, decision analysis, and other formal decision-aiding approaches* (Vol. 2). Springer Science & Business Media.
Michalik, B., Keutel, M., & Mellis, W. (2014). Coping with Requirements Uncertainty--A Case Study of an Enterprise-Wide Record Management System. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on* (pp. 4024–4033).
Moløkken-Østvold, K., Haugen, N. C., & Benestad, H. C. (2008). Using planning poker for combining expert estimates in software projects. *Journal of Systems and Software*, *81*(12), 2106–2117.
Na, K.-S., Li, X., Simpson, J. T., & Kim, K.-Y. (2004). Uncertainty profile and software project performance: A cross-national comparison. *Journal of Systems and Software*, *70*(1), 155–163.
Nidumolu, S. R. (1996). Standardization, requirements uncertainty and software project performance. *Information & Management*, *31*(3), 135–150.
Nijkamp, P., & van Delft, A. (1977). *Multi-criteria analysis and regional decision-making* (Vol. 8). Springer Science & Business Media.
Nolan, A. J., Abrahão, S., Clements, P. C., & Pickard, A. (2011). Requirements uncertainty in a software product line. In *Software Product Line Conference (SPLC), 2011 15th International* (pp. 223–231).
Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632.
Nurmuliani, N., Zowghi, D., & Powell, S. (2004). Analysis of requirements volatility during software development life cycle. In *Software Engineering Conference, 2004. Proceedings. 2004 Australian* (pp. 28–37).
O’Donoghue, T., & Somerville, J. (2018). Modeling Risk Aversion in Economics. *Journal of Economic Perspectives*, *32*(2), 91–114.
Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, *42*(1), 15–29.
Palmquist, S., Lapham, M. A., Garcia-Miller, S., Chick, T. A., & Ozkaya, I. (2013). Parallel worlds: Agile and waterfall differences and similarities.
Project Management Institute. (2013). A guide to the project management body of knowledge (PMBOK guide). Newtown Square, Pennsylvania : Project Management Institute,. Retrieved from http://ebookcentral.proquest.com/lib/AUT/detail.action?docID=4603786
Pyster, A., Adcock, R., Ardis, M., Cloutier, R., Henry, D., Laird, L., … Wade, J. (2015). Exploring the Relationship between Systems Engineering and Software Engineering. *Procedia Computer Science*, *44*, 708–717. https://doi.org/https://doi.org/10.1016/j.procs.2015.03.016
Regnell, B., Svensson, R., & Wnuk, K. (2008). Can we beat the complexity of very large-scale requirements engineering? *Requirements Engineering: Foundation for Software Quality*, 123–128.
Rijkswaterstaat. (2017). Toelichting op Toetskader v1.23 - Veiligheidsbeambte wegtunnels Rijkswaterstaat. Retrieved from http://publicaties.minienm.nl/download-bijlage/91314/toelichting-op-toetskader-versie-1-23.pdf
Rijkswaterstaat, ProRail, Bouwend Nederland, Vereniging van Waterbouwers, NLingenieurs, & Uneto VNI. (2013). *Guideline for Systems Engineering within the civil engineering sector*. Retrieved from https://www.leidraadse.nl/assets/files/downloads/LeidraadSE/V3_EN/boek_Leidraad_SE_ENG_3_72dpi-def.pdf
Sadia, H., Beg, M. R., & Faisal, M. (2014). Requirement Risk Identification: A Practitioner’s Approach. *International Journal of Computer Applications*, *102*(15).
Salay, R., Chechik, M., Horkoff, J., & Di Sandro, A. (2013). Managing requirements uncertainty with partial models. *Requirements Engineering*, *18*(2), 107–128.
Scheithauer, D., & Forsberg, K. (2013). V-Model Views. *INCOSE International Symposium TA -*, *23*(1), 502–516.
Schmidt, R., Lyytinen, K., & Mark Keil, P. C. (2001). Identifying software project risks: An international Delphi study. *Journal of Management Information Systems*, *17*(4), 5–36.
Scott, W., Cook, S. C., & others. (2004). *A context-free requirements grammar to facilitate automatic assessment*. UniSA.
SEBoK. (2017a). Systems Engineering and Software Engineering.
SEBoK. (2017b). Systems Engineering Overview.
Seyedhosseini, S., & Keyghobadi, A. (2014). An integrated model for mechatronic products in agile manufacturing system. *Decision Science Letters*, *3*(4), 535–550.

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research*, *6*, 1–21.

Sommerville, I., & Ransom, J. (2005). An empirical study of industrial requirements engineering process assessment and improvement. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, *14*(1), 85–117.

Stephenson, Z., Attwood, K., & McDermid, J. (2011). Product-Line Models to Address Requirements Uncertainty, Volatility and Risk. In *Relating Software Requirements and Architectures* (pp. 111–131). Springer.

Sullivan, G. M., & Artino Jr, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542.

Suresh Kute, S., & Thorat, S. D. (2014). A Review on Various Software Development Life Cycle (SDLC) Models. *IJRCCT*, *3*(7), 776–781.

Terry Bahill, A., & Henderson, S. J. (2005). Requirements development, verification, and validation exhibited in famous failures. *Systems Engineering*, *8*(1), 1–14.

Thunnissen, D. P. (2003). Uncertainty classification for the design and development of complex systems. In *3rd annual predictive methods conference* (pp. 1–16).

Thunnissen, D. P. (2005). *Propagating and mitigating uncertainty in the design of complex multidisciplinary systems*. California Institute of Technology.

Toth, K. (2008). Selecting Software Estimating Techniques that Fit the Software Process. PNSQC.

Walden, D. D., Roedler, G. J., Forsberg, K. J., Hamelin, R. D., & Shortell, T. M. (2015). *INCOSE Systems Engineering Handbook* (4th ed.). Hoboken: Wiley.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & von Krauss, M. P. (2003). Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, *4*(1), 5–17.

Walton, M. A. (2002). *Managing uncertainty in space systems conceptual design using portfolio theory*. Massachusetts Institute of Technology.

Wheatcraft, L. S., Ryan, M. J., & Dick, J. (2016). On the use of attributes to manage requirements. *Systems Engineering*, *19*(5), 448–458.

Wohlin, C., & Aurum, A. (2005). What is important when deciding to include a software requirement in a project or release? In *Empirical Software Engineering, 2005. 2005 International Symposium on* (p. 10--pp).

Wulf, W. A. (1998). The urgency of engineering education reform. In *Proceedings, Conference on Realizing the New Paradigm for Engineering Education* (pp. 28–30).

Xu, L., & Yang, J.-B. (2001). *Introduction to multi-criteria decision making and the evidential reasoning approach*. Manchester School of Management Manchester.

Yadav, R. S. (2012). Improvement in the V-Model. *International Journal of Scientific & Engineering Research*, *3*(1), 2229–5518. Retrieved from http://www.ijser.org

Appendix A Questionnaire design Round 1

Fields on assessment criteria
The assessment criteria area of the questionnaire contains the following fields:

Provided:
- ID of the criterion.
- Criterion. Read only for the criteria found in the literature review.
- Perspective found in the literature review.
- Meta aspect found in the literature review.
- Researcher remarks on the criterion from the literature review.

Participant input:
- Relevance. The participant can indicate if criterion makes sense to use. This is measured on a scale from 1-5:
  - 1=Not relevant
  - 2=Slightly relevant
  - 3=Fairly relevant
  - 4=Relevant
  - 5=Highly relevant
- Perspective as seen by the participant. If the participant does not agree with the perspective from the literature review, the preferred perspective can be filled in here.
- Meta aspect as seen by the participant. If the participant does not agree with the meta aspect from the literature review, the preferred meta aspect can be filled in here.
- Comments by the participant on the criterion. This input has a maximum of 500 characters per criterion to keep the input manageable within the boundaries of the research.

The participants are allowed to add up to ten additional criteria each in the first round. This is limited to ten to keep the input manageable within the boundaries of the research and to make sure the most important criteria will be mentioned if there are many.

Fields on uncertainty measurement
The questionnaire area with the measurement types of uncertainty, contains the following fields:

Provided:
- ID of the measurement type.
- Short description. Read only for the measurement types found in the literature review. Detailed information will be provided in the background information.
- Type. Type of measurement method, added to support grouping.
- Researcher remarks on the measurement method from the literature review.

Participant input:
- Relevance. The participant can indicate if the measurement type makes sense. This is measured on a scale from 1-5:
  - 1=Not relevant
  - 2=Slightly relevant
  - 3=Fairly relevant
  - 4=Relevant
  - 5=Highly relevant
- Comments by the participant on the measurement type. This input has a maximum of 500 characters per measurement type to keep the input manageable within the boundaries of the research.

The participants are allowed to add up to five additional measurement types each in the first round. This is limited to five to keep the input manageable within the boundaries of the research and to make sure the most important criteria will be mentioned if there are many.

Fields on factors that influence acceptable uncertainty
The questionnaire area with the factors that influence acceptable requirements uncertainty, contains the following fields:

Provided:

- ID of the factor.
- Short description. Read only for the factor found in the literature review. Detailed information will be provided in the background information.
- Researcher remarks on the influence factor from the literature review.

Participant input:

- Relevance. The participant can indicate if the factor makes sense. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant
- Comments by the participant on the factor. This input has a maximum of 500 characters per factor to keep the input manageable within the boundaries of the research.

The participants are allowed to add up to five additional factors each in the first round. This is limited to five to keep the input manageable within the boundaries of the research and to make sure the most important factors will be mentioned if there are many.

Appendix B Questionnaire design Round 2

Fields on the assessment criteria
The assessment criteria area of the questionnaire contains the following fields:

Provided:
- ID of the criterion.
- Criterion
- Results from Round 1
    - Relevance based on the results of Round 1.
    - For each participant:
        - ID
        - Relevance
        - Perspective
        - Meta aspect
        - Comment
    - Researcher remarks from the literature review.

Participant input:
- Relevance. The participant can indicate (dis)agreement on the relevance from Round1 by specifying relevance again. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant
- Perspective. This is a choice from the perspectives from research and round 1 for the criterion.
- Meta aspect. This is a choice from the meta aspects from research and round 1 for the criterion.
- Comments. This input has a maximum of 500 characters per criterion to keep the input manageable within the boundaries of the research.

No additional criteria can be added.

Fields on uncertainty measurement
The questionnaire area with the measurement types of uncertainty, contains the following fields:

Provided:
- ID of the measurement type.
- Short description
- Type
- Results from Round 1
    - Relevance based on the results of Round 1.
    - For each participant:
        - ID
        - Relevance
        - Comment
    - Researcher remarks from the literature review.

Participant input:
- Relevance. The participant can indicate (dis)agreement on the relevance from Round 1 by specifying relevance again. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant
- Comments. This input has a maximum of 500 characters per measurement type to keep the input manageable within the boundaries of the research.

No additional measurement types can be added.

Fields on factors that influence acceptable uncertainty
The questionnaire area with the factors that influence acceptable requirements uncertainty, contains the following fields:

Provided:

- ID of the factor.
- Short description, recoded (if necessary) based on the results of Round 1. Detailed information will be provided in the background information.
- Results from Round 1
    - Relevance based on the results of Round 1.
    - For each participant:
        - ID
        - Relevance
        - Comment
    - Researcher remarks from the literature review.

Participant input:

- Relevance. The participant can indicate (dis)agreement on the relevance from Round 1 by specifying relevance again. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant
- Comments by the participant on the factor. This input has a maximum of 500 characters per factor to keep the input manageable within the boundaries of the research.

Appendix C Questionnaire design Round3

Fields on the assessment criteria
The assessment criteria area of the questionnaire contains the following fields:

Provided:
- ID of the criterion.
- Criterion
- Perspective, based on the results of Round 2.
- Meta aspect, based on the results of Round 2.
- Results from Round 2
    - Relevance based on the results of Round 2.
    - For each participant:
        - ID
        - Relevance
        - Perspective
        - Meta aspect
        - Comment

Participant input:
- Relevance. The participant can indicate (dis)agreement on the relevance from Round 2 by specifying relevance again. If the value is agreed with, the same score must be filled in. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant

No additional criteria can be added.

Fields on uncertainty measurement
The questionnaire area with the measurement types of uncertainty, contains the following fields:

Provided:
- ID of the measurement type.
- Short description.
- Results from Round 2
    - Relevance based on the results of Round 2.
    - For each participant:
        - ID
        - Relevance
        - Comment

Participant input:
- Relevance. The participant can indicate (dis)agreement on the relevance from Round 2 by specifying relevance again. If the value is agreed with, the same score must be filled in. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant

No additional measurement types can be added.

Fields on factors that influence acceptable uncertainty
The questionnaire area with the factors that influence acceptable requirements uncertainty, contains the following fields:

Provided:

- ID of the factor.
- Short description.
- Results from Round 2
    - Relevance based on the results of Round 2.
    - For each participant:
        - ID
        - Relevance
        - Comment

Participant input:

- Relevance. The participant can indicate (dis)agreement on the relevance from Round 2 by specifying relevance again. If the value is agreed with, the same score must be filled in. This is measured on a scale from 1-5:
    - 1=Not relevant
    - 2=Slightly relevant
    - 3=Fairly relevant
    - 4=Relevant
    - 5=Highly relevant

No additional factors can be added.

Appendix D Criteria from literature review

Table 26: "Criteria found in the article by Ebert & De Man (2005)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
C_1	Conflict of interest; commitments not maintained	Business	Supported by all stakeholders	The requirement is supported by all stakeholders	If there is doubt about interest / commitments (by stakeholders), this might undermine the business case for the requirement. The contractor should verify this with the client.
C_2	Unexpected dependencies between requirements	Risk	No unexpected dependencies	The requirement does not have unexpected dependencies with other requirements	The requirement themselves might be of good quality, but dependencies between them might cause conflicts, which is a project risk.
C_3	Cost/benefit of individual requirement unclear	Business	Cost/benefit	The cost/benefit of the requirement is clear	If there is doubt about the cost/benefit ratio, there might be a good reason to drop the requirement. The contractor should verify this with the client.
C_4	Incomplete requirements	Quality	Complete	The requirement is complete	If the requirement is incomplete, it is of poor quality*

* There is a thin line between the Risk and Quality perspective a s a requirement with a poor quality is also a risk to the project. The definition used for this research will be that the Quality perspective will be used for how well the requirement was written and the Risk perspective for the impact of the requirement (and its consequences) on the project.

Table 27: "Criteria found in the article by Kamalrudin & Sidek (2015)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_5$	Describes the correspondence of that specification with the real needs of the intended users […]	Business	Reflects the needs	The requirement reflects the needs of the stakeholders[18]	If the requirement does not reflect the needs of the stakeholders, it should probably be discarded
$\mathcal{C}_6$	Implies that all customer's needs will be met when the system is constructed.	Business	Reflects the needs	The requirement reflects all the needs of the stakeholders	If the requirement does not reflect the needs of the stakeholders, it should probably be discarded
$\mathcal{C}_7$	A requirement must have all relevant components	Quality	Complete	The requirement contains all relevant components	What "relevant" components are might be project specific. A comprehensive list of 44 requirement attributes is specified in the paper "On the use of attributes to manage requirements" by Wheatcraft, Ryan & Dick (2016)
$\mathcal{C}_8$	It specifies required behaviour and output for all possible states under all possible constraints	Quality	Complete	The requirement specifies required behaviour and output for all possible states under all possible constraints	It could depend on the project if it is necessary before the design phase to specify all possible states.
$\mathcal{C}_9$	No two or more requirements in a specification contradict each other	Risk	Consistency	The requirement does not contradict another requirement	Contradicting requirements might lead to a wrong implementation
$\mathcal{C}_{10}$	Words and terms have the same meaning throughout the requirement's specifications (consistent	Quality	Consistency	The requirement uses consistent terminology	Inconsistent terminology could lead to wrong interpretations

[18] The word "users" is replaced with "stakeholders" to make the criterion more generic

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
	use of terminology)				
$\mathcal{C}_{11}$	Requirement uses terms in a manner consistent with their specified meanings	Quality	Consistency	The requirement uses appropriate terminology	Wrong terminology could lead to wrong interpretations
$\mathcal{C}_{12}$	Requirement should be understood precisely in the same way by every person who reads it	Quality	Consistency	The requirement can be understood the same by all readers	This is very subjective but could be made objective by looking at consistent and appropriate terminology
$\mathcal{C}_{13}$	Requirements in the document should not conflict	Risk	Consistency	The requirement does not conflict with other requirements	Conflicting requirements can lead to a problem with verification
$\mathcal{C}_{14}$	Consistent specification exists when there is a computational model for its implementatio n	Business	Consistency	The requirements are consistent when a computational model for the implementation exist	A computational model requires consistency to give valid results
$\mathcal{C}_{15}$	[…] the specification is valid when it satisfies the user requirements	Business	Reflects the needs	The requirement reflects the needs of the stakeholders	If the requirement does not reflect the needs of the stakeholders, it should probably be discarded

Table 28: "Criteria found in the article by Scott et al (2004)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{16}$	All requirements annotated for relative necessity (Must have, desirable, optional)	Business	Necessity	The requirement is annotated for relative necessity (Must have, desirable, optional)	If requirements are not annotated for relative necessity, it might occur that important requirements are not fulfilled by the system when the project comes under pressure, but less important ones are
$\mathcal{C}_{17}$	For each requirement, all requirement parts are necessary and sufficient	Quality	Complete	The requirement is complete and does not contain unnecessary information	What "complete" means might be project specific. A comprehensive list of 44 requirement

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
					attributes is specified in the paper “On the use of attributes to manage requirements” by Wheatcraft, Ryan & Dick (2016)
$\mathcal{C}_{18}$	No two requirements are in conflict with each other	Risk	Consistency between requirements	The requirement does not conflict with other requirements	Contradicting requirements might lead to a wrong implementation
$\mathcal{C}_{19}$	For each requirement, there is no conflict between any of its parts	Quality	Consistency within itself	Parts of the requirement do not conflict with each other	Contradicting requirements might lead to a wrong implementation
$\mathcal{C}_{20}$	No requirement is in conflict with a Stakeholder’s needs statement	Business	Consistency with needs	The requirement reflects the needs of the stakeholders	If the requirement does not reflect the needs of the stakeholders, it should probably be discarded
$\mathcal{C}_{21}$	For each requirement, all requirement parts have just the right information (accurate to the right level)	Quality	Correct	All parts of the requirement have the right information (accurate to the right level)	It’s not clear what “the right level” actually means
$\mathcal{C}_{22}$	Requirements do not unnecessarily constrain the design	Risk	Design Independent	The requirement does not unnecessarily constrain the design	If the requirement is concerned with the “how”, it might lead to a suboptimal design
$\mathcal{C}_{23}$	At least one system could be built to meet the requirements	Risk	Feasible	It is possible to build a system that meets the requirement	If no system can be built to meet the requirements, the requirement should be discarded as it will cause a great risk to the contractor
$\mathcal{C}_{24}$	Requirements are organized according to a document standard	Quality	Organized and Formatted	The requirement is organized according to a document standard	The document standard can be company or project specific
$\mathcal{C}_{25}$	Requirements are electronically stored to allow	Quality	Electronically Storable	The requirement is electronically stored	It would be impossible to work without an electronically stored

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
	basic modification				requirement in a digital work place
$\mathcal{C}_{26}$	For each requirement, a procedure can be found to verify that the system meets the requirements	Risk	Testable	For the requirement, a procedure can be found to verify that the system meets the requirement	If there is no procedure to test the requirement, it will probably not be tested
$\mathcal{C}_{27}$	For each basic requirement, there is a link to the Stakeholder's need statement	Business	Traceable to need	The requirement links to the needs statements of the stakeholders	If there is no link to the needs statements of the stakeholders, the need might not be fully understood
$\mathcal{C}_{28}$	Each requirement has a link to its parent in the Informal Requirements Document	Risk	Traceable to parent requirement	The requirement contains a link to its parent requirement (if any)	If there is no link to the parent requirement, verification of the parent requirement by its children could be difficult to prove
$\mathcal{C}_{29}$	Each requirement has exactly one interpretation to all reviewers	Quality	Unambiguous	The requirement has exactly one interpretation to all reviewers	When more than one person is concerned with reviewing the requirement, collaboration is required to assess this criterion
$\mathcal{C}_{30}$	Each requirement can be comprehended by its viewers	Quality	Understandable	The requirement can be comprehended by its viewers	When more than one person is concerned with reviewing the requirement, collaboration is required to assess this criterion

Table 29: "Criteria found in the article by Grenn, Sarkani and Mazzuchi (2014)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{31}$	The requirement can be met within the cost and schedule constraints of the program.	Business	Realistic	The requirement can be met within the cost and schedule constraints of the program.	If the requirement cannot be met within the cost and schedule of the program, there is an issue with the business case
$\mathcal{C}_{32}$	The requirement	Business	Correct	The requirement represents	Necessity indicates this is a

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
	represents something necessary for the system to be built and satisfaction of some need.			something necessary for the system to be built and satisfaction of some need.	business perspective
$\mathcal{C}_{33}$	The stakeholders can easily comprehend the meaning of the requirement with minimal explanation	Quality	Understandable	The stakeholders can easily comprehend the meaning of the requirement with minimal explanation	Review by multiple stakeholders would be required to verify this
$\mathcal{C}_{34}$	The requirement captures everything that is intended with not to-be-determined or to-be-resolved	Quality	Complete	The requirement captures everything that is intended with not to-be-determined or to-be-resolved	
$\mathcal{C}_{35}$	The requirement is written in such a way that enables cost effective verification that the system fulfils it	Quality	Verifiable	The requirement is written in such a way that enables cost effective verification that the system fulfils it	This might seem a Business perspective at first, but the criterion is about *how* the requirement is written and is therefore more quality related
$\mathcal{C}_{36}$	The requirement is necessary to build the right system that meets user/customer needs	Business	Validated	The requirement is necessary to build the right system that meets user/customer needs	Necessity indicates this is a business perspective
$\mathcal{C}_{37}$	The requirement is allocated and linked to all other supporting requirements at higher and lower levels.	Risk	Traced/traceable	The requirement is allocated and linked to all other supporting requirements at higher and lower levels.	This could also be a Quality perspective, but Risk was chosen as it is more about the relation with other requirements that the inner parts of the requirement
$\mathcal{C}_{38}$	The requirement has only one possible interpretation	Quality	Unambiguous	The requirement has only one possible interpretation	Review by multiple stakeholders would be

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
					required to verify this
$\mathcal{C}_{39}$	The requirement is not in conflict with any other requirement	Risk	Consistent	The requirement is not in conflict with any other requirement	Conflicting requirements pose a risk to the project
$\mathcal{C}_{40}$	The requirement numeric quantities are accurate and specified to an appropriate level.	Quality	Precise	The requirement numeric quantities are accurate and specified to an appropriate level.	What the "appropriate" level is, is not defined. This might be set as a project specific level.
$\mathcal{C}_{41}$	The requirement communicates the information using as few words as possible.	Quality	Concise	The requirement communicates the information using as few words as possible.	This could lead to discussions as each reviewer might have his own opinion on how many words are enough
$\mathcal{C}_{42}$	The requirement enables flexibility in the design process by defining what is necessary, not how it should be designed or implemented.	Quality	Design independent	The requirement enables flexibility in the design process by defining what is necessary, not how it should be designed or implemented.	This means the requirement should be concerned with the black box, with the behaviour on the outside, not the inside. However, it is likely that when it comes to standardisation and interfacing, there might be requirements that do set limits on the implementation.
$\mathcal{C}_{43}$	The requirement necessity and relative importance are documented.	Quality	Annotated	The requirement necessity and relative importance are documented.	Necessity is a Business perspective, but stating that the requirement should mention it, is quality related.
$\mathcal{C}_{44}$	The requirement information is not stated more than once without a good reason to do so.	Risk	Not redundant	The requirement information is not stated more than once without a good reason to do so.	Stating the same information in different places might cause discrepancies, which are a risk to the project.

Table 30: "Criteria found in the article by Hirshorn (2017)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{45}$	Are the requirements clear and unambiguous?	Quality	Clarity	The requirement is clear and unambiguous	If the requirement is not clear and unambiguous, this might lead to discussions with the client about the implementation
$\mathcal{C}_{46}$	Are the requirements concise and simple?	Quality	Clarity	The requirement is concise and simple	If the requirement is not concise and simple, it will take more time to interpret correctly
$\mathcal{C}_{47}$	Do the requirements express only one thought per requirement statement, a stand-alone statement as opposed to multiple requirements in a single statement, or a paragraph that contains both requirements and rationale?	Risk	Clarity	The requirement does not contain multiple statements	If the requirement contains multiple statements, it is harder to break down in work packages and to verify and test with single test cases
$\mathcal{C}_{48}$	Does the requirement statement have one subject and one predicate?	Quality	Clarity	The requirement has one subject and one predicate	If the requirement has more than one subject or predicate, it probably consists of multiple requirements and should be split
$\mathcal{C}_{49}$	Are requirements stated as completely as possible? Have all incomplete requirements been captured as TBDs or TBRs and a complete listing of them maintained with the requirements?	Quality	Completeness	The requirement is complete.	The item also concerns an aspect on the set of requirements
$\mathcal{C}_{50}$	Are any requirements missing?	Risk	Completeness	There are no missing requirements	This concerns the set of requirements, not the individual requirement

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{51}$	Have all assumptions been explicitly stated?	Risk	Completeness	All assumptions have been explicitly stated	This criterion is needed to make sure the requirement is implemented properly
$\mathcal{C}_{52}$	Are all requirements at the correct level (e.g., system, segment, element, subsystem)?	Risk	Compliance	The requirement is at the correct level	If the requirement is at the wrong level, this might pose a risk if cannot be fulfilled at that level
$\mathcal{C}_{53}$	Are requirements free of implementation specifics? (Requirements should state what is needed, not how to provide it.)	Risk	Compliance	The requirement is not implementation specific	If the requirement is concerned with the “how”, the consequence might be that a suboptimal implementation follows
$\mathcal{C}_{54}$	Are requirements free of descriptions of operations?	Quality	Compliance	The requirement is free of descriptions of operations	The same remarks as for c_{53} apply here
$\mathcal{C}_{55}$	Are requirements free of personnel or task assignments?	Quality	Compliance	The requirement is free of personnel or task assignments	The same remarks as for c_{53} apply here
$\mathcal{C}_{56}$	Are the requirements stated consistently without contradicting themselves or the requirements of related systems?	Risk	Consistency	The requirement is stated consistently without contradicting itself or the requirements of related systems	Conflicting requirement are a risk to the project as they might lead to discussions or wrong implementations
$\mathcal{C}_{57}$	Is the terminology consistent with the user and sponsor’s terminology? With the project glossary?	Risk	Consistency	The requirement’s terminology is consistent with the user and sponsor’s terminology and the project glossary	This could also be a Quality perspective but inconsistent terminology could lead to wrong interpretations
$\mathcal{C}_{58}$	Is the terminology consistently used throughout	Risk	Consistency	The requirement’s terminology is consistent	This could also be a Quality perspective but inconsistent

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
	the document? Are the key terms included in the project's glossary?			throughout the document and key terms are included in the glossary.	terminology could lead to wrong interpretations
$\mathcal{C}_{59}$	Are all requirements needed?	Business	Traceability	The requirement is needed	An unnecessary requirement uses valuable resources
$\mathcal{C}_{60}$	Is each requirement necessary to meet the parent requirement?	Business	Traceability	The requirement is needed to meet a parent requirement	An unnecessary requirement uses valuable resources
$\mathcal{C}_{61}$	Is each requirement a needed function or characteristic?	Business	Necessity	The requirement is a needed function or characteristic	An unnecessary requirement uses valuable resources
$\mathcal{C}_{62}$	Are all requirements bidirectionally traceable to higher-level requirements or mission or system-of-interest scope (i.e., need(s), goals, objectives, constraints, or concept of operations)?	Business	Traceability	The requirement is bidirectionally traceable to a higher-level requirement or mission or system-of-interest scope (i.e., need(s), goals, objectives, constraints, or concept of operations)	If the requirement cannot be traced, it might not be needed
$\mathcal{C}_{63}$	Is each requirement stated in such a manner that it can be uniquely referenced (e.g., each requirement is uniquely numbered) in subordinate documents?	Risk	Traceability	The requirement is stated in such a manner that it can be uniquely referenced in subordinate documents	A requirement that cannot be uniquely referenced, might be lost or wrongly interpreted
$\mathcal{C}_{64}$	Is each requirement correct?	Quality	Correctness	The requirement is correct	It is not clear how "correct" is defined
$\mathcal{C}_{65}$	Is each stated assumption correct?	Quality	Correctness	The stated assumptions are correct	It is not clear how "correct" is defined
$\mathcal{C}_{66}$	Are the requirements technically feasible?	Risk	Correctness	The requirements technically feasible	If the requirement is not technically feasible, this will be a risk to the project as it would be impossible to

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
					fulfil the requirement
$\mathcal{C}_{67}$	Are all described functions necessary and together sufficient to meet mission and system goals and objectives	Business	Necessity	All described functions are necessary and together sufficient to meet mission and system goals and objectives	An unnecessary function uses valuable resources
$\mathcal{C}_{68}$	Are all required performance specifications and margins listed?	Risk	Performance	All required performance specifications and margins are listed	This goes beyond an individual requirement but concerns the whole set of requirements
$\mathcal{C}_{69}$	Is each performance requirement realistic?	Risk	Performance	The performance requirement is realistic	If not realistic, this will be a risk to the project as it would be impossible to fulfil the requirement
$\mathcal{C}_{70}$	Are the tolerances overly tight?	Risk	Performance	Tolerances mentioned in the requirement are not overly tight	If too tight, this will be a risk to the project as it would be impossible to fulfil the requirement
$\mathcal{C}_{71}$	Are the tolerances defendable and cost-effective?	Business	Cost/Benefit	Tolerance mentioned in the requirement are defendable and cost-effective	If not, they might not be needed as implementing them would take up resources
$\mathcal{C}_{72}$	Are all external interfaces clearly defined?	Risk	Interfaces	All external interfaces are clearly defined	If all external interfaces are not clearly defined, there is a risk of rework or a wrong implementation
$\mathcal{C}_{73}$	Are all internal interfaces clearly defined?	Risk	Interfaces	All internal interfaces are clearly defined	If all internal interfaces are not clearly defined, there is a risk of rework or a wrong implementation
$\mathcal{C}_{74}$	Are all interfaces necessary, sufficient, and consistent with each other?	Risk	Interfaces	All interfaces are necessary, sufficient, and consistent with each other	An unnecessary interface uses valuable resources
$\mathcal{C}_{75}$	Have the requirements for maintainability of the system been specified	Risk	Maintainability	The requirements concerning maintainability of the system	If not measurable, there is a risk of rework or a wrong implementation

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
	in a measurable, verifiable manner?			have been specified in a measurable, verifiable manner	
$\mathcal{C}_{76}$	Are requirements written so that ripple effects from changes are minimized (i.e., requirements are as weakly coupled as possible)?	Risk	Consistency	The requirement is as weakly coupled as possible to minimize ripple effects from changes	If a change of a requirement causes a ripple-effect, there is a high risk of scope change or rework throughout the system
$\mathcal{C}_{77}$	Are clearly defined, measurable, and verifiable reliability requirements specified?	Risk	Reliability	There are clearly defined, measurable, and verifiable reliability requirements specified	If no reliability requirements are specified, there is a risk that these might pop up during testing and cause discussion and rework
$\mathcal{C}_{78}$	Are there error detection, reporting, handling, and recovery requirements?	Risk	Reliability	There are error detection, reporting, handling, and recovery requirements	If no error detection, reporting, handling, and recovery requirements are specified, there is a risk that these might pop up during testing and cause discussion and rework
$\mathcal{C}_{79}$	Are undesired events considered and their required responses specified?	Risk	Reliability	There are requirements concerning undesired events and their response.	If there are no requirements concerning undesired events and their response, there is a risk that these might pop up during testing and cause discussion and rework
$\mathcal{C}_{80}$	Are there error detection, reporting, handling, and recovery requirements?	Risk	Reliability	There are requirements concerning error detection, reporting, handling, and recovery	If there are no requirements concerning error detection, reporting, handling, and recovery, there is a risk that these might pop up during testing and

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
					cause discussion and rework
$\mathcal{C}_{81}$	Do these requirements adequately address the survivability after a software or hardware fault of the system from the point of view of hardware, software, operations, personnel and procedures?	Risk	Reliability	There are requirements concerning the survivability after a software or hardware fault of the system from the point of view of hardware, software, operations, personnel and procedures	If there are no requirements concerning the survivability, there is a risk that these might pop up during testing and cause discussion and rework
$\mathcal{C}_{82}$	Can the system be tested, demonstrated, inspected, or analyzed to show that it satisfies requirements?	Risk	Verifiability/Test ability	The requirement can be verified in the system	If the requirement cannot be verified, the client might not sign off for acceptance
$\mathcal{C}_{83}$	Can the system be tested, demonstrated, inspected, or analyzed to show that it satisfies requirements at the level of the system at which the requirement is stated?	Risk	Verifiability/Test ability	The requirement can be verified at the correct level in the system	If the requirement cannot be verified at the right level, the client might not sign off for acceptance
$\mathcal{C}_{84}$	Does a means exist to measure the accomplishment of the requirement and verify compliance?	Risk	Verifiability/Test ability	The accomplishment of the requirement can be measured, and its compliancy verified.	If the accomplishment of the requirement cannot be measured and is compliance not verified, there is a risk of discussions on the scope and deliverable with the client
$\mathcal{C}_{85}$	Can the criteria for verification be stated?	Risk	Verifiability/Test ability	The criteria for verification can be stated	If there are no criteria for verification, there is a risk of discussions with the client about signing off the requirement's implementation

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{86}$	Are the requirements stated precisely to facilitate specification of system test success criteria and requirements?	Quality	Verifiability/Testability	The requirement is stated precisely to facilitate specification of system test success criteria	This is a Quality perspective as it defines an element of how the requirement is written but can lead to a risk if it's not
$\mathcal{C}_{87}$	Are the requirements free of unverifiable terms (e.g., flexible, easy, sufficient, safe, ad hoc, adequate, accommodate, user-friendly, usable, when required, if required, appropriate, fast, portable, light-weight, small, large, maximize, minimize, sufficient, robust, quickly, easily, clearly, other "ly" words, other "ize" words)?	Quality	Verifiability/Testability	The requirement is free of unverifiable terms (e.g., flexible, easy, sufficient, safe, ad hoc, adequate, accommodate, user-friendly, usable, when required, if required, appropriate, fast, portable, light-weight, small, large, maximize, minimize, sufficient, robust, quickly, easily, clearly, other "ly" words, other "ize" words)	These terms are open for different interpretations
$\mathcal{C}_{88}$	Where applicable, are "don't care" conditions truly "don't care"? ("Don't care" values identify cases when the value of a condition or flag is irrelevant, even though the value may be important for other cases.)	Risk	Data Usage	In the requirement, "don't care" conditions are truly "don't care" ("Don't care" values identify cases when the value of a condition or flag is irrelevant, even though the value may be important for other cases.)	If "don't care" conditions are not truly "don't care", this might lead to wrong behaviour of the system
$\mathcal{C}_{89}$	Are "don't care" conditions values explicitly stated? (Correct identification of "don't care" values may improve a	Risk	Data Usage	In the requirement, don't care" conditions values are explicitly stated (Correct identification of	If don't care" conditions values are not explicitly stated, this might lead to wrong behaviour of the system

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
	design's portability.)			"don't care" values may improve a design's portability.)	

Table 31: "Criteria found in the article by Stephenson, Attwood and McDermid (2011)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{90}$	Unfinished requirement	Quality	Incompleteness	The requirement is complete	An incomplete requirement cannot be accepted
$\mathcal{C}_{91}$	Placeholders	Risk	Incompleteness	The requirement does not contain placeholders	Placeholders can be used to identify a gap, which means the requirement is not complete yet. Often, a placeholder is used to point information stored elsewhere, which makes it harder to make to fully understand the requirement or to verify its implementation as the other information might contain requirements as well.
$\mathcal{C}_{92}$	Missing counterpart	Business	Consistency	The requirement has counterpart requirements	In many cases, requirements come as a set. For example, there may be a start-up requirement for each mode of a system. Even with little domain knowledge, it should be apparent when part of the set is missing. (Stephenson et al., 2011)
$\mathcal{C}_{93}$	Under-specification	Risk	Ambiguity	The requirement does not leave options open	A requirement that leaves option open might cause scope discussions

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{94}$	Terminology	Risk	Ambiguity	The terminology in the requirement is well defined	If the terminology is not well defined, the requirement might not be interpreted properly and implemented the wrong way
$\mathcal{C}_{95}$	Syntactic structure	Risk	Ambiguity	The requirement can only be read in one plausible way	If the requirement can be read in more than one plausible way, it might not be interpreted properly and implemented the wrong way
$\mathcal{C}_{96}$	Incorrectness	Quality	Commitment	The information in the requirement is demonstrably correct	If the information is demonstrably incorrect, the requirement should be sent back into the requirements engineering process
$\mathcal{C}_{97}$	Over specification	Risk	Commitment	The requirement does not include more detail than necessary and does not give awkward or infeasible constraints	Too much detail or constraints might constrain required options in the design phase.
$\mathcal{C}_{98}$	Misplaced requirement	Quality	Mislabelling	The positioning of the requirement (section heading, informative context) does not conflicts with its contents	If the positioning is incorrect, it might take more time to find the correct information to interpret the requirement
$\mathcal{C}_{99}$	Mislabelled domain information	Quality	Mislabelling	The requirement does not just contain definitions	A requirement with just definitions does not require any action and is therefore obsolete.

Table 32: "Criteria found in the article by Sadia, Beg & Faisal (2014)"

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
$\mathcal{C}_{100}$	Poorly defined requirements	Quality	Properly defined	The requirement is properly defined	If the requirement is poorly defined, it is very hard to assess and should be send

#	Criterion	Perspective	Meta aspect	Rephrased	Review remarks
					back to the requirements engineering process
C_{101}	Ambiguous requirements	Quality	Unambiguous	The requirement is not ambiguous	If the requirement is not unambiguous, this might lead to discussions with the client about the implementation
C_{102}	Inadequate requirements	Quality	Adequate	The requirement is adequate	If the requirement is inadequate, it is very hard to assess and should be send back to the requirements engineering process
C_{103}	Impossible requirements	Risk	Possible	The requirement is possible to implement	If the requirement is impossible, it cannot be verified and will cause discussions about fulfilling the contractual obligations
C_{104}	Invalid requirements	Business	Valid	The requirement is valid	The requirement should not go against laws and regulations and should be within contractual scope.
C_{105}	Requirement not necessary	Business	Necessary	The requirement is necessary	If the requirement is not necessary, it will use resources that could better be used on other requirements
C_{106}	Requirement not verifiable	Risk	Verifiable	The requirement can be verified	If the requirement is not verifiable, it should probably not have been implemented. If not verifiable, this can cause discussions with the client.
C_{107}	Requirements not traceable to a parent	Business	Traceable to parent	The requirement can be traced to a parent requirement	If there is no link to the parent requirement, verification of the parent requirement by its children could be difficult to prove

Appendix E **Processed assessment criteria**

Table 33: "Processing criteria"

Original	Processed into	Original	Processed into	Original	Processed into	Original	Processed into
C_1	CP_2	C_{28}	CP_{45}	C_{55}	CP_{32}	C_{82}	CP_{52}
C_2	CP_{41}	C_{29}	CP_{16}	C_{56}	CP_{37}	C_{83}	CP_{53}
C_3	CP_3	C_{30}	CP_{16}	C_{57}	CP_{39}	C_{84}	CP_{62}
C_4	CP_{13}	C_{31}	CP_7	C_{58}	CP_{38}	C_{85}	CP_{63}
C_5	CP_1	C_{32}	CP_1	C_{59}	CP_1	C_{86}	CP_{33}
C_6	CP_1	C_{33}	CP_{22}	C_{60}	CP_5	C_{87}	CP_{30}
C_7	CP_{13}	C_{34}	CP_{36}	C_{61}	CP_{10}	C_{88}	CP_{64}
C_8	CP_{15}	C_{35}	CP_{23}	C_{62}	CP_9	C_{89}	CP_{65}
C_9	CP_{37}	C_{36}	CP_1	C_{63}	CP_{48}	C_{90}	CP_{13}
C_{10}	CP_{38}	C_{37}	CP_{45}	C_{64}	CP_{28}	C_{91}	CP_{36}
C_{11}	CP_{39}	C_{38}	CP_{16}	C_{65}	CP_{29}	C_{92}	CP_{66}
C_{12}	CP_{16}	C_{39}	CP_{40}	C_{66}	CP_{42}	C_{93}	CP_{16}
C_{13}	CP_{40}	C_{40}	CP_{24}	C_{67}	CP_8	C_{94}	CP_{39}
C_{14}	CP_{43}	C_{41}	CP_{14}	C_{68}	CP_{50}	C_{95}	CP_{16}
C_{15}	CP_1	C_{42}	CP_{25}	C_{69}	CP_{42}	C_{96}	CP_{28}
C_{16}	CP_4	C_{43}	CP_4	C_{70}	CP_{51}	C_{97}	CP_{13}
C_{17}	CP_{14}	C_{44}	CP_{46}	C_{71}	CP_{11}	C_{98}	CP_{34}
C_{18}	CP_{40}	C_{45}	CP_{16}	C_{72}	CP_{54}	C_{99}	CP_{35}
C_{19}	CP_{17}	C_{46}	CP_{26}	C_{73}	CP_{55}	C_{100}	CP_{28}
C_{20}	CP_1	C_{47}	CP_{27}	C_{74}	CP_{56}	C_{101}	CP_{16}
C_{21}	CP_{18}	C_{48}	CP_{27}	C_{75}	CP_{57}	C_{102}	CP_{28}
C_{22}	CP_{25}	C_{49}	CP_{13}	C_{76}	CP_{49}	C_{103}	CP_{42}
C_{23}	CP_{42}	C_{50}	CP_{47}	C_{77}	CP_{58}	C_{104}	CP_{12}
C_{24}	CP_{19}	C_{51}	CP_6	C_{78}	CP_{59}	C_{105}	CP_{10}
C_{25}	CP_{20}	C_{52}	CP_{18}	C_{79}	CP_{60}	C_{106}	CP_{52}
C_{26}	CP_{44}	C_{53}	CP_{25}	C_{80}	CP_{59}	C_{107}	CP_9
C_{27}	CP_{21}	C_{54}	CP_{31}	C_{81}	CP_{61}		

Table 34: "Processed criteria"

ID	Criterion	Perspective	Meta aspect	Remarks
CP_1	The requirement reflects the needs of the stakeholders	Business	Necessity	If the requirement does not reflect the needs of the stakeholders, it should probably be discarded
CP_2	The requirement is supported by all stakeholders	Business	Stakeholder support	If there is doubt about interest / commitments (by stakeholders), this might undermine the business case for the requirement. The contractor should verify this with the client.
CP_3	The cost/benefit of the requirement is clear	Business	Cost/Benefit	If there is doubt about the cost/benefit ratio, there might be a good reason to

ID	Criterion	Perspective	Meta aspect	Remarks
				drop the requirement. The contractor should verify this with the client.
CP_4	The requirement is annotated for relative necessity (Must have, desirable, optional)	Business	Necessity	If requirements are not annotated for relative necessity, it might occur that important requirements are not fulfilled by the system when the project comes under pressure, but less important ones are
CP_5	The requirement is needed to meet a parent requirement	Business	Traceability	An unnecessary requirement uses valuable resources
CP_6	All assumptions have been explicitly stated	Risk	Completeness	This criterion is needed to make sure the requirement is implemented properly
CP_7	The requirement can be met within the cost and schedule constraints of the program.	Business	Cost/Benefit	If the requirement cannot be met within the cost and schedule of the program, there is an issue with the business case
CP_8	All described functions are necessary and together sufficient to meet mission and system goals and objectives	Business	Necessity	Implementing an unnecessary function uses valuable resources. If together they are not sufficient, the might be a problem with the business case
CP_9	The requirement is bidirectionally traceable to a higher-level requirement or mission or system-of-interest scope (i.e., need(s), goals, objectives, constraints, or concept of operations)	Business	Traceability	If the requirement cannot be traced, it might not be needed
CP_{10}	The requirement is necessary	Business	Necessity	Implementing an unnecessary function uses valuable resources
CP_{11}	Tolerance mentioned in the requirement are defendable and cost-effective	Business	Cost/Benefit	If not, they might not be needed as implementing them would take up resources
CP_{12}	The requirement is valid	Business	Validity	The requirement should not go against laws and regulations and should be within contractual scope.
CP_{13}	The requirement contains all relevant components	Quality	Completeness	What “relevant” components are might be project specific. A comprehensive list of 44 requirement attributes is specified in the paper “On the use of attributes to manage requirements” by

ID	Criterion	Perspective	Meta aspect	Remarks
				Wheatcraft, Ryan & Dick (2016)
CP_{14}	The requirement does not contain unnecessary information	Quality	Completeness	Unnecessary information might cause confusion
CP_{15}	The requirement specifies required behaviour and output for all possible states under all possible constraints	Quality	Completeness	It could depend on the project if it is necessary before the design phase to specify all possible states.
CP_{16}	The requirement does not contain ambiguities and can be understood the same by all readers	Quality	Consistency	If the requirement is not clear and unambiguous, this might lead to discussions with the client about the implementation
CP_{17}	Parts of the requirement do not conflict with each other	Risk	Consistency	Contradicting information within the requirement might lead to a wrong implementation
CP_{18}	All parts of the requirement have the right information (accurate to the right level)	Quality	Consistency	This requires group discussion on the requirement as what is the right level, should be discussed
CP_{19}	The requirement is organized according to a document standard	Quality	Administration	The document standard can be company or project specific
CP_{20}	The requirement is electronically stored	Quality	Administration	It would be impossible to work without an electronically stored requirement in a digital work place
CP_{21}	The requirement links to the needs statements of the stakeholders	Quality	Traceability	If there is no link to the needs statements of the stakeholders, the need might not be fully understood
CP_{22}	The requirement can be comprehended with minimal explanation	Quality	Consistency	Review by multiple stakeholders would be required to verify this
CP_{23}	The requirement is written in such a way that enables cost effective verification that the system fulfils it	Quality	Verification	If not written properly, the requirement would require a lot of unnecessary costs to verify it
CP_{24}	The requirement's numeric quantities are accurate and specified to an appropriate level.	Quality	Consistency	What the “appropriate” level is, is not defined. This might be set as a project specific level.

ID	Criterion	Perspective	Meta aspect	Remarks
CP_{25}	The requirement enables flexibility in the design process by defining what is necessary, not how it should be designed or implemented.	Risk	Design independency	This means the requirement should be concerned with the black box, with the behaviour on the outside, not the inside. However, it is likely that when it comes to standardisation and interfacing, there might be requirements that do set limits on the implementation.
CP_{26}	The requirement is concise and simple	Quality	Consistency	If the requirement is not concise and simple, it will take more time to interpret correctly
CP_{27}	The requirement does not contain multiple statements	Quality	Consistency	If the requirement contains multiple statements, it is harder to break down in work packages and to verify and test with single test cases
CP_{28}	The requirement and its information is correct	Quality	Correctness	The requirement does not contain false information
CP_{29}	The stated assumptions are correct	Quality	Correctness	The requirement does not contain false assumptions
CP_{30}	The requirement is free of unverifiable terms (e.g., flexible, easy, sufficient, safe, ad hoc, adequate, accommodate, user-friendly, usable, when required, if required, appropriate, fast, portable, light-weight, small, large, maximize, minimize, sufficient, robust, quickly, easily, clearly, other “ly” words, other “ize” words)	Quality	Consistency	These terms are open for different interpretations
CP_{31}	The requirement is free of descriptions of operations	Risk	Design independency	If the requirement is concerned with the “how”, the consequence might be that a suboptimal implementation follows
CP_{32}	The requirement is free of personnel or task assignments	Risk	Design independency	If the requirement is concerned with the “how”, the consequence might be that a suboptimal implementation follows
CP_{33}	The requirement is stated precisely to facilitate specification of system test success criteria	Quality	Precision	This is a Quality perspective as it defines an element of how the requirement is written but can lead to a risk if it’s not
CP_{34}	The positioning of the requirement (section heading, informative	Quality	Consistency	If the positioning is incorrect, it might take more time to find the correct information to interpret the requirement

ID	Criterion	Perspective	Meta aspect	Remarks
	context) does not conflicts with its contents			
CP_{35}	The requirement does not just contain definitions	Quality	Consistency	A requirement with just definitions does not require any action and is therefore obsolete.
CP_{36}	The requirement does not contain placeholders	Quality	Consistency	Placeholders (like "TBD") can be used to identify a gap, which means the requirement is not complete yet. Often, a placeholder is used to point information stored elsewhere, which makes it harder to make to fully understand the requirement or to verify its implementation as the other information might contain requirements as well.
CP_{37}	The requirement does not contradict another requirement	Risk	Consistency	Contradicting requirements might lead to a wrong implementation
CP_{38}	The requirement uses consistent terminology	Quality	Consistency	Inconsistent terminology could lead to wrong interpretations
CP_{39}	The requirement uses appropriate terminology	Quality	Consistency	Wrong terminology could lead to wrong interpretations
CP_{40}	The requirement does not conflict with other requirements	Risk	Consistency	Conflicting requirements can lead to a problem with verification
CP_{41}	The requirement does not have unexpected dependencies with other requirements	Risk	Consistency	The requirement themselves might be of good quality, but dependencies between them might cause conflicts, which is a project risk.
CP_{42}	The requirement is (technically) feasible / realistic	Risk	Feasibility	If no system can be built to meet the requirements, the requirement should be discarded as it will cause a great risk to the contractor
CP_{43}	The requirements are consistent when a computational model for the implementation exist	Risk	Consistency	A computational model requires consistency to give valid results
CP_{44}	For the requirement, a procedure can be found to verify that the system meets the requirement	Risk	Testability	If there is no procedure to test the requirement, it will probably not be tested
CP_{45}	The requirement is allocated and linked to all other supporting requirements at higher and lower levels.	Business	Traceability	If there is no link to a parent requirement, verification of the parent requirement by its children could be difficult to prove.
CP_{46}	The requirement information is not stated	Risk	Consistency	Stating the same information in different places might cause

ID	Criterion	Perspective	Meta aspect	Remarks
	more than once without a good reason to do so.			discrepancies, which are a risk to the project.
CP_{47}	There are no missing requirements	Risk	Completeness	Missing requirements are a project risk as they might lead to an implementation that does not fulfill the needs of the stakeholders
CP_{48}	The requirement is stated in such a manner that it can be uniquely referenced in subordinate documents	Quality	Administration	A requirement that cannot be uniquely referenced, might be lost or wrongly interpreted
CP_{49}	The requirement is as weakly coupled as possible to minimize ripple effects from changes	Risk	Consistency	If a change of a requirement causes a ripple-effect, there is a high risk of scope change or rework throughout the system
CP_{50}	All required performance specifications and margins are listed	Risk	Performance	
CP_{51}	Tolerances mentioned in the requirement are not overly tight	Risk	Performance	If too tight, this will be a risk to the project as it would be impossible to fulfil the requirement
CP_{52}	The requirement can be verified in the system	Risk	Verification	If the requirement cannot be verified, the client might not sign off for acceptance
CP_{53}	The requirement can be verified at the correct level in the system	Risk	Verification	If the requirement cannot be verified at the right level, the client might not sign off for acceptance
CP_{54}	All external interfaces are clearly defined	Risk	Interfaces	If all external interfaces are not clearly defined, there is a risk of rework or a wrong implementation
CP_{55}	All internal interfaces are clearly defined	Risk	Interfaces	If all external interfaces are not clearly defined, there is a risk of rework or a wrong implementation
CP_{56}	All interfaces are necessary, sufficient, and consistent with each other	Risk	Interfaces	Implementing an unnecessary interface uses valuable resources in the project
CP_{57}	The requirements concerning maintainability of the system have been specified in a measurable, verifiable manner	Risk	Maintainability	If not measurable, there is a risk of rework or a wrong implementation
CP_{58}	There are clearly defined, measurable, and verifiable reliability requirements specified	Risk	Reliability	If no reliability requirements are specified, there is a risk that these might pop up during testing and cause discussion and rework

ID	Criterion	Perspective	Meta aspect	Remarks
CP_{59}	There are error detection, reporting, handling, and recovery requirements	Risk	Reliability	If no error detection, reporting, handling, and recovery requirements are specified, there is a risk that these might pop up during testing and cause discussion and rework
CP_{60}	There are requirements concerning undesired events and their response.	Risk	Reliability	If there are no requirements concerning undesired events and their response, there is a risk that these might pop up during testing and cause discussion and rework
CP_{61}	There are requirements concerning the survivability after a software or hardware fault of the system from the point of view of hardware, software, operations, personnel and procedures	Risk	Reliability	If there are no requirements concerning the survivability, there is a risk that these might pop up during testing and cause discussion and rework
CP_{62}	The accomplishment of the requirement can be measured, and its compliancy verified.	Risk	Verification	If the accomplishment of the requirement cannot be measured and is compliance not verified, there is a risk of discussions on the scope and deliverable with the client
CP_{63}	The criteria for verification can be stated	Risk	Verification	If there are no criteria for verification, there is a risk of discussions with the client about signing off the requirement's implementation
CP_{64}	In requirements, "don't care" conditions are truly "don't care" ("Don't care" values identify cases when the value of a condition or flag is irrelevant, even though the value may be important for other cases.)	Risk	Verification	If "don't care" conditions are not truly "don't care", this might lead to wrong behaviour of the system
CP_{65}	In the requirement, don't care" conditions values are explicitly stated (Correct identification of "don't care" values may improve a design's portability.)	Risk	Verification	If don't care" conditions values are not explicitly stated, this might lead to wrong behaviour of the system
CP_{66}	The requirement has counterpart requirements	Business	Consistency	In many cases, requirements come as a set. For example, there may be a start-up requirement for

ID	Criterion	Perspective	Meta aspect	Remarks
				each mode of a system. Even with little domain knowledge, it should be apparent when part of the set is missing.

Appendix F Requirement components

Some assessment criteria state that the requirement should contain all necessary components. A comprehensive list of 44 of these components is specified in the paper “On the use of attributes to manage requirements” by Wheatcraft, Ryan & Dick (2016).

Table 35: "Requirement components"

No.	Area	Component
1	Definition and intent	Rationale
2		SOI[19] Primary verification method
3		SOI Verification Approach
4		Parent Requirements
5		Source
6		Condition of Use
7		States and Modes
8	SOI Verification	SOI Verification Level
9		SOI Verification Phase
10		SOI Verification Results
11		SOI Verification Status
12	Maintainability of the requirement	Unique Identifier
13		Unique Name
14		Originator/Author
15		Date Requirement Entered
16		Owner
17		Stakeholders
18		Change Board
19		Change Status
20		Version Number
21		Approval Date
22		Date of Last Change
23		Stability
24		Responsible Person
25		Requirement Verification Status
26		Requirement Validation Status
27		Status of requirement
28		Status of implementation
29		Trace to Interface Definition
30		Trace to Peer Requirements
31		Priority
32		Criticality
33		Risk
34		Key Driving Requirement
35		Additional Comments
36		Type/Category
37	Applicability and reusability	Applicability
38		Region
39		Country
40		State/Province
41		Application
42		Market Segment
43		Business Unit

[19] System of Interest

No.	Area	Component
44		Business Line

Appendix G **Search statistics**

In Table 36, statistics on the search queries as produced by Publish or Perish can be found. The number of papers is limited due to the maximum of search results allowed by the Google Scholar API. When searching in Google Scholar, the results pages might indicate thousands of search results, but it is not possible to scroll further than the first one thousand results.

Table 36: "Search statistics"

ID	Papers	Citations	Cites_ Paper	Cites_ Author	h_ index	Cites_ Author_ Year	h_ coverage	year_ first	Year _ last
RQ1_1	434	3387	7.80	1659.10	30	8.33	61.9	1819	2018
RQ1_2	26	3256	125.23	1716.17	7	59.17	99.3	1989	2017
RQ1_3	437	10621	24.30	5647.30	38	106.55	84.4	1965	2018
RQ1_4	84	2169	25.82	1314.33	14	37.55	94.7	1983	2017
RQ1_5	980	23769	24.25	12459.85	59	259.58	76.8	1970	2018
RQ1_6	6	13	2.17	4.00	1	0.57	84.6	2011	2017
RQ1_7	15	28	1.87	14.50	3	1.03	78.6	2004	2017
RQ1_8	18	2163	120.17	1081.23	5	98.29	99.1	2007	2017
RQ1.1_1	980	25793	26.32	13040.35	69	217.33	72.5	1958	2018
RQ1.1_2	600	10393	17.32	5360.85	55	74.45	69.8	1946	2018
RQ1.2_1	117	6525	55.77	3424.20	18	122.29	95.9	1990	2017
RQ1.2_2	215	7800	36.28	4152.82	26	148.31	90.8	1990	2017
RQ1.2_3	191	7453	39.02	4052.12	24	115.77	93.2	1983	2017
RQ1.2_4	438	10635	24.28	5655.50	38	106.70	84.3	1965	2018

Appendix H Data analysis in SPSS

Table 37: "SPSS variables for the assessment criteria"

Name	Type	Width	Decimals	Label	Values	Measure
Round	Numeric	1	0	Round	None	Nominal
ParticipantID	Numeric	2	0	Participant ID	None	Nominal
CP_ID	Numeric	3	0	Criterion ID	None	Nominal
IsNew	Numeric	1	0	Is new	0=No 1=Yes	Scale
CP_PerspectiveOriginal	String	50	0	Original perspective	None	Nominal
CP_MetaAspectOriginal	String	50	0	Original meta aspect	None	Nominal
CP_Relevance	Numeric	8	2	Relevance score	1=Not relevant 2=Slightly relevant 3=Moderately relevant 4=Relevant 5=Highly relevant	Ordinal
CP_Perspective	String	50	0	Perspective	None	Nominal
CP_MetaAspect	String	50	0	Meta aspect	None	Nominal

Table 38: "SPSS variables for the measurement methods"

Name	Type	Width	Decimals	Label	Values	Measure
Round	Numeric	1	0	Round	None	Nominal
ParticipantID	Numeric	2	0	Participant ID	None	Nominal
M_ID	Numeric	3	0	Measurement ID	None	Nominal
IsNew	Numeric	1	0	Is new	0=No 1=Yes	Scale
M_Relevance	Numeric	8	2	Relevance score	1=Not relevant 2=Slightly relevant 3=Moderately relevant 4=Relevant 5=Highly relevant	Ordinal

Table 39: "SPSS variables for the influencing factors"

Name	Type	Width	Decimals	Label	Values	Measure
Round	Numeric	1	0	Round	None	Nominal
ParticipantID	Numeric	2	0	Participant ID	None	Nominal
F_ID	Numeric	3	0	Factor ID	None	Nominal
IsNew	Numeric	1	0	Is new	0=No 1=Yes	Scale
M_Relevance	Numeric	8	2	Relevance score	1=Not relevant 2=Slightly relevant 3=Moderately relevant 4=Relevant 5=Highly relevant	Ordinal

Appendix I **Statistical analysis data on assessment criteria Round 1**

The assessment criteria were processed in SPSS. To create the frequency tables, the file was split by criterion ID:

```
SORT CASES  BY CP_ID
SPLIT FILE LAYERED BY CP_ID
```

Only data from Round 1 was entered when analyzing, so no filter on cases was needed. Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=CP_Relevance
  /STATISTICS=RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 40: "Frequencies on assessment criteria Round 1"

Criterion ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
1	11	0	5.00	.218	1	4	5
2	11	0	3.00	1.291	3	1	4
3	11	0	3.00	.673	3	2	5
4	10	1	3.00	1.956	4	1	5
5	10	1	3.00	1.433	4	1	5
6	11	0	4.00	1.673	4	1	5
7	11	0	4.00	.800	2	3	5
8	11	0	4.00	1.691	4	1	5
9	11	0	4.00	1.655	3	2	5
10	11	0	4.00	1.655	4	1	5
11	11	0	3.00	2.018	4	1	5
12	11	0	5.00	1.455	4	1	5
13	11	0	3.00	1.600	4	1	5
14	11	0	4.00	.873	3	2	5
15	11	0	4.00	1.655	4	1	5
16	11	0	5.00	.655	2	3	5
17	11	0	5.00	.855	3	2	5
18	11	0	4.00	1.255	4	1	5
19	11	0	3.00	.655	3	1	4
20	11	0	3.00	1.564	4	1	5
21	11	0	3.00	1.455	4	1	5
22	10	1	3.50	1.344	4	1	5
23	11	0	4.00	1.418	4	1	5
24	11	0	4.00	.618	2	3	5
25	11	0	4.00	2.255	4	1	5
26	11	0	4.00	1.073	4	1	5
27	11	0	4.00	1.218	3	2	5
28	11	0	5.00	1.418	4	1	5
29	11	0	5.00	2.400	4	1	5
30	11	0	4.00	.691	3	2	5
31	11	0	3.00	1.400	4	1	5
32	11	0	3.00	1.291	4	1	5
33	11	0	4.00	1.255	4	1	5
34	11	0	4.00	.855	3	2	5
35	11	0	3.00	1.255	3	2	5
36	11	0	4.00	1.000	3	2	5
37	11	0	5.00	.455	2	3	5
38	9	2	4.00	.250	2	3	5

Criterion ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
39	11	0	4.00	1.400	4	1	5
40	11	0	5.00	.255	1	4	5
41	11	0	4.00	.600	2	3	5
42	11	0	4.00	.564	2	3	5
43	11	0	3.00	.873	3	1	4
44	11	0	4.00	1.255	4	1	5
45	11	0	3.00	.655	2	3	5
46	11	0	3.00	.855	3	2	5
47	11	0	4.00	1.200	3	2	5
48	11	0	5.00	2.000	4	1	5
49	11	0	3.00	1.291	3	1	4
50	10	1	3.50	1.122	3	2	5
51	11	0	3.00	1.400	4	1	5
52	11	0	5.00	.473	2	3	5
53	10	1	3.50	1.789	4	1	5
54	10	1	5.00	.500	2	3	5
55	11	0	3.00	1.618	4	1	5
56	11	0	4.00	1.673	4	1	5
57	11	0	4.00	1.273	4	1	5
58	11	0	4.00	1.564	3	1	4
59	11	0	4.00	1.564	4	1	5
60	11	0	4.00	1.655	4	1	5
61	10	1	3.00	1.433	4	1	5
62	11	0	4.00	1.455	4	1	5
63	11	0	4.00	1.491	4	1	5
64	11	0	3.00	2.218	4	1	5
65	11	0	4.00	1.273	4	1	5
66	10	1	3.00	1.122	4	1	5
67	1	0	4.00		0	4	4
68	1	0	5.00		0	5	5
69	1	0	5.00		0	5	5
70	1	0	3.00		0	3	3
71	1	0	5.00		0	5	5
72	1	0	4.00		0	4	4
73	1	0	4.00		0	4	4
74	0	1					
75	0	1					
76	0	1					
77	1	0	4.00		0	4	4
78	1	0	4.00		0	4	4
79	1	0	5.00		0	5	5
80	1	0	4.00		0	4	4

Table 41: "Detailed relevance statistics on assessment criteria Round 1"

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
1	Valid	Relevant	3	27.3	27.3	27.3
		Highly relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
2	Valid	Not relevant	2	18.2	18.2	18.2

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
3	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	5	45.5	45.5	54.5
		Relevant	4	36.4	36.4	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
4	Valid	Not relevant	2	18.2	20.0	20.0
		Slightly relevant	2	18.2	20.0	40.0
		Moderately relevant	4	36.4	40.0	80.0
		Highly relevant	2	18.2	20.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
5	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	2	18.2	20.0	30.0
		Moderately relevant	3	27.3	30.0	60.0
		Relevant	3	27.3	30.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
6	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	3	27.3	27.3	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
7	Valid	Moderately relevant	4	36.4	36.4	36.4
		Relevant	3	27.3	27.3	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
8	Valid	Not relevant	1	9.1	9.1	9.1

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	2	18.2	18.2	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
9	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	2	18.2	18.2	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
10	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
11	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	2	18.2	18.2	54.5
		Relevant	2	18.2	18.2	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
12	Valid	Not relevant	1	9.1	9.1	9.1
		Relevant	3	27.3	27.3	36.4
		Highly relevant	7	63.6	63.6	100.0
		Total	11	100.0	100.0	
13	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	1	9.1	9.1	27.3
		Moderately relevant	4	36.4	36.4	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
14	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Total	11	100.0	100.0	
15	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	4	36.4	36.4	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
16	Valid	Moderately relevant	2	18.2	18.2	18.2
		Relevant	3	27.3	27.3	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
17	Valid	Slightly relevant	1	9.1	9.1	9.1
		Relevant	4	36.4	36.4	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
18	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
19	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	6	54.5	54.5	90.9
		Relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
20	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	4	36.4	36.4	63.6
		Relevant	2	18.2	18.2	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
21	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	4	36.4	36.4	54.5

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
22	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	1	9.1	10.0	20.0
		Moderately relevant	3	27.3	30.0	50.0
		Relevant	4	36.4	40.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
23	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
24	Valid	Moderately relevant	5	45.5	45.5	45.5
		Relevant	4	36.4	36.4	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
25	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	1	9.1	9.1	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	3	27.3	27.3	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
26	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
27	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Total	11	100.0	100.0	
28	Valid	Not relevant	1	9.1	9.1	9.1
		Relevant	4	36.4	36.4	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
29	Valid	Not relevant	2	18.2	18.2	18.2
		Relevant	3	27.3	27.3	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
30	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	1	9.1	9.1	18.2
		Relevant	7	63.6	63.6	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
31	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	3	27.3	27.3	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
32	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	4	36.4	36.4	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
33	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
34	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	4	36.4	36.4	45.5
		Relevant	4	36.4	36.4	81.8
		Highly relevant	2	18.2	18.2	100.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Total	11	100.0	100.0	
35	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	3	27.3	27.3	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
36	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	2	18.2	18.2	27.3
		Relevant	4	36.4	36.4	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
37	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	2	18.2	18.2	27.3
		Highly relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
38	Valid	Moderately relevant	1	9.1	11.1	11.1
		Relevant	7	63.6	77.8	88.9
		Highly relevant	1	9.1	11.1	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
39	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	1	9.1	9.1	18.2
		Relevant	5	45.5	45.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
40	Valid	Relevant	4	36.4	36.4	36.4
		Highly relevant	7	63.6	63.6	100.0
		Total	11	100.0	100.0	
41	Valid	Moderately relevant	3	27.3	27.3	27.3
		Relevant	5	45.5	45.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
42	Valid	Moderately relevant	2	18.2	18.2	18.2

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Relevant	5	45.5	45.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
43	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	2	18.2	18.2	36.4
		Moderately relevant	6	54.5	54.5	90.9
		Relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
44	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
45	Valid	Moderately relevant	6	54.5	54.5	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
46	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	4	36.4	36.4	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
47	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	2	18.2	18.2	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
48	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	1	9.1	9.1	27.3
		Relevant	2	18.2	18.2	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
49	Valid	Not relevant	2	18.2	18.2	18.2

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
50	Valid	Slightly relevant	1	9.1	10.0	10.0
		Moderately relevant	4	36.4	40.0	50.0
		Relevant	2	18.2	20.0	70.0
		Highly relevant	3	27.3	30.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
51	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	3	27.3	27.3	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
52	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	3	27.3	27.3	36.4
		Highly relevant	7	63.6	63.6	100.0
		Total	11	100.0	100.0	
53	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	2	18.2	20.0	30.0
		Moderately relevant	2	18.2	20.0	50.0
		Relevant	3	27.3	30.0	80.0
		Highly relevant	2	18.2	20.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
54	Valid	Moderately relevant	1	9.1	10.0	10.0
		Relevant	3	27.3	30.0	40.0
		Highly relevant	6	54.5	60.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
55	Valid	Not relevant	1	9.1	9.1	9.1

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	3	27.3	27.3	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
56	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	3	27.3	27.3	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
57	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
58	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	1	9.1	9.1	27.3
		Moderately relevant	1	9.1	9.1	36.4
		Relevant	7	63.6	63.6	100.0
		Total	11	100.0	100.0	
59	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	1	9.1	9.1	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
60	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	4	36.4	36.4	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
61	Valid	Not relevant	1	9.1	10.0	10.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Slightly relevant	2	18.2	20.0	30.0
		Moderately relevant	3	27.3	30.0	60.0
		Relevant	3	27.3	30.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
62	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	4	36.4	36.4	45.5
		Relevant	3	27.3	27.3	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
63	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	2	18.2	18.2	27.3
		Relevant	4	36.4	36.4	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
64	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	1	9.1	9.1	27.3
		Moderately relevant	3	27.3	27.3	54.5
		Relevant	2	18.2	18.2	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
65	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
66	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	3	27.3	30.0	40.0
		Moderately relevant	5	45.5	50.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
	Total		11	100.0		
67	Valid	Relevant	1	100.0	100.0	100.0
68	Valid	Highly relevant	1	100.0	100.0	100.0
69	Valid	Highly relevant	1	100.0	100.0	100.0
70	Valid	Moderately relevant	1	100.0	100.0	100.0
71	Valid	Highly relevant	1	100.0	100.0	100.0
72	Valid	Relevant	1	100.0	100.0	100.0
73	Valid	Relevant	1	100.0	100.0	100.0
74	Missing	System	1	100.0		
75	Missing	System	1	100.0		
76	Missing	System	1	100.0		
77	Valid	Relevant	1	100.0	100.0	100.0
78	Valid	Relevant	1	100.0	100.0	100.0
79	Valid	Highly relevant	1	100.0	100.0	100.0
80	Valid	Relevant	1	100.0	100.0	100.0

Appendix J Content analysis on assessment criteria Round 1

Content analysis techniques were used to discover common findings from the participant comments. These comments were scanned for generic remarks and underlined if one was found. An annotation was made for each underlined remark. These annotations were then scanned for similarities, indicated by colour coding. From the similar annotations, concepts were distilled.

Table 42: "Content analysis on assessment criteria Round 1"

ID	Quotes	Annotations	Concept
CP_1	Not every requirement is equaly necessary for all stakeholders. In RUP beschrijven ze ook het concept van supplementary specifications. Dit zijn requirements die niet direct zijn te herleiden naar de behoefte van de stakeholders.	Relevance and support of all stakeholders not required	Business domain
CP_2	In general not all the stakeholders agree with every requirement. It is morfe than likely that some requirements meet the need of one stakeholder and not all. The business case should point this out. It depends from which business perspective the assesor operates. all stakeholders not necessary. The problem might be that stakeholders have different interests/requirements	Different interests between stakeholders	Business domain
CP_3	A requirement on its own does not have costs. The project has been calculated as a whole, including an overall estimate of all requirements. So every requirement can not be assessed on its individual costs in perspective to the estimate of the whole. Niet alle requirements zijn goed te vertalen naar cost/benefit. It is not possible to make a clear cost/ benefit analysis at an indivudual requirement level.	Cost/benefit not always related to one requirement, it's more related to the **whole**	Business domain
CP_4	This is primarily the reason why we should make a business case in the first place. In the elicitation phase of the requirements process this aspect is very important Requirements should all be necessary. [..]contract is contract. Client should do his prioritsing	relevance should be part of the elicitation phase. Client should do prioritizing	Interaction with client Business domain
CP_6	Or you understand what is required or you do not. Een interpretatiefout (aanname) kan, als niet op tijd getackled zeer kostbaar zijn in tijd, geld en business value.	No assumptions	Interaction with client

ID	Quotes	Annotations	Concept
CP_7	As stated earlier the cost estimation is done for the project as a whole and not per requirement. Er zijn weinig individuele requirements die je direct kunt koppelen aan kosten en planning.	Related to the whole → business case	Business domain
CP_8	Depends a lot an how detailed the mission and system goals are formulated. If a function is required it has to be bulit as is and not as part of a mission goal. So, actually our client is responsible for describing all necessary functions in the requirements of the contract	Part of business domain. Contractual obligation	Business domain Contract
CP_9	excellent means for objective analysis of validity of requirements or the proof that the designer doesn't understand the business case	Needs interaction with client to understand the business case	Interaction with client
CP_{10}	What is necessary is subjective. In the initial phase it is very difficult to judge whether a requirement is really necessary or not (customer just says everything is necessary). What is necessary?	Need client to tell what is necessary	Interaction with client
CP_{11}	tolerances shoud be directly aligned to the business case	Business case	Business domain
CP_{12}	Necessity seems to focus on resources and Validity on regulatations/laws. However both aspects should be within the contractual scope.	contractual scope	Contract
CP_{13}	It depends on the type of requirement and the type of contract.	Contract	Contract
CP_{14}	Better have a bit more information than less (within reasonable amounts). It is in the interest of the business that a requirement does only contain necessary information. Can only be determined from the business perspective.	More information Business	Business domain
CP_{15}	Defining the required states is part of the engineering work. It is important that only the behavior and output of all functional states and contraints that are visible to the end users of a system (as a black box) are in the requirements Design freedom must stay in place.	Black box Design freedom	Black box
CP_{16}	Essential that the clien't's expectations matchthe contractor's expectations	matching expectations.	Interaction with client
CP_{17}	It is worse when the contradiction is wthin the requirement because in that case the requirement is	Discuss with client	Interaction with client

ID	Quotes	Annotations	Concept
	ambiguous. Contractor needs to discuss this immediately.		
CP_{18}	It is the professional who needs to be critical and who should acquire clarification.	Professional skills Acquire clarification	Professional skills
CP_{21}	the proof that the designer doesn't understand the business case	Understanding the business case	Interaction with client
CP_{22}	The leading question will always be : Is my interpretation of your requirement correct ?	Interpretation	Interaction with client
CP_{25}	Business experts are not qualified to define technical requirements and technological experts are not qualified to make the business case. This criteria seems to look at the quality of the tender process or the project strategy and not the requirement itself.	Design freedom Business domain Project level	Black box Business domain
CP_{28}	How do you assess this criterion whithout having interaction with the client?	Interaction with client	Interaction with client
CP_{29}	Requirements are part of contract!	Contract	Business domain
CP_{33}	Testing is not the goal of the requirement, Functionality needed is the goal.	What the business needs	Business domain
CP_{42}	Therefore the requirement as such should not be discarded before agreement is found between parties as to the (functional) need it adresses.	Agreement with client	Interaction with client
CP_{47}	This criteria seems to look at the quality of the tender process or the project strategy and not the requirement itself.	Project level	Project level
CP_{51}	Let the contractor solve this (original tekst: “Dit is aan de contractor om op te lossen.”)	Contractor design freedom	Black box
CP_{52}	A requirement that can't be verified should be discussed with customer as soon as possible.	Discuss with client	Interaction with client
CP_{53}	The real problem being lack of agreement between cpntractor and client.	Agreement with client	Interaction with client
CP_{58}	This criterion seems to look at the quality of the tender process or the project strategy and not the requirement itself.	Project level	Project level
CP_{59}	This criterion seems to look at the quality of the tender process or the project strategy and not the requirement itself.	Project level	Project level
CP_{60}	Describing too much in requirements will unwantedly restrain solutions (original text: “Te veel in requirements vastleggen beperkt ongewenst de oplossingsrichtingen”).	Need design freedom Project level	Black box Project level

ID	Quotes	Annotations	Concept
	This criterion seems to look at the quality of the tender process or the project strategy and not the requirement itself.		
CP_{61}	This criterion seems to look at the quality of the tender process or the project strategy and not the requirement itself.	Project level	Project level
CP_{64}	true but a question of prfessional capabilities to avoid the pitfall This criterion seems to look at the quality of the tender process or the project strategy and not the requirement itself.	Professional qualities Project level	Professional skills
CP_{65}	true but a question of prfessional capabilities to avoid the pitfall	Professional qualities	Professional skills
CP_{68}	Provide rationale to explain and justify why a requirement exists	Business	Business domain
CP_{79}	When requirements are open to more than one interpretation, difficulty always arise in the engineering process […]	Need interaction	Interaction with client
CP_{80}	If there is no business need, a equirement is simply not valid	Business need	Business domain

Appendix K Statistical analysis data on measurement methods Round 1

The measurement methods were processed in SPSS. To create the frequency tables, the file was split by Measurement ID:

```
SORT CASES  BY M_ID
SPLIT FILE LAYERED BY M_ID
```

Only data from Round 1 was entered when analyzing, so no filter on cases was needed. Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=M_Relevance
  /STATISTICS=RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 43: "Frequencies on measurement methods Round 1"

Measurement ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
1	11	0	3.00	.600	2	2	4
2	11	0	3.00	1.164	3	1	4
3	11	0	3.00	.873	3	2	5
4	11	0	3.00	1.091	4	1	5
5	11	0	4.00	1.564	4	1	5
6	11	0	2.00	1.473	4	1	5
7	1	0	4.00		0	4	4

Table 44: "Detailed relevance statistics on measurement methods Round 1"

Measurement ID			Frequency	Percent	Valid Percent	Cumulative Percent
1	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	5	45.5	45.5	72.7
		Relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
2	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	1	9.1	9.1	27.3
		Moderately relevant	5	45.5	45.5	72.7
		Relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
3	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	6	54.5	54.5	63.6
		Relevant	2	18.2	18.2	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
4	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	6	54.5	54.5	72.7
		Relevant	2	18.2	18.2	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
5	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	1	9.1	9.1	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	

Measurement ID			Frequency	Percent	Valid Percent	Cumulative Percent
6	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	4	36.4	36.4	54.5
		Moderately relevant	3	27.3	27.3	81.8
		Relevant	1	9.1	9.1	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
7	Valid	Relevant	1	100.0	100.0	100.0

Appendix L 1 Content analysis on measurement methods Round

Content analysis techniques were used to discover common findings from the participant comments. These comments were scanned for generic remarks and underlined if one was found. An annotation was made for each underlined remark. These annotations were then scanned for similarities, indicated by colour coding. From the similar annotations, concepts were distilled.

Table 45: "Content analysis on measurement methods Round 1"

ID	Quotes	Annotations	Concept
M_1	As many business needs are hard to quantify but still relevant, the measurements need to be in the same terms. Too many values are not useful. Concept of uncertainty not recognized.	Hard Too many values Concept	Complexity
M_2	Only with very experienced reviewers this can be useful. Otherwise the results wil vary a lot and are probably less useful. KISS! So one dimensional only	Hard Not simple	Complexity
M_3	Two is not enough to make a good distinction Simplicity is often good. […]Therefore, less is more and change management is crucial. The problem I have with this is that there are 2 values but intermediate values are possible (Original text: “Mijn probleem is dat we hier 2 waarden hebben, en tussensmaken zijn denkbaar.”) Under the assumption that in a good requirements specification the large majority of requirements is quite OK, just flagging the problematic ones could be a suitable method.	Not elaborate enough Simple Not enough values Not simple enough	Complexity
M_4	I think this method can be useful, however also relatively complex Simpel	Complex Simple	Complexity
M_5	Too many options (Original text: “Te veel opties”) Risk based approach that appeals to me most.	Too many options Appealing	Complexity
M_6	This measurment seems overly complex on first glance. Good way to review requirements because it takes into account more factors (such as experience of the company), with easy to understand result. This method however is also quite time consuming […] It looks like an excessively heavy method to map the uncertainties of a requirement (Original text: “Het lijkt me een overdreven zware methode om de uncertainty van requirements in kaart te brengen”)	Complex Appealing Time consuming Heavy	Complexity

Appendix M Statistical analysis data on influencing factors Round 1

The influencing factors were processed in SPSS. To create the frequency tables, the file was split by Factor ID:

```
SORT CASES  BY F_ID
SPLIT FILE LAYERED BY F_ID
```

Only data from Round 1 was entered when analyzing, so no filter on cases was needed. Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=F_Relevance
  /STATISTICS=RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 46: "Frequencies on influencing factors Round 1"

Factor ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
1	11	0	5.00	.818	2	3	5
2	11	0	4.00	.891	3	2	5
3	1	0	3.00		0	3	3
4	1	0	4.00		0	4	4
5	1	0	4.00		0	4	4
6	1	0	3.00		0	3	3
7	1	0	5.00		0	5	5

Table 47: "Detailed relevance statistics on influencing factors Round 1"

Factor ID			Frequency	Percent	Valid Percent	Cumulative Percent
1	Valid	Moderately relevant	3	27.3	27.3	27.3
		Relevant	2	18.2	18.2	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
2	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	2	18.2	18.2	27.3
		Relevant	5	45.5	45.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
3	Valid	Moderately relevant	1	100.0	100.0	100.0
4	Valid	Relevant	1	100.0	100.0	100.0
5	Valid	Relevant	1	100.0	100.0	100.0
6	Valid	Moderately relevant	1	100.0	100.0	100.0
7	Valid	Highly relevant	1	100.0	100.0	100.0

Appendix N Content analysis on influencing factors Round 1

Content analysis techniques were used to discover common findings from the participant comments. These comments were scanned for generic remarks and underlined if one was found. An annotation was made for each underlined remark. These annotations were then scanned for similarities, indicated by colour coding. From the similar annotations, concepts were distilled.

Table 48: "Content analysis on influencing factors Round 1"

ID	Quotes	Annotations	Concept
F1	Cost/benefit is part of the risk assessment as risks will be measured in cost/benefits and are therefore less its own factor	Cost/benefit – Risk relation	Cost/Benefit-Risk
F2	Very often risk aversion is solely focussed on cost and very often leads to suboptimal benefits.	Risk Cost Benefit	Cost/Benefit-Risk
F3	If it is known how to deal with a requirement from past projects, more uncertainty is probably allowed.	Experience from past projects	Experience
F6	How familiar the project organization, colleagues, etc are with the rules and context of the environment in which the implementation should operate defines how certain (how elaborate) the requirements need to be	Familiarity/experience	Experience

Appendix O **Statistical analysis data on assessment criteria Round 2**

The measurement methods were processed in SPSS. To create the frequency tables, the file was split by Measurement ID:

```
SORT CASES  BY CP_ID.
SPLIT FILE LAYERED BY CP_ID.
```

Only cases for Round 2 were selected:

```
USE ALL.
COMPUTE filter_$=(Round = 2).
VARIABLE LABELS filter_$ 'Round = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=CP_Relevance
  /STATISTICS=VARIANCE RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 49: "Frequencies on measurement methods Round 2"

Criterion ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
1	10	1	5.00	.100	1	4	5
2	11	0	3.00	.691	3	1	4
3	11	0	3.00	.364	2	2	4
4	11	0	3.00	1.255	4	1	5
5	11	0	3.00	1.055	3	1	4
6	11	0	3.00	1.455	4	1	5
7	11	0	4.00	.564	2	3	5
8	11	0	4.00	1.400	4	1	5
9	11	0	4.00	.564	2	3	5
10	11	0	4.00	1.873	4	1	5
11	11	0	3.00	1.418	4	1	5
12	11	0	5.00	.473	2	3	5
13	11	0	3.00	1.455	4	1	5
14	11	0	3.00	.855	3	2	5
15	11	0	4.00	.818	3	2	5
16	11	0	5.00	.655	2	3	5
17	11	0	4.00	.273	1	4	5
18	11	0	4.00	.273	1	3	4
19	11	0	3.00	.655	2	1	3
20	11	0	3.00	1.964	4	1	5
21	11	0	4.00	.473	2	2	4
22	11	0	3.00	.673	3	2	5
23	11	0	4.00	.618	2	3	5
24	11	0	4.00	.764	3	2	5
25	11	0	3.00	1.255	4	1	5
26	11	0	4.00	1.073	4	1	5
27	11	0	4.00	.818	3	2	5
28	11	0	4.00	1.364	4	1	5
29	11	0	4.00	2.164	4	1	5
30	11	0	4.00	.200	2	3	5

31	11	0	3.00	.818	3	2	5
32	10	1	3.00	1.067	4	1	5
33	11	0	3.00	.291	2	2	4
34	11	0	4.00	.455	2	3	5
35	11	0	3.00	.691	2	2	4
36	10	1	4.00	.544	2	3	5
37	11	0	5.00	.091	1	4	5
38	11	0	4.00	.091	1	4	5
39	11	0	4.00	.491	2	3	5
40	11	0	5.00	.164	1	4	5
41	11	0	4.00	.818	3	2	5
42	11	0	4.00	.455	2	3	5
43	11	0	2.00	.691	2	1	3
44	11	0	4.00	.600	2	3	5
45	11	0	3.00	.655	2	3	5
46	11	0	3.00	.655	3	2	5
47	11	0	5.00	.818	2	3	5
48	11	0	4.00	2.255	4	1	5
49	11	0	3.00	1.255	3	1	4
50	11	0	4.00	.491	2	3	5
51	11	0	3.00	1.455	4	1	5
52	11	0	5.00	.455	2	3	5
53	11	0	3.00	1.073	3	2	5
54	11	0	5.00	.455	2	3	5
55	11	0	3.00	.818	3	1	4
56	11	0	4.00	1.000	3	2	5
57	11	0	4.00	1.055	3	1	4
58	11	0	4.00	1.273	4	1	5
59	11	0	3.00	1.091	3	1	4
60	11	0	4.00	1.018	3	1	4
61	10	1	3.50	1.344	4	1	5
62	11	0	4.00	.564	2	3	5
63	11	0	4.00	1.491	4	1	5
64	11	0	3.00	1.055	3	2	5
65	11	0	4.00	.673	3	2	5
66	10	1	2.00	1.156	4	1	5
67	10	1	4.00	1.378	4	1	5
68	11	0	4.00	.818	3	2	5
69	10	1	4.50	2.933	4	1	5
70	11	0	2.00	.618	2	1	3
71	11	0	4.00	2.618	4	1	5
72	10	1	4.00	2.000	3	1	4
73	9	2	3.00	1.750	3	1	4
74	9	2	4.00	2.278	4	1	5
75	9	2	2.00	1.000	3	1	4
76	9	2	2.00	1.250	3	1	4
77	9	2	4.00	2.000	4	1	5
78	9	2	4.00	2.000	4	1	5
79	8	3	4.00	2.554	4	1	5
80	9	2	4.00	2.444	4	1	5

Table 50: "Detailed relevance statistics on assessment criteria Round 2"

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
1	Valid	Relevant	1	9.1	10.0	10.0
		Highly relevant	9	81.8	90.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
2	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	7	63.6	63.6	72.7
		Relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
3	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	7	63.6	63.6	72.7
		Relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
4	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	2	18.2	18.2	36.4
		Moderately relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
5	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	2	18.2	18.2	36.4
		Moderately relevant	5	45.5	45.5	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
6	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
7	Valid	Moderately relevant	4	36.4	36.4	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
8	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	1	9.1	9.1	18.2
		Relevant	5	45.5	45.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
9	Valid	Moderately relevant	4	36.4	36.4	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
10	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	1	9.1	9.1	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
11	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	5	45.5	45.5	63.6
		Relevant	2	18.2	18.2	81.8
		Highly relevant	2	18.2	18.2	100.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Total	11	100.0	100.0	
12	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	3	27.3	27.3	36.4
		Highly relevant	7	63.6	63.6	100.0
		Total	11	100.0	100.0	
13	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
14	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	4	36.4	36.4	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
15	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
16	Valid	Moderately relevant	2	18.2	18.2	18.2
		Relevant	3	27.3	27.3	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
17	Valid	Relevant	6	54.5	54.5	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
18	Valid	Moderately relevant	5	45.5	45.5	45.5
		Relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
19	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	3	27.3	27.3	45.5
		Moderately relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
20	Valid	Not relevant	3	27.3	27.3	27.3
		Slightly relevant	1	9.1	9.1	36.4
		Moderately relevant	3	27.3	27.3	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
21	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	4	36.4	36.4	45.5
		Relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
22	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	5	45.5	45.5	54.5
		Relevant	4	36.4	36.4	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
23	Valid	Moderately relevant	5	45.5	45.5	45.5
		Relevant	4	36.4	36.4	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
24	Valid	Slightly relevant	1	9.1	9.1	9.1

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Moderately relevant	2	18.2	18.2	27.3
		Relevant	6	54.5	54.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
25	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	6	54.5	54.5	63.6
		Relevant	2	18.2	18.2	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
26	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
27	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
28	Valid	Not relevant	1	9.1	9.1	9.1
		Relevant	5	45.5	45.5	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
29	Valid	Not relevant	2	18.2	18.2	18.2
		Relevant	5	45.5	45.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
30	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	9	81.8	81.8	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
31	Valid	Slightly relevant	5	45.5	45.5	45.5
		Moderately relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
32	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	2	18.2	20.0	30.0
		Moderately relevant	6	54.5	60.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
33	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	8	72.7	72.7	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
34	Valid	Moderately relevant	5	45.5	45.5	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
35	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	4	36.4	36.4	63.6
		Relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
36	Valid	Moderately relevant	2	18.2	20.0	20.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Relevant	5	45.5	50.0	70.0
		Highly relevant	3	27.3	30.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
37	Valid	Relevant	1	9.1	9.1	9.1
		Highly relevant	10	90.9	90.9	100.0
		Total	11	100.0	100.0	
38	Valid	Relevant	10	90.9	90.9	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
39	Valid	Moderately relevant	2	18.2	18.2	18.2
		Relevant	6	54.5	54.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
40	Valid	Relevant	2	18.2	18.2	18.2
		Highly relevant	9	81.8	81.8	100.0
		Total	11	100.0	100.0	
41	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
42	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	5	45.5	45.5	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
43	Valid	Not relevant	3	27.3	27.3	27.3
		Slightly relevant	4	36.4	36.4	63.6
		Moderately relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
44	Valid	Moderately relevant	3	27.3	27.3	27.3
		Relevant	5	45.5	45.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
45	Valid	Moderately relevant	6	54.5	54.5	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
46	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	6	54.5	54.5	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
47	Valid	Moderately relevant	3	27.3	27.3	27.3
		Relevant	2	18.2	18.2	45.5
		Highly relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
48	Valid	Not relevant	2	18.2	18.2	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	3	27.3	27.3	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
49	Valid	Not relevant	3	27.3	27.3	27.3
		Moderately relevant	6	54.5	54.5	81.8

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
50	Valid	Moderately relevant	3	27.3	27.3	27.3
		Relevant	6	54.5	54.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
51	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	3	27.3	27.3	45.5
		Moderately relevant	4	36.4	36.4	81.8
		Relevant	1	9.1	9.1	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
52	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	2	18.2	18.2	27.3
		Highly relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
53	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	4	36.4	36.4	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
54	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	2	18.2	18.2	27.3
		Highly relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
55	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	5	45.5	45.5	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
56	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	2	18.2	18.2	27.3
		Relevant	4	36.4	36.4	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
57	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	7	63.6	63.6	100.0
		Total	11	100.0	100.0	
58	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
59	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	3	27.3	27.3	54.5
		Relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
60	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	1	9.1	9.1	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	6	54.5	54.5	100.0

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Total	11	100.0	100.0	
61	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	1	9.1	10.0	20.0
		Moderately relevant	3	27.3	30.0	50.0
		Relevant	4	36.4	40.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
62	Valid	Moderately relevant	4	36.4	36.4	36.4
		Relevant	5	45.5	45.5	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
63	Valid	Not relevant	1	9.1	9.1	9.1
		Moderately relevant	2	18.2	18.2	27.3
		Relevant	4	36.4	36.4	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
64	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	5	45.5	45.5	63.6
		Relevant	2	18.2	18.2	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
65	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	4	36.4	36.4	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
66	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	6	54.5	60.0	70.0
		Moderately relevant	2	18.2	20.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
67	Valid	Not relevant	1	9.1	10.0	10.0
		Slightly relevant	1	9.1	10.0	20.0
		Moderately relevant	2	18.2	20.0	40.0
		Relevant	5	45.5	50.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
68	Valid	Slightly relevant	1	9.1	9.1	9.1
		Relevant	5	45.5	45.5	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
69	Valid	Not relevant	2	18.2	20.0	20.0
		Slightly relevant	1	9.1	10.0	30.0
		Moderately relevant	1	9.1	10.0	40.0
		Relevant	1	9.1	10.0	50.0
		Highly relevant	5	45.5	50.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
70	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	4	36.4	36.4	54.5
		Moderately relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
71	Valid	Not relevant	3	27.3	27.3	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	3	27.3	27.3	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
72	Valid	Not relevant	3	27.3	30.0	30.0
		Moderately relevant	1	9.1	10.0	40.0
		Relevant	6	54.5	60.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
73	Valid	Not relevant	3	27.3	33.3	33.3
		Moderately relevant	3	27.3	33.3	66.7
		Relevant	3	27.3	33.3	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
74	Valid	Not relevant	2	18.2	22.2	22.2
		Relevant	5	45.5	55.6	77.8
		Highly relevant	2	18.2	22.2	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
75	Valid	Not relevant	2	18.2	22.2	22.2
		Slightly relevant	3	27.3	33.3	55.6
		Moderately relevant	3	27.3	33.3	88.9
		Relevant	1	9.1	11.1	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
76	Valid	Not relevant	2	18.2	22.2	22.2
		Slightly relevant	4	36.4	44.4	66.7
		Moderately relevant	1	9.1	11.1	77.8
		Relevant	2	18.2	22.2	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
77	Valid	Not relevant	2	18.2	22.2	22.2
		Moderately relevant	1	9.1	11.1	33.3
		Relevant	5	45.5	55.6	88.9
		Highly relevant	1	9.1	11.1	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
78	Valid	Not relevant	2	18.2	22.2	22.2
		Moderately relevant	1	9.1	11.1	33.3
		Relevant	5	45.5	55.6	88.9
		Highly relevant	1	9.1	11.1	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
79	Valid	Not relevant	2	18.2	25.0	25.0
		Moderately relevant	1	9.1	12.5	37.5
		Relevant	3	27.3	37.5	75.0
		Highly relevant	2	18.2	25.0	100.0
		Total	8	72.7	100.0	
	Missing	System	3	27.3		
	Total		11	100.0		
80	Valid	Not relevant	2	18.2	22.2	22.2
		Slightly relevant	1	9.1	11.1	33.3
		Moderately relevant	1	9.1	11.1	44.4
		Relevant	3	27.3	33.3	77.8
		Highly relevant	2	18.2	22.2	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		

Duplicate results were identified via content analysis and will be removed from the list, see Table 51 for the criteria concerned.

Table 51: "Duplicate assessment criteria identified in Round 2"

ID	Criterion	Duplicate of	Relevance	Perspective	Meta aspect
CP_{67}	It is clear who the stakeholder of the requirement is	CP_{13}	4.0	Business	Stakeholder support
CP_{73}	failure behaviour have been specified	CP_{58}, CP_{59}, CP_{60}	3.0	Risk	Reliability
CP_{74}	SMART	All	4.0	quality	Completeness
CP_{76}	Originator requirement	CP_{21}, CP_{67}	2.0	quality	#N/A
CP_{77}	Does the requirement add value	CP_{1}, CP_{3}, CP_{4}, CP_{8}, CP_{10}	4.0	Business	Validity
CP_{78}	Is the requirement unique	CP_{46}, CP_{48}	4.0	Quality	Consistency
CP_{79}	Is the requirement unambiguous	CP_{16}, CP_{26}, CP_{27}, CP_{30}, CP_{39}	4.0	Quality	Risk
CP_{80}	Does the requirement meet a business need	CP_{1}, CP_{3}, CP_{4}, CP_{8}, CP_{10}	4.0	Business	Validity

Appendix P Content analysis on assessment criteria Round 2

Content analysis techniques were used to discover common findings from the participant comments. These comments were scanned for generic remarks and underlined if one was found. An annotation was made for each underlined remark. These annotations were then scanned for similarities, indicated by colour coding. From the similar annotations, concepts were distilled.

Table 52: "Content analysis on measurement methods Round 2"

ID	Quotes	Annotations	Concept
CP_6	By investing in making these assumptions explicit, the chance of non acceptance is reduced considerably	Make assumptions explicit	Assumptions
CP_{25}	However, the business case is definitive. If a customer wants a cow, he shouldn't define his need for a 4 legged animal that eats grass and produces milk. So , it's relative to the business case. like the comments of round 1. Criteria is true but not practiable workable in a project on a indivual requirement level	Business Set of requirements	Set Business domain
CP_{29}	Moving assumptions into validated requirements or remove them altogether should be your aim during the requirements review. no assumptions allowed in a requirement Requirements should never be based on assumptions but on facts or calculations.	No assumptions	Assumptions
CP_{66}	Most of the quality criteria for requirements also apply to the entire set of requirements.	Set of requirements	Set
CP_{71}	This questionair is about the quality of individual requirements, not about the completeness of a requirement set. not a requirement criteria but a project criteria	Project level	Set
CP_{72}	This questionair is about the quality of individual requirements, not about the completeness of a requirement set. not a requirement criteria but a project criteria	Project level	Set

Appendix Q Statistical analysis data on measurement methods Round 2

The measurement methods were processed in SPSS. To create the frequency tables, the file was split by Measurement ID:

```
SORT CASES  BY M_ID.
SPLIT FILE LAYERED BY M_ID.
```

Only cases for Round 2 were selected:

```
USE ALL.
COMPUTE filter_$=(Round = 2).
VARIABLE LABELS filter_$ 'Round = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=M_Relevance
  /STATISTICS=VARIANCE RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 53: "Frequencies on measurement methods Round 2"

Measurement ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
1	9	2	3.00	.528	2	1	3
2	11	0	3.00	.655	3	1	4
3	11	0	3.00	.673	3	2	5
4	11	0	3.00	.291	2	2	4
5	11	0	3.00	1.200	3	1	4
6	11	0	2.00	1.018	3	1	4
7	11	0	4.00	.655	2	2	4

Table 54: "Detailed relevance statistics on measurement methods Round 2"

Measurement ID			Frequency	Percent	Valid Percent	Cumulative Percent
1	Valid	Not relevant	1	9.1	11.1	11.1
		Slightly relevant	2	18.2	22.2	33.3
		Moderately relevant	6	54.5	66.7	100.0
		Total	9	81.8	100.0	
	Missing	System	2	18.2		
	Total		11	100.0		
2	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	6	54.5	54.5	90.9
		Relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
3	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	5	45.5	45.5	54.5
		Relevant	4	36.4	36.4	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
4	Valid	Slightly relevant	1	9.1	9.1	9.1

Measurement ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Moderately relevant	8	72.7	72.7	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
5	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	2	18.2	18.2	54.5
		Relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
6	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	6	54.5	54.5	72.7
		Moderately relevant	1	9.1	9.1	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
7	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	3	27.3	27.3	45.5
		Relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	

Appendix R Statistical analysis data on influencing factors Round 2

The influencing factors were processed in SPSS. To create the frequency tables, the file was split by Factor ID:

```
SORT CASES  BY F_ID
SPLIT FILE LAYERED BY F_ID
```

Only cases for Round 2 were selected:

```
USE ALL.
COMPUTE filter_$=(Round = 2).
VARIABLE LABELS filter_$ 'Round = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=F_Relevance
  /STATISTICS=RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 55: "Frequencies on influencing factors Round 2"

Factor ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
1	11	0	4.00	.455	2	3	5
2	11	0	4.00	.964	3	2	5
3	11	0	3.00	.655	2	3	5
4	11	0	4.00	.873	3	2	5
5	10	1	3.50	.489	2	3	5
6	11	0	3.00	.491	2	2	4
7	11	0	4.00	2.018	4	1	5

Table 56: "Detailed relevance statistics on influencing factors Round 2"

Factor ID			Frequency	Percent	Valid Percent	Cumulative Percent
1	Valid	Moderately relevant	1	9.1	9.1	9.1
		Relevant	5	45.5	45.5	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
2	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	3	27.3	27.3	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
3	Valid	Moderately relevant	6	54.5	54.5	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
4	Valid	Slightly relevant	2	18.2	18.2	18.2

Factor ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
5	Valid	Moderately relevant	5	45.5	50.0	50.0
		Relevant	4	36.4	40.0	90.0
		Highly relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		
6	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	6	54.5	54.5	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
7	Valid	Not relevant	2	18.2	18.2	18.2
		Relevant	6	54.5	54.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	

Appendix S **Statistical analysis data on assessment criteria Round 3**

The measurement methods were processed in SPSS. To create the frequency tables, the file was split by Measurement ID:

```
SORT CASES  BY CP_ID.
SPLIT FILE LAYERED BY CP_ID.
```

Only cases for Round 3 were selected:

```
USE ALL.
COMPUTE filter_$=(Round = 3).
VARIABLE LABELS filter_$ 'Round = 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=CP_Relevance
  /STATISTICS=VARIANCE RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 57: "Frequencies on measurement methods Round 3"

Criterion ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
4	11	0	3.00	.764	3	1	4
5	11	0	2.00	.655	2	2	4
6	11	0	4.00	1.218	3	2	5
8	11	0	4.00	.764	3	2	5
10	11	0	4.00	2.564	4	1	5
11	11	0	3.00	.964	3	2	5
13	11	0	3.00	1.764	4	1	5
20	11	0	3.00	1.091	3	1	4
25	11	0	3.00	1.055	3	2	5
26	11	0	4.00	.564	3	2	5
28	11	0	4.00	.818	3	2	5
29	11	0	4.00	1.291	4	1	5
32	11	0	3.00	.691	3	2	5
48	11	0	4.00	1.455	3	2	5
49	11	0	3.00	1.073	3	1	4
51	11	0	3.00	1.200	4	1	5
53	11	0	4.00	.418	2	3	5
56	11	0	4.00	.564	2	3	5
57	11	0	4.00	.673	2	2	4
58	11	0	4.00	.873	2	2	4
59	11	0	3.00	.764	3	1	4
60	11	0	4.00	1.164	3	1	4
61	11	0	4.00	1.164	3	1	4
63	11	0	4.00	.800	3	2	5
64	11	0	3.00	.655	3	2	5
66	11	0	2.00	1.218	4	1	5
69	11	0	4.00	1.218	3	2	5
71	11	0	4.00	1.891	4	1	5
72	11	0	4.00	1.964	3	1	4

Criterion ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
75	10	1	2.00	.989	3	1	4

Table 58: "Detailed relevance statistics on assessment criteria Round 3"

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
4	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	6	54.5	54.5	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
5	Valid	Slightly relevant	6	54.5	54.5	54.5
		Moderately relevant	3	27.3	27.3	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
6	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
8	Valid	Slightly relevant	1	9.1	9.1	9.1
		Relevant	6	54.5	54.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
10	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	3	27.3	27.3	45.5
		Relevant	3	27.3	27.3	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
11	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	4	36.4	36.4	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
13	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	3	27.3	27.3	36.4
		Moderately relevant	2	18.2	18.2	54.5
		Relevant	3	27.3	27.3	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
20	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	3	27.3	27.3	54.5
		Relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
25	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	5	45.5	45.5	63.6
		Relevant	2	18.2	18.2	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
26	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	1	9.1	9.1	18.2
		Relevant	8	72.7	72.7	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
28	Valid	Slightly relevant	1	9.1	9.1	9.1
		Relevant	5	45.5	45.5	54.5
		Highly relevant	5	45.5	45.5	100.0
		Total	11	100.0	100.0	
29	Valid	Not relevant	1	9.1	9.1	9.1
		Relevant	6	54.5	54.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
32	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	7	63.6	63.6	81.8
		Relevant	1	9.1	9.1	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
48	Valid	Slightly relevant	3	27.3	27.3	27.3
		Moderately relevant	1	9.1	9.1	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
49	Valid	Not relevant	3	27.3	27.3	27.3
		Slightly relevant	1	9.1	9.1	36.4
		Moderately relevant	6	54.5	54.5	90.9
		Relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
51	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	5	45.5	45.5	72.7
		Relevant	2	18.2	18.2	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
53	Valid	Moderately relevant	4	36.4	36.4	36.4
		Relevant	6	54.5	54.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
56	Valid	Moderately relevant	2	18.2	18.2	18.2
		Relevant	5	45.5	45.5	63.6
		Highly relevant	4	36.4	36.4	100.0
		Total	11	100.0	100.0	
57	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	1	9.1	9.1	27.3
		Relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
58	Valid	Slightly relevant	3	27.3	27.3	27.3
		Relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
59	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	6	54.5	54.5	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
60	Valid	Not relevant	1	9.1	9.1	9.1
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
61	Valid	Not relevant	1	9.1	9.1	9.1

Criterion ID			Frequency	Percent	Valid Percent	Cumulative Percent
		Slightly relevant	2	18.2	18.2	27.3
		Moderately relevant	2	18.2	18.2	45.5
		Relevant	6	54.5	54.5	100.0
		Total	11	100.0	100.0	
63	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	1	9.1	9.1	18.2
		Relevant	6	54.5	54.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
64	Valid	Slightly relevant	1	9.1	9.1	9.1
		Moderately relevant	6	54.5	54.5	63.6
		Relevant	3	27.3	27.3	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
66	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	6	54.5	54.5	72.7
		Moderately relevant	2	18.2	18.2	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
69	Valid	Slightly relevant	2	18.2	18.2	18.2
		Moderately relevant	2	18.2	18.2	36.4
		Relevant	4	36.4	36.4	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
71	Valid	Not relevant	2	18.2	18.2	18.2
		Slightly relevant	2	18.2	18.2	36.4
		Moderately relevant	1	9.1	9.1	45.5
		Relevant	5	45.5	45.5	90.9
		Highly relevant	1	9.1	9.1	100.0
		Total	11	100.0	100.0	
72	Valid	Not relevant	3	27.3	27.3	27.3
		Relevant	8	72.7	72.7	100.0
		Total	11	100.0	100.0	
75	Valid	Not relevant	3	27.3	30.0	30.0
		Slightly relevant	4	36.4	40.0	70.0
		Moderately relevant	2	18.2	20.0	90.0
		Relevant	1	9.1	10.0	100.0
		Total	10	90.9	100.0	
	Missing	System	1	9.1		
	Total		11	100.0		

Appendix T Statistical analysis data on measurement methods Round 3

The measurement methods were processed in SPSS. To create the frequency tables, the file was split by Measurement ID:

```
SORT CASES  BY M_ID.
SPLIT FILE LAYERED BY M_ID.
```

Only cases for Round 3 were selected:

```
USE ALL.
COMPUTE filter_$=(Round = 3).
VARIABLE LABELS filter_$ 'Round = 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=M_Relevance
  /STATISTICS=VARIANCE RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 59: "Frequencies on measurement methods Round 3"

Measurement ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
5	11	0	3.00	.618	2	2	4
6	11	0	2.00	.164	1	2	3

Table 60: "Detailed relevance statistics on measurement methods Round 3"

Measurement ID			Frequency	Percent	Valid Percent	Cumulative Percent
5	Valid	Slightly relevant	5	45.5	45.5	45.5
		Moderately relevant	4	36.4	36.4	81.8
		Relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	
6	Valid	Slightly relevant	9	81.8	81.8	81.8
		Moderately relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	

Appendix U Statistical analysis data on influencing factors Round 3

The influencing factors were processed in SPSS. To create the frequency tables, the file was split by Factor ID:

```
SORT CASES  BY F_ID
SPLIT FILE LAYERED BY F_ID
```

Only cases for Round 3 were selected:

```
USE ALL.
COMPUTE filter_$=(Round = 3).
VARIABLE LABELS filter_$ 'Round = 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

Frequency tables were generated using the following command:

```
FREQUENCIES VARIABLES=F_Relevance
  /STATISTICS=RANGE MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

Table 61: "Frequencies on influencing factors Round 3"

Factor ID	N		Median	Variance	Range	Min	Max
	Valid	Missing					
2	11	0	4.00	.491	2	3	5
7	11	0	4.00	1.091	4	1	5

Table 62: "Frequencies on influencing factors Round 3"

Factor ID			Frequency	Percent	Valid Percent	Cumulative Percent
2	Valid	Moderately relevant	2	18.2	18.2	18.2
		Relevant	6	54.5	54.5	72.7
		Highly relevant	3	27.3	27.3	100.0
		Total	11	100.0	100.0	
7	Valid	Not relevant	1	9.1	9.1	9.1
		Relevant	8	72.7	72.7	81.8
		Highly relevant	2	18.2	18.2	100.0
		Total	11	100.0	100.0	