K. van Limbergen, July 2020

Assessing collaborative behaviour as partner selection mechanism

An explorative study into team assessments used in the tender process of infrastructure projects

TU Delft

AT OSBORNE

Assessing collaborative behaviour as partner selection mechanism

An explorative study into team assessments used in the tender process of infrastructure projects

Master Thesis
July 2020

In partial fulfilment of the requirements for the degree of
Master of Science in Construction, Management and Engineering,
at Delft University of Technology, the Netherlands

to be defended on July 16, 2020

by

Koen VAN LIMBERGEN

Cover page (Suurhoffbrug)	© Peter de Krom

Permission from the photographer to use the photo for this report

Cover Part I (A2 Hooggelegen)	© Rijkswaterstaat
Cover Part II (Markermeerdijken)	© Arcadis
Cover Part III (Rantaväylä tunnel)	© Väylä
Cover Part IV (Nijkerkerbrug)	© ProjectDOEN
Cover Part V (TTT)	© Thames Water
Cover Part VI (Suurfhoffbrug)	© ProRail

Author

K. (Koen) van Limbergen
K.vanlimbergen@tudelft.nl
Koen.vanlimbergen@atosborne.nl
Student number: 4276124

Graduation committee

Chairman
Prof. Dr.ir. M.J.C.M. (Marcel) Hertogh Faculty of Civil Engineering and Geosciences, TU Delft

1st supervisor
Dr. ir. M.G.C. (Marian) Bosch-Rekveldt Faculty of Civil Engineering and Geosciences, TU Delft

2nd supervisor
Dr.ir. B.M. (Bauke) Steenhuisen Faculty of Technology, Policy and Management, TU Delft

1st company supervisor
Ir. T.T.E.R. (Thomas) Neijenhuis Infrastructure, Planning & Environment, AT Osborne

2nd company supervisor
Ir. M.T.A. (Mark) Roelofsen Infrastructure, Planning & Environment, AT Osborne

MSc Construction Management and Engineering

Delft University of Technology
/ Faculty of Civil Engineering and Geoscience
/ Faculty of Technology, Policy and Management
/ Faculty of Architecture & the Built Environment

Research is conducted at:

Delft University of Technology
Mekelweg 5, 2628 CD Delft

Graduation company:

AT Osborne
John F. Kennedylaan 100, 3741 EH Baarn

Preface

This report is the result of my graduation research project and was carried out for the master program Construction Management & Engineering at TU Delft. Finishing this study is a special moment. After years of following courses at three of our different faculties; the Faculty of Technology, Policy and Management, the Faculty of Architecture & the Built Environment and the Faculty of Civil Engineering and Geoscience, it is now time to say goodbye.

Over the last seven months, I had the chance to apply my knowledge, be challenged and to gain more insights in my fields of interest. It was a real privilege to interview sixteen senior professionals in this domain, both in the Netherlands and abroad. Since the potential of team assessments has gained increased attention in the Dutch construction sector, there could not have been a better time to conduct this study. I already had the opportunity to share my lessons and put them into practice at Dutch projects that are designing their tender process.

My previous internship at TwynstraGudde in Amersfoort can be seen as the starting point for becoming acquitted with team assessments for partner selection purposes. It was my former colleague Elske who introduced me to this topic and showed me the interface of assessment and procurement practices. Selecting a partner on how you will work together in the future is perhaps the most obvious and oldest selection criterion there is. Yet recent years have brought new attention to this mechanism in large infrastructural projects across Europe.

I would like to express my gratitude to everyone who made this research possible. Of course, my committee; Marcel, Marian, Bauke, Thomas and Mark for providing me the tools and feedback to perform this research. I would also like to thank all colleagues and interviewees for providing me the opportunity to conduct this research and having a great time doing so. A special thanks goes to Elske, Marieke and all other people who gave me direction to the practical relevancy of this topic, of which I am confident, added to the value of this research.

On a personal note, I want to thank my family, friends, my 'graduating buddies' Alexander, Sybren and Niek, and my boyfriend Luuk. The Covid-19-situation has made this a quite unique ending of all those sociable and eventful years in Delft but your support was very helpful.

Wishing you a pleasant reading,

Koen van Limbergen

July, 2020

Summary

Investing in the relationship between the client and contractor cannot guarantee a successful outcome of a construction project. Nevertheless, several studies in this domain have put forward the positive relation between the collaborative relationship and the end-result of a project (Lloyd-Walker and Walker, 2015; Eriksson, 2015). Especially in large-scale infrastructure projects, where people from the client (the owner) and contractor (the supplier) work together over a more extended period, the performance of the project organisation has been reported to be crucial for the success of the project (Suprapto, 2016).

Therefore, several studies and magazines highlight the relevance of assessing contractors in the partner-selection stage on their capability to relate and collaborate (Cobouw, 2016; NCE, 2015; Rahman et al, 2005). Such award criteria, which are called *collaborative criteria* in this study, aim to select the contractor(s) who could potentially have the best relational fit with the public client in terms of collaboration and team performance. One of the most intensive forms of evaluation on these parameters can be performed via assessments, including fictitious cases, games or workshops (Portilla, 2020; Kamminga, 2012). The intent of using assessments as part of the tender process has been mentioned as "*to gauge whether a contractor's behaviour and working style will be a good fit with the client team*" (Dewberry et al, 2018, p2).

Research objective and question

Construction-related literature does not properly address what kind of instruments are used to assess these collaborative criteria and what the client's experiences are afterwards. The objective of this study is to give both more academic and practical insights by (SQ1) identifying the motives of public clients that have applied these collaborative criteria; (SQ2) describing the instruments and formats that are used to assess the contractor's collaborative behaviour; (SQ3) collecting the experiences, and based on these observations and lessons; (SQ4) deriving the (improved) procedure for (behavioural) team assessments.

The main question is <u>what are the client's motives and experiences of using assessments to evaluate the contractor's collaborative behaviour as part of the tender process in infrastructure projects?</u>

While the Dutch context forms the starting point, observations from Finland and the United Kingdom are collected to put the Dutch application into a broader perspective and derive potential improvements.

An exploratory desk and case study

The first part of this study focusses on a quick desk study using both scientific literature as well as project-related articles and documents. The observations were helpful to ensure this research complies with the current practice and latest knowledge on team assessments and in this way, design a proper case selection and interview protocol for the case study. One of the findings concerned the variety of team assessments across those countries. A wide range of applications and weighting (10%-50%) was observed. Most notably, the desk study showed a still-growing use of assessments while literature lags behind. Over a hundred European infrastructure projects have conducted workshops or assessments in their tender phase.

The second and empirical part of this thesis consists of a case study, including semi-structured interviews with senior professionals of client organisations. The case study focussed on ten different cases across the Netherlands (six cases), as well as the United Kingdom and Finland (two cases each). Every single case embodied a part of the procurement phase of a medium to a large scale infrastructure project. Due to the limited amount of resources and time, it was chosen to include more cases and clients (abroad), instead of including the contractor's view as well. After all, it is the owner to decide on the delivery method and tender criteria during future construction projects (Eriksson and Laan, 2007). In order to enrich and validate the observations from the client's representatives, five qualified assessors across the Netherlands, Finland and the United Kingdom were interviewed subsequently. These independent assessors design the assessment exercises and score the tenderers' teams on a predefined list of sub-criteria.

The client's motives

Based on a thematic analysis, this study identified three motives that can be related to a majority of the cases: 1) applying collaborative criteria because the (relational) contract and project demand for an intensive collaboration in the design/realisation phase, which could be via an integrated team; 2) applying these criteria because other criteria focusing on price and content were possible to a lesser extent; 3) applying these criteria to emphasise on the importance of soft skills, as signalling function to the market. Those three motives form the answer to the first sub-question and are shown in the upper row in Figure 1.

Based on these observations, it can be concluded that collaborative criteria follow the development of relational contracting and early contractor involvement (ECI). All cases used some relational contract form (e.g. project alliances, design teams), without having selected the cases on these characteristics.

The instruments and formats

Since it was a requirement to be selected, all ten cases used a variant of the team assessment (also called *selection workshops* or *behavioural assessments*) as pre-selection or award criterion. Next to this instrument, multiple cases also made use of cultural alignment instruments; like written documents and (panel) interviews, and individual matching instruments like an EQ-i test. The team assessments within the ten cases all consisted of involving independent assessors to score the bidding parties on a predefined list of competencies. Guidelines and a robust procedure including several exercises are used to ensure a level playing field and conduct the assessment according to the procurement regulations.

In one of the cases, it was chosen to assess the tenderer without the client's team participating. In all other cases, the client's team provided the context and interaction that the tenderer would face during realisation (Kamminga, 2012). The number of participants (per side) varied from 4 until 25 individuals. Results revealed that the variety of applications relates to the different definitions of the sub-criteria. Team assessments that focus on the cultural fit between client and tenderer bring different competencies on the scoring card compared to assessments emphasising on the tenderer's alliance capability.

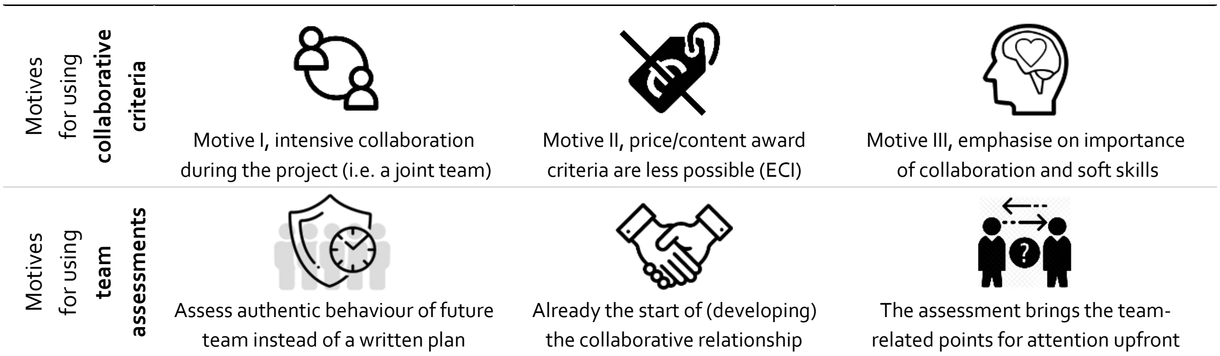


Figure 1 | The motives of using collaborative criteria and team assessments

In all cases, a conscious choice was made to use a team assessment next to the other instruments for three reasons: 1) it assesses authentic behaviour instead of an ambitious written plan, 2) the future collaborative relationship has already been brought forward and developed in the tender phase, 3) the challenging environment of simulated assessments results in knowing where one stands, the good and bad points.

The lower row in Figure 1 shows these three distinguishing elements of a team assessment. Together with the answer on the first sub-question, Figure 1 shows the six drivers of applying the team assessments.

Having applied team assessments

Both clients and assessors consider the team assessment as a valuable instrument that has a strong signalling function. Across several cases, it was observed that the bidding parties were looking towards collaboration and their team-composition in a different way: more attention is paid to soft elements compared to more traditional tenders. Moreover, the subsequent meetings and project start-up after the team assessment benefit from the interaction and reflection of the team assessment. However, clients who applied an assessment to develop and enhance collaboration see this instrument as just one of the ways to do so. A tender follow-up, the competitive dialogue and team-building could serve that as well.

Despite some prerequisites for further use, in all cases the reflection on the assessments that were being used was predominantly positive. This study found seven more operational experiences of the specific course of team assessments. Some of these lessons are used for developing recommendations.

- Expectation – Bidders need to know the rationale for the criteria and what to expect at the assessment;
- Preparation – A dry run and practising by the client's team is crucial to ensure a level playing field;
- Implementation – Usually a full assessment day and a mix of exercises is required to score properly;
- Distinction – Each tenderer's team brings a different interplay and outcome, despite the same case;
- Evaluation – Putting effort in the review of the scoring is greatly appreciated and serves transparency;
- Interaction – The following competitive dialogue and interaction benefit from the assessment interplay;
- Collaboration – Ongoing attention is needed to utilise the assessment focus on collaborative behaviour.

Interpretation and improved team assessment procedure

The desk and case study showed that the British sector is leading in this domain. The Dutch clients in the infrastructure domain could learn from the comprehensive guidelines and application of British team assessments. The British approach is not about a higher weighting of these criteria; on the contrary, the cornerstones are more accurate and more detailed. Dutch team assessments could be enlarged by including prior risk assessments, more instruments and observers. Moreover, the embedding of assessments can be improved by appointing a behavioural team on the project and maximise transparency. Future Dutch project alliances could also look at the extensive workshops by Finnish clients. These (two-day) workshops involve several levels of the future alliance organisation, which is not seen in the Netherlands yet.

The answer on the fourth sub-question provided the observed steps of the assessment procedure, including improvements based on previous observations (see Figure 2). Steps number 1, 2 and 3 correspond to pre-assessment activities and take several months, step 4 is the team assessment itself, and step 5 is related to the scoring and completion of the assessment. Some improvements are shown at the bottom of Figure 2.

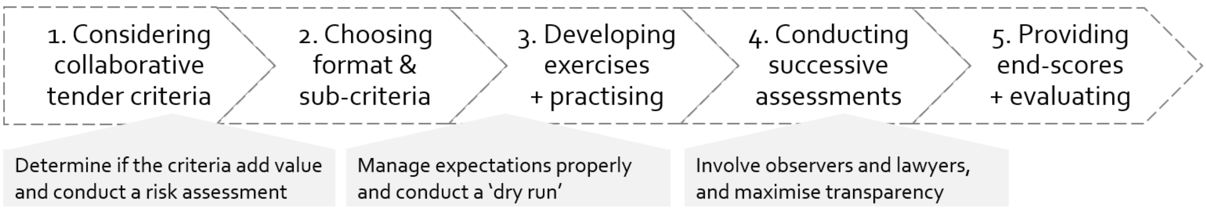


Figure 2 | Procedure for implementing team assessments and its improvements

Conclusion and recommendations

In this study, it was empirically found that the (increasing) use of collaborative criteria can be linked to the developments of relational contracting and early contractor involvement. Across these projects, public clients prefer the use of team assessments to assess authentic behaviour of the tenderer's project team that will be working on the project after awarding. Besides, the team assessment puts 'collaboration' high on the agenda of all parties right from the start (Bloemendaal & Van der Geest, 2011). However, these instruments are no miracle cure. Other instruments can lead to the same goal, and ongoing attention is needed to enhance collaboration. Especially the attitude of the client organisation is considered to be crucial.

The instrument has proven to enable the client to make a more informed choice for a contractor on its collaborative behaviour. In this way, team assessments can fulfil the potential of being one of the instruments to improve the client-contractor collaboration within the infrastructure sector from the beginning of a tender. The main recommendations of this study for the working field are displayed the list above and in Figure 2. Future research could build on this study by linking the factors related to project success to the sub-criteria that are used for the assessment scoring.

Table of contents

List of tables and figures

- this page is intentionally left blank-

Part I

Evaluating a potential partner's collaborative behaviour and cultural fit is seen more and more in infrastructure procurement exercises. This study aims to identify the client's motives, instruments and experiences of assessing the tenderer's collaborative behaviour by focussing on the team assessments in particular. In this first part, the research context, objective, questions and method are described in more detail.

1.Introduction

Investing in the relationship between the client and contractor cannot guarantee a successful outcome of a project. However, several studies have put forward the positive effect of a good collaborative relationship on the end-result of a project (Lloyd-Walker et al., 2015; Eriksson, 2015). Especially in large infrastructure projects, where people from both parties work together over a more extended period, the performance of the organisation has been reported to be crucial for the success of a project (Suprapto, 2016).

The growing attention for this collaborative relationship is reflected in new relational contracting forms that are progressively focussing on the acknowledgement of mutual dependence and benefits between client and contractor (Florence, 2014). In the Netherlands, the necessity to change procurement practises improve future collaboration and recognize the mutual benefits and is reflected in the 'Marktvisie' (2016). This vision document was an initiative developed by Rijkswaterstaat and several market parties within the construction sector. One of the pillars is of this market vision focuses on tendering; "*using the tender to lay the foundation of and emphasize on a future collaboration*" (Rijkswaterstaat, 2016 p3.).

This study aims to contribute to this higher goal of improving (and not automatically increasing) the client-contractor relationship from the perspective of the tender phase. In any case, the relation between those two parties is seen as the most critical partnership between all parties involved in a construction project (Dorée, 2001). The overall project team, consisting of client and contractor team members, is crucial for this relationship and performs better "*when relational and collaborative traits are present*" (Bresnen & Marshall, 2000). Therefore, public clients focus more and more on mechanisms to select the best potential contractor in terms of building a collaborative relationship and working effectively together (Portilla, 2020).

The most intensive form of such a mechanism is the evaluation of tenderers teams during team assessments. These assessments include simulated cases and make it possible to assess teams on their collaborative skills and their potential match with the client (Kamminga, 2012; Doyle and Jones, 2009). The precise goal of using assessments as part of the tender phase within construction projects has been described in various ways: "to *isolate the precise attitudes and personality traits they're [clients are] after*", "to *evaluate how potential partners would perform in the future*" and, "*to gauge whether a contractor's behaviour and working style will be a good fit with the client team*" (Dewberry et al, 2018, p2; Mitchell, 2016).

1.1 Introduction to the subject

The team assessment, as introduced above, is just one of the variety of instruments that focusses on 'collaboration' as part of the tender evaluation. Besides these assessments, selecting a contractor on their collaborative profile and fit is also done by, among other instruments, written proposals and interviews (Kamminga, 2012; Puckett, 2007). This study uses the term 'collaborative criteria' to label these soft tender criteria focussing on the client-contractor collaboration somehow.

> **Collaborative criteria** "*help the bid evaluator – the client – to identify a tenderer with good collaborative performance*" (Liang et al, 2019), i.e. focus on "*the capability of contractors to collaborate and perform with the client in a joint project team*" (Portilla, 2020).

Eriksson (2010) state that the use of these soft criteria could be a valuable impulse of improved client-contractor collaboration. The term 'collaborative criteria' is not used often in construction-related literature but can be understood in the context of concepts like 'collaborative contracting' and 'collaborative procurement'. These two concepts have gained increased attention over the last decades and recognised the influence of procurement on the collaborative environment of a construction process.

Since the early 1980s, clients across the construction industry have been focusing on improving the collaborative process throughout the project. In the search for ways to improve this process and the future project performance, ideas of collaborative and relational-contracting have been used. These theories are based on the idea that the relationship itself is the most important incentive to comply with agreements in long-lasting relationships between contracting parties (Kamminga, 2009). It has also been described as a 'trust-based way of working together' (Rahman & Kumaraswamy, 2002). Research has shown that these mechanisms offer flexibility, enable team building and simplify ongoing contractual relationships. In recent decades, relational contracting has been introduced at construction sites in the United States, Australia, and extended to Europe, Asia and Africa (Kamminga, 2009). The principles are noticed to be the foundation of partnering, alliancing, joint venturing and other models (Rahman & Kumaraswamy, 2002). From a more theoretical viewpoint, the concepts of relational-contracting and collaborative procurement can be seen as the ideological opposite of competitive procurement (Cheung et al., 2003; Eriksson and Laan, 2007).

> **Collaborative procurement** (or cooperative procurement) aims to enhance collaboration between the client and -contractor(s) throughout the procurement procedure (Eriksson, 2010).

The link of this concept with 'collaborative criteria' becomes apparent when Eriksson and Westerberg (2010) describe different aspects of collaborative procurement, among which, selected tendering, soft parameters in bid evaluation and collaborative instruments (Eriksson & Westerberg, 2010).

Based on a literature study, Eriksson & Westerberg (2010) suggest that collaborative procurement approaches have a positive effect on the project outcome (Eriksson & Westerberg, 2010). However, researchers state that the specific impact and eligibility of a collaborative procurement approach depends per project (Walker et al, 2013; Bosch-Sijtsema and Postma; 2009). The benefits gained from collaborative procurement have relatively more value in (large) construction projects that are characterised by uncertainty, complexity and a high form of customisation. In these projects, investment in collaboration throughout the tender phase is more beneficial (Eriksson, 2010).

Identified research gap

Several researchers have highlighted the possibility of assessing bidding parties with team assessment or other collaborative criteria (Dewberry et al, 2008; Kamminga, 2012; Doyle et al, 2009). However, those studies do not clarify the domain of projects that should apply these criteria in the tender process. It is not clear if this is the same domain of complex and dynamic projects that Eriksson (2010) described.

This gap raises the research question of what kind of projects use collaborative criteria and which motives are underlying in terms of the project characteristics and governance form. The second question that is not answered in construction-related literature focusses on the instruments that are used for assessing these criteria. Some studies highlight the relevance of assessing project partners (Pansiri, 2007; Rahman et al, 2005), but it is not stated what different assessment formats are used to do so. Moreover, the client's experiences of applying this instrument are not collected yet in more scientific literature. Figure 3 shows these three questions represented by black boxes. It is not clear what considerations and processes take place within these three boxes, although the fact that collaborative criteria are applied frequently nowadays.

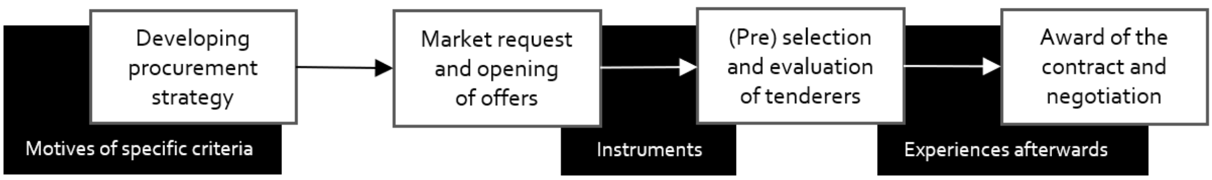


Figure 3 | Illustration of the identified three research gaps

Since 2010 assessing 'the collaborative potential of contractors' has grown significantly. On a European level already more than a hundred public infrastructure projects assessed the tenderer collaborative behaviour over the last decade (Straatbeeld 2020; Financial Times, 2019; Cobouw; 2016; Jensen, 2015). This explorative (and exploratory) study is a first response (on an European level) to this increased use of collaborative criteria. It aims to collect and describe the instruments, the matching motives to do so and the client's experiences afterwards. In this way, the three black boxes are coloured.

1.2 Research objective & questions

Based on the identified research gaps, the main objective of this study is to give more academic and practical insights by learning from public infrastructure projects that have conducted team assessments in their tender processes. This objective mainly focusses on exploring the motives and experiences of several Dutch, British and Finnish clients. These three European countries are considered to be leading in this domain.

Since the client's motives and experiences are (potentially) closely related to the implementation of the assessment, this study also discusses the different instruments and formats to assess contractors on their capability to collaborate and relate with the client. This third black box is crucial to understand and interpret all observations that are related to the other black boxes (the client's motives and experiences). A last objective focusses on interpreting all these observations and derive a general assessment procedure.

Clients could apply the more operational findings of this study to assess bidders based on their collaborative behaviour. Recommendations of this study will be relevant for public clients in the European infrastructure but also act as a starting point for further research on this topic. The explorative character of this study is reflected in the absence of any theoretical framework and broad focus rather than an in-depth focus.

The objective of this study is translated into a research question below. It must be noted that last objective is seen in the fourth sub-question but not clearly reflected in the main question. This is due to the fact that the fourth sub-question acts as a link between the primary findings and the recommendations of this study. The research strategy will focus on the main research question.

1.2.1 Research questions

<u>The main research question: what are the client's motives and experiences of assessments that evaluate the contractor's collaborative behaviour as part of the tender process in infrastructure projects?</u>

In order to answer this main question, the research question has been divided into several sub-questions. This helps to make the research and interviews comprehensible and structured.

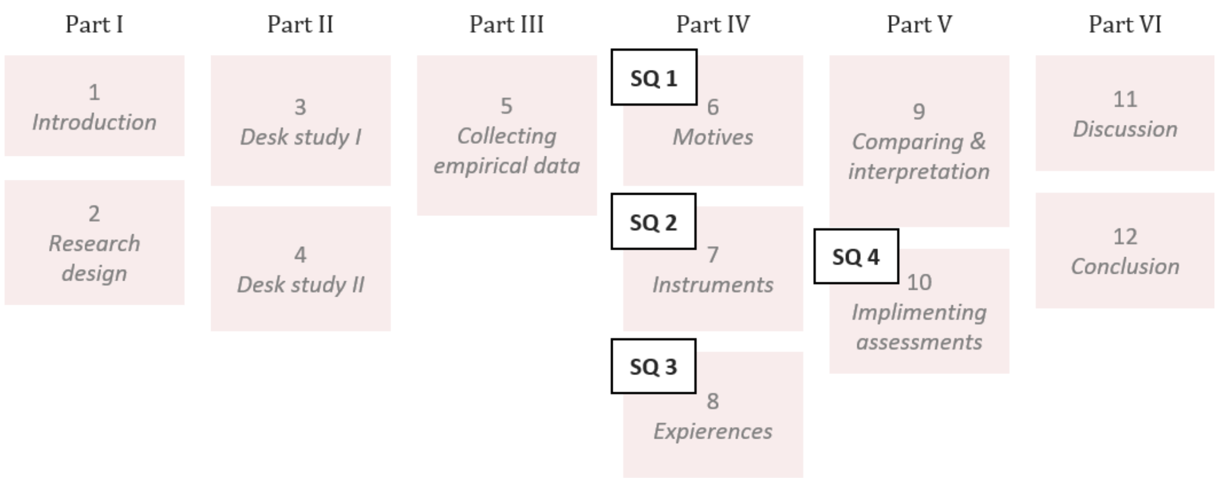
SQ 1. What are the client's motives to apply collaborative criteria in the tender process?

SQ 2. What instruments and formats are used to assess the contractor's collaborative behaviour?

SQ 3. What are the client's experiences from conducting team assessments?

SQ 4. What is the procedure of implementing a team assessment, and how can this be improved?

The sub-questions will be answered in Part IV and V of this report:



2. Research Design

This second chapter describes the used methodology and scope of the research; both are essential to conducting proper research. First, (2.1) the research strategy of this study is presented. It is based on the considerations and methods of '*Designing a Research Project*' by Verschuren and Doorewaard (2015). Subsequently, the methods applied, a twofold desk study (2.2) and case study (2.3), are agued and described in more detail. This is followed by the scope of the research (2.4) and reading guide for this report (2.4).

2.1 Research Strategy

This research is exploratory in nature since little knowledge on the specific topic is present in other studies. Because there has been little scientific research conducted, it is not expected that the answer lies in theoretical sources, and this research is practice-oriented.

The study will use a qualitative method to observe and identify collaborative criteria applied in current projects. The research objective, as set out in the first chapter, asks for practical insight, and this can only be found when a research focus is on depth (qualitative) rather than breadth (quantitative). However, the main research question also requires an exploration of the domain in terms of studying a considerable amount of different cases. This balance between in-depth and breadth is discussed later on.

According to Yin (2003), the preferred research method is based on three different aspects: a) the nature of the research question(s), b) the level that the research has a control on the events that are being studied and c) the focus of the study in terms of current or previous events. Following these three answers, several research methods could be considered – from a survey and case study until a desk study and experiment (Verschuren & Doorewaard, 2010). This also depends on the data required to be able to answer the research question. Since this research includes several "what?" questions, but also the contemporary events cannot be controlled, a case study seems the best approach. It is believed that a case study could show what and why certain decisions were taken, how they were implemented and with what consequences. These "decisions" could also be "events" or in this case, the different assessments.

A case-study is chosen because this fits the qualitative and practice-oriented approach of this study best. A twofold desk study is conducted to understand the observations of the following case study and give input for the interview questions. The two following paragraphs describe the methods being used to collect data.

2.2 Desk study

As the first method, the desk study provides the context for the rest of the study. As described above, the desk study is needed to understand but also interpret the data from the following case study.

Scientific literature

For this reason, the desk study will start in the field of collaboration and procurement in the construction industry. This also includes a short explanation of the procurement law since it sets the boundaries of tender processes, evaluation criteria and reviews – mainly since this study focuses on soft criteria that have the image of being hard to quantify and hard to assess (Kamminga, 2012; Erikson, 2010).

Search engines like Google (Scholar), Scopus and repository are used to find scientific papers, reports, books, theses and procurement legislation to gain a first insight into these topics.

Practitioner literature

A second desk study is conducted to gain some first insights into the current application of collaborative criteria in the three countries that are being studied. This prior knowledge is needed before the case study selection and designing of the interview protocol. Without having an idea of the concrete developments over the last years in this domain, the interviews would have less value.

The desk study can be seen as a first short review into collaborative criteria, while the empirical data from the interviewees will make it possible to answer the research (sub) question(s). So in short, it will act as needed fundament for the following main part of this study.

To be able to find the suitable and relevant literature, project documents, governmental reports and news articles, different keywords and synonyms are searched for (i.e. 'Behavioural assessment + infrastructure', 'Team assessment + tender', 'Assessing tenderers collaborative behaviour' and 'Workshop with client and contractor). The different keywords are used in several search engines such as online Construction Magazines, Google (Scholar), and the websites of public clients and ministries in those three countries. By studying magazines and websites within the construction industry and through contact with assessors in this industry, the feasibility of this study was also re-confirmed in this early stage.

2.3 Case Study

As described above, this is empirical and qualitative research balancing between in-depth and broad focus. The third characteristic of this study is that it is exploratory. Exploratory case studies are mostly done in the initial phase of the research to explore the area, which correspondents to this research objective.

This exploratory case study includes ten cases across the infrastructure sector of three different countries. The ten cases in this research will be studied primarily by interviews, and when relevant, a short document analysis of the tender evaluation report. These interviews will be semi-structured since both structured and unstructured approach have more limitations. In a semi-structured approach, there is a list of some 'cornerstone' questions which are asked every interview and used to conduct some thematic analysis, but there is also the chance to dive into new observations. This increases the value of the observations generated in these interviews. The interview protocol(s) can be found in Appendix B.

In-depth versus more cases

Focussing on just a few cases, each having several interviewees, is mostly seen in fully in-depth case studies. This exploratory study aims to collect motives and experiences that are primarily case-related and not interviewee-related. In other words, multiple interviewees per case would increase reliability, but have less added value for answering the research questions compared to an extra case.

Conducting interviews is highly time-consuming, so this study had to find the optimal balance between the input per case and the number of cases, which will lead to a larger set of experiences. For almost every case, just one interviewee was conducted within the client organisation. Most of the time, this person was in charge of the process of choosing for and implementing the collaborative criteria. So the design of this case study is primarily focused on gathering a large set of experiences - from different projects, different cases.

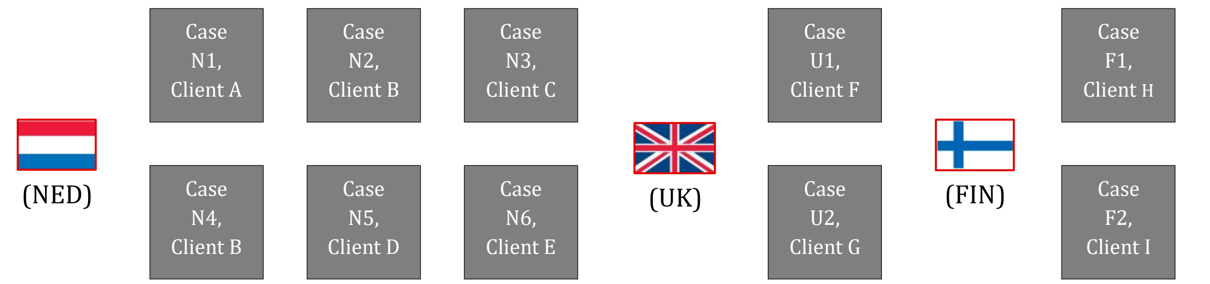


Figure 4 | Overview of the ten cases of this study

The first chapter already elaborated on the aim and purpose of this research. As shown in Figure 4, six Dutch cases act as the main focus of this study and give an impression of the current state of this domain in the

Netherlands. The four cases abroad are used to study this domain from a wider perspective. Every single case embodies a part of a tender process of a medium till large scale infrastructure project.

Assessors involved

To get a closer understanding of the instruments used for collaborative criteria but also involve the view of the client's supporting party in those assessments, this study involved five assessors as well. This amount of five is proportionally distributed across the Netherlands (3), the United Kingdom (1) and Finland (1). The interviews with the assessors have the objective to a) get an understanding of the whole process of partner selection based on collaborative criteria b) discuss the different instruments used and its developments, c) derive improvements for the team-assessment in particular.

For both sets of interviews, a different interview protocol is used. The interviews are recorded, written down and analysed via a thematic content analysis method (Braun & Clarke, 2006). Since this study is conducted from a construction management (client) perspective, a more in-depth view of relevant competencies from a psychological perspective is out of scope. In this way, the five assessor-interviews of this study form the starting point for new research combining these two domains.

2.4 Scope of the research

It is essential to make some well-balanced decisions regarding the boundaries of the study being conducted. The objective of this explorative research brings the risk of having too less focus. The following paragraph describes what is excluded and why. Firstly, this study is limited to the infrastructure domain of the Netherlands, United Kingdom and Finland. These countries, as described in the introduction, are leading in this domain of relational contracts. European procurement law can be seen as a common base. However, the (legal and cultural) context differs great per country. For this reason, the motives for using collaborative criteria (Chapter 6) are discussed within the context of their country.

This study is conducted in the infrastructure sector, which can be seen as a part of the broader construction sector. Real estate and renovation projects are also experimenting and using more collaborative criteria. However, they are operated in a different context: a domain with different challenges (Kamminga, 2012). Moreover, this research focusses on public clients of infrastructure projects, such as the national department for transport, railway authority, municipalities or waterboards. This scoping follows the fact that the (legal) procurement procedure and project characteristics/values of public projects differ from private infrastructure projects (Arlbjørn and Freytag, 2012).

While designing this research strategy, it is considered to include the view of the contractor as well. Collaboration is always about the interaction between both parties. However, it is the limited amount of resources and time that makes it reasonable to focus on just the motives and experiences of the client involved. When it is the client's perspective that a project has benefited from the use of those collaborative criteria, that perspective has an absolute value on its own – without knowing the contractor's perspective. It will be the client to choose for the tender procedure and criteria in the projects to come and setting the basis for the governance of construction projects (Eriksson and Laan, 2007). Involving assessors can be argued since they act as a sort of 'extension' of the client. Assessors support the client with their expertise in realising the goal of using collaborative criteria. That makes it necessary to include the detailed considerations of the assessors involved (instead of those of tenderers).

2.41 Definition of team assessment

Since literature and guidelines in the Netherlands, the United Kingdom and Finland use different words for the assessment of a team in the tender process, the Dutch term is used as the over-all term. The different words partly relate to the different content of those assessments. This will be elaborated in Section 7.4

The Netherlands:	*'Team assessment' or 'collaboration assessment' (Assessor 1, 2 and 3)*
The United Kingdom:	*'Behavioural assessment' or 'behavioural workshop'(Assessor 4)*
Finland:	*'(Selection) workshop' (Assessor 5)*
Term used in this study:	***'Team assessment'***

2.5 Research Report

This report is structured via multiple chapters in six different parts. This part - Part I - introduced the topic and set out the research design that's followed. In Part II the context of procurement and relevant legal aspects are summed up in order to set the boundaries of this domain. This is followed by Chapter 4, which is a first investigation into 'collaborative criteria' via practitioner literature.

These first two parts form the input for the case study design as set out in Part III. Chapter 5 includes the case study selection criteria, the set-up of the interviews and a brief description of each case. The following results are presented in the three different chapters of Part IV. The different motives, instruments and experiences of assessing the contractor's capability to collaborate will be discussed chronologically in Chapter 6, 7 and 8. At the end of each chapter, the corresponding sub-question is answered shortly. These observations and conclusions will be used for the interpretation in Part V.

The derived procedure for team assessments in Chapter 10 is one of these recommendations and forms the answer to the last research sub-question. Part VI provides the conclusions, practical implications, and recommendations for further research. Discussing the limitations and findings in Chapter 11 is needed to put the final conclusion of this research (Chapter 12) into the right perspective.

Part II

This second part consists of a desk study that addresses literature on procurement of infrastructure projects, as well as the use of collaborative criteria nowadays. The literature perspective of Chapter 3 is used to understand the possibilities of soft award criteria while the review of Chapter 4 helps to identify the development and different views on team assessment.

3. Procurement of infrastructure

This study examines the team assessment and broader domain of collaborative criteria from the perspective of public procurement and project partnership. Understanding these terms and its developments is not only crucial for interpreting the observations in this report. It also provides a kind of framework which helps make appropriate recommendations towards the end of this study that are in line with current practice.

After introducing the procedure of (3.1) procurement of public infrastructure projects, including the procurement process, this domain is viewed from (3.2) a legal perspective.

3.1 Procurement of public infrastructure projects

Procurement is defined as the process of finding a suitable supplier for the execution of the desired product, delivery or service, finalised in a contractual agreement (IBR, 2016). It can be seen as a method or part of purchasing. Procurement within a public context – projects with one or more governmental bodies that act as client or purchaser – is called *public procurement*. Infrastructure projects like bridges, roads and railways are typical examples of works that are publicly procured.

It is in that line of reasoning that clients must focus on open and fair procurement and best value for money. The European procurement law regulates all procurement in Europe (above the thresholds and only for (semi) public owners). This is done in order to provide a level playing field between all supplies within the European market (European Commission, 2017). The essence is based on the three well-known pillars of transparency, non-discrimination and competition, and is discussed in the following sections. Public clients are entities (states, provinces, municipalities and water boards) that have to follow this law.

The European procurement law is still developing and has a positive influence on innovation and competition within the European construction sector (Chao-Duivis et al, 2012). A significant change that is relevant for the objective of this study (soft parameters; collaborative criteria) is the new regulations from 2014 that aims to ensure that public owners across Europe will procure and purchase best quality and value for (public) money (European Commission, 2017). The new legislation included the use of the so-called 'Most Economical and Advantageous Tender' [MEAT/EMAT] method as the new standard. Additionally, it includes the possibility of an innovation partnerships delivery model as a mechanism for enhancing innovation – this is done by also facilitating some form of dialogue with tenderers during the tender phase.

3.1.2 The procurement process and criteria

The general (public) *procurement process* follows six general steps and is characterised by putting the contract out to tender to the market at the same time, on the same conditions, terms and procedure rules (PIANOo, 2019a). Figure 6 below shows these steps from the view of the client.

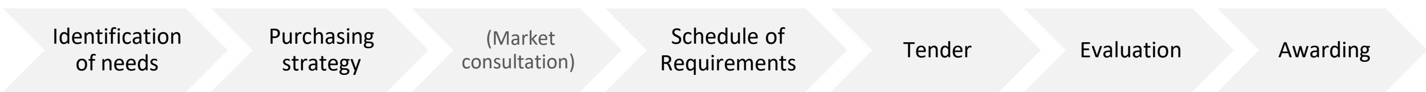


Figure 6 | Procurement process, based on IBR, 2016

Despite these general steps of this process, the specific *procurement procedure* (with the help of which a contracting authority tries to come to awarding a contract) differs for each project. It is crucial to review per project which procedure is the most suitable and proportional since it will be the procurement procedure that will define the specific evaluation of the bidding parties on pre-defined criteria (Bruggemen et al, 2010). For this study, it's relevant to elaborate on the difference between award criteria and selection criteria. This difference is relevant because it influences the potential impact of the criteria used - like those focusing on collaboration aspects. The relevant legal boundaries for using collaboration aspects in both selection and award/evaluation criteria are discussed in the following paragraph.

Selection criteria

Selection criteria are used by (public) clients to limit the number of parties that will become tendering parties throughout the following tender phase (Bochenek, 2014) As the Dutch procurement platform PIANOo described, the possibility of selection criteria "*exists in the restricted procedure, the competitive dialogue procedure, the competitive procedure with negotiation and the innovation partnership procedure*" (PIANOo, 2016). It must be noted that this study does not focus on specific prequalification requirements. These requirements ['geschiktheidseisen'] have a different function and are used to examine whether the subscribed parties are capable for the job (Pianoo, 2020).

Evaluation/award criteria

Evaluation/award criteria are applied with their own percentages to define the winning tenderer in the end of the tender phase, so to select the best tender proposal. All bidding parties that have passed the prequalification requirements and pre-selection criteria are measured by these evaluation criteria. As described previously, the use of EMAT is the new standard ('EMVI' or now 'BKPV' in Dutch). Roughly, there are two domains of award criteria: criteria focusing on project characteristics (like sustainability, risks or planning solutions) and the more soft criteria focusing on the constructor itself (experience, competence).

Each award criteria, sometimes called evaluation criteria or "wishes", cannot be legally binding (like a requirement). However, it must be noted that these criteria (could) have a crucial impact. In order to have a bigger chance of winning the project, tenderers will make sure to answer the criteria in a balanced way (Chao-Duivis et al, 2010).

3.2 Relevant procurement regulations

When the award criteria include an assessment of collaborative behaviours, cultural fit or individuals – all of which can naturally give rise to subjective interpretation - it is important that the regulations about procurement criteria are observed.

European case law ('jurisprudence') states that it is permitted to take into account soft criteria that could be seen as subjective elements, such as a presentation or interview (Brinkman, 2020). In the recently reintroduced two-phase process, it is possible that both the client and the contractor have an important substantive involvement, for example during the design phase. The way in which both organisations interact and collaborate during this design phase can have a major impact on the success of this first phase (Lloyd-Walker and Walker, 2015). Brinkman (2020) argues that it is therefore conceivable that a contracting authority deems it desirable and necessary to introduce a criterion relating to the quality of the collaboration between the client and the tenderer. This is also mentioned in the recently published report 'Guide to phased tendering' by Rijkswaterstaat (Rijkswaterstaat, 2019)

Brinkman also quotes the Pursuant to Directive 2014/24/ EU stating that following from legislation and case law, award criteria must relate to the tender and not to the tenderers. Therefore, award criteria should, in principle, not refer to experience requirements and the technical and professional capacities of tenderers (Brinkman, 2020). However, exceptions to this are possible. If the quality of the team members is relevant to the performance of the contract, the client is able to use the organisation, qualification and experience of the team members involved in the performance of the contract as an award or selection criterion.

The distinction between this selection and award criteria is relevant because the impact is different (Construction Manager, 2016). According to regulations, the client should argue not only the applied criteria in general but also the place within the procedure.

On the basis of case law, it must be completely clear to candidates which quality requirements must be met, and the assessment must be made on the basis of the most objective system possible (Mitchell, 2016). Therefore, the challenge lies with the client and involved assessors to translate 'subjective criteria' about collaboration and the relational fit into criteria that are as objective as possible. The clear link between the assessment to the contract terms and intended collaboration (delivery model) is also addressed by Stella Mitchell, a lawyer at Berwin Leighton Paisner in the *Construction Manager* (a magazine) (see Citation 1).

> *… Evaluation criteria must in turn be clearly linked to contract terms and outputs. Behavioural assessment cannot be used to assess general issues of culture and behaviour unless these are clearly relevant to the contract. … Both the behavioural criteria being tested and the relative weight given to those scores must be linked and proportionate to their likely impact on contract performance. So, for example, a weighting of 20/25% associated with behavioural criteria is unlikely to be appropriate unless a long-term collaborative relationship is vital to successful content delivery….*

Citation 1 | Stella Mitchell in *Construction Manager*, February 2016

As stated previously, the core of the European procurement law is to safeguard transparency, competition and non-discrimination (Chao-Duivis, Koning and Ubink, 2013). The guiding principles 'transparency' and 'objective' of these EU regulations are particularly relevant for the domain of this study. Eriksson, Volker, Kadefors and Larsson (2018) stated that the use of soft criteria by public clients make it even more critical to follow European procurement regulations and ensure the assessment rating could be defended in court.

This includes that the evaluation of all criteria should be clear, objective, transparent and argued very well. In this way, a level playing field is ensured (Eriksson et al, 2018). These more abstract principles are also translated into a piece of advice that is more focussed on the domain of this study: *"each element of a team assessment should be administered and scored in a manner which is objective and accountable and which treats all bidders equally*" (Mitchell, 2016).

This context of the procurement procedures and procurement law was provided in this chapter to a) understand the considerations within the tender process as shared by public clients, and b) be able to give recommendations that comply with the requirements and principles of procurement regulation. It must be noted that this study did not focus on the specific argumentation of clients who applied collaborative criteria. Studies like those of Brinkman (2020) that are conducted on the interface of law and procurement practices are considered to be very relevant for this domain.

4. Review of the subject

In order to select appropriate cases and derive a proper interview protocol for the case study, it is essential to be familiar with the domain of collaborative criteria and its development. Comments and conclusions that have been drawn in previous studies or more popular articles can help sharpen the focus of this research. While Chapter 3 described the context of procurement, this review discusses publicly available sources about the topic itself. In this way, the quick review makes sure this research, including its recommendations, complies with the current practice and latest knowledge on team assessments.

Based on a quick desk study, (4.1) the domain of collaborative criteria is described and (4.2) an overview of the introduction of the team assessment in the infrastructure sector is provided. For this purpose, both papers and project documents within the three countries studied (the Netherlands, the United Kingdom and Finland) are examined. Subsequently, (4.3) the current views and comments on team assessments are collected as they are the starting point for the recommendations and discussions in Part V. In section 4.3, also the results of this review are summed up so they can be incorporated in the interviews.

The three steps of this first review are shown in Figure 7.

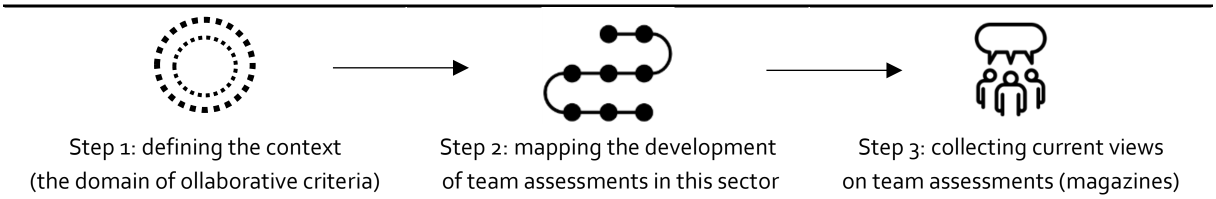


Figure 7 | Set-up of the quick scan

4.1 The domain of collaborative criteria

From a historical perspective, public clients set a very high weight on price and lower weight on soft parameters (Kadefors, 2005). As pointed out in Chapter 1 and Section 3.2, this has changed over the last two decades. There has been growing interest in a shift from lowest price selections to multi-criteria selections, also considering soft parameters (Kumaraswamy and Anvuur, 2008).

The set of collaborative criteria [criteria helping the bid evaluator to identify a tenderer on its collaborative behaviour] can be seen as a significant part of this domain of soft parameters. Several researchers mentioned examples of those soft parameters with a clear presence of relationship traits.

> "*Soft elements are mostly those related to behavioural aspects such as trust, honesty, communication, relationship and job satisfaction"* (Cheung et al., 2003)

> "*… soft parameters such as collaborative ability, reputation, earlier experience of the supplier, and shared values, which decrease opportunism and enhance trust*" (Eriksson, 2008).

> "*If hard elements are defined as discrete, quantifiable and easily observed, soft elements exist more towards the qualitative end of the spectrum* (…)" (Engebø, Skatvedt and Torp, 2019)

Sarhan, Pasquire, King and Manu (2018) have defined soft parameters as "*e.g. competence, reputation, capacity, collaborative-ability, and experience*" and put these into a one-dimensional spectrum of price on one side and the soft parameters on the other side (as can be seen in Figure 8).

High weight on the tender price	Equal weight on price and soft parameters	High weight on soft parameters

Figure 8 | Selection methodologies: bid-evaluations. Based on Sarhan et al (2018)

4.1.1 New procurement methods and instruments

Over the last decade, several procurement instruments have been created or taken over from abroad to include these soft parameters like competence, capacity, collaborative-ability, and experience.

A well-known example is the *Best Value Procurement* system. Potential contribution to collaboration is if the contractor's collaborative attitude and values can also be discussed during the presentation or interview (Kamminga, 2012). A second instrument is weighting past performance. By means of 'past performance', the tenderer can include his own experience or that of third parties with the contractor in the selection. The procurement regulations offer the possibility of taking reference works into account when selecting contractors. In this way, clients can use 'past performance' to weigh up their wish to take account of the contractor's past performance in terms of collaboration (Petit, 2010).

Conducting a team assessment to assess the cooperative skills of tenderers is one step further. By testing the 'click factor' between parties based on an assessment, it can be objectified and made part of the award criteria. A team assessment consists of a game simulation: the fictitious case is used and is not played on content, but on behaviour (Kamminga 2012). The intention is that the tenderer and contractors learn from their experiences in the simulation by, among other things, becoming aware of effective and ineffective behaviour. Parties not only get to know themselves better but also each other start (Bloemendaal & Van der Geest, 2011). Furthermore, commissioning authorities have also requested written documents focusing on a plan for collaboration or panel/team interviews.

4.2 The introduction of team assessments

Conducting team assessments within the infrastructure sector for the purpose of tender evaluation has two origins: the partner selection by other public clients and the alliance model as started in the private sector.

It was the British Ministry of Defense that started in 2001 with (behavioural) team assessments for the selection of their construction partners (Association for Project Management, 2019). Following early 2006, the construction partners of the Olympic Delivery Authority (ODA) were selected with the use of team assessments. It would take a few years, until 2014, before the British construction sector itself would apply an assessment. Over the six years to follow, the British construction sector would make a tremendous development in terms of designing the assessments (B2B, 2017). Public clients like Thames Water and Highways Agency prepared procurement procedures including these assessments.

Around that time in 2000, the potential of alliance contracting was gaining increased attention in the public sector in Australia (Wondimu et al. 2016). The procurement of one of their first public alliances in the infrastructure domain included a 2-day workshop with each contractor and focused on the assessment of both leadership and collaboration skills, but also project risks and opportunities (Commonwealth Department of Infrastructure and Regional Development, 2006).

This alliance model of the Australian government in 2006 corresponds to studies of the Finish Technical Research Center (VTT) in 2009 a few years later (Lahdenperä, 2009). The first public sector project alliance in Finland (which also applied the team assessment workshop) was the Tampere–Kokemäki rail renovation in 2011. Meanwhile, the Netherlands used a team assessment in one of their first large project alliance in the infrastructure domain, A2 Hoogelegen in 2009 (Van Leeuwen, Van der Veen, Huizer, & Maliepaard, 2011). The alliance experiences in Australia were one of the reasons to apply this model in Europe as well.

The following paragraphs will describe the developments and specific implementation of team assessments across the Netherlands, the United Kingdom and Finland.

4.2.1 Team assessments in the Netherlands

The first Dutch infrastructure project with a team assessment in the tender phase was road project A2 Hooggelegen (2009). Not only was the team assessment new, but the project alliance model was also relatively new for Rijkswaterstaat. Before the start of 'Alliance A2 Hooggelegen', there were just a few experiences with alliancing in the Netherlands: the Waardse Alliance and the Alliance N201.

In order to assess this collaborative criterion, Rijkswaterstaat made use of an assessment developed and carried out by independent assessors, in which both employees from the client's side and potential contractors participated. By the use of practical assignments and role-plays, the potential of the various possible contractor combinations was determined. This approach not only improved the partner selection for the coming project alliance but also put collaboration high on the agenda of all parties right from the start (Bloemendaal & Van der Geest, 2011). Compared to the more standard Australian Victoria alliance model, the Netherlands applied a different form of the project alliance model. Generally speaking, the procurement process of Dutch project alliances can be seen as a light variant, because it also focuses on price elements or an agreement on price before the execution phase (Elemans, 2016; Trafikverket, 2016).

Team assessments were not only used in the procurement of project alliances in the Netherlands. The more recent pilot projects under de umbrella of 'Project DOEN' also embraced more soft criteria focusing on collaboration. The first project under this umbrella used a team assessment, including case studies with fictitious differences of interest and concluded by rounds of reflection. Ten Hoeve (2018) conducted a case study focused on the experiences of the involved parties in this first Project DOEN. One participant underlined the importance of the human factor for project success and described that "*assessments under the guidance of a phycologist helped to get to know yourself and the each other*" (Ten Hoeve, 2018).

In recent years, several Dutch water boards and provinces also conducted team assessments. Especially the last three years, the attention for team assessments has increased on the level op procurement directors and project managers. Two water boards allocated no less than 40% to the team assessment weight in the tender evaluation. It can be concluded that this application makes it (probably) a decisive criterion.

4.2.2 Team assessments in the United Kingdom

The use of a behavioural assessment in the United Kingdom infrastructure sector has increasingly grown in the past decade. Doyle and Jones (2009) described that a behavioural assessment approach was first used in the construction domain when a nuclear power station had to be demolished in 2007 (Doyle & Jones, 2009). From 2014 on, Thames Water and Highways Agency prepared procurement procedures, including these behavioural (team) assessments (Construction Manager, 2016).

Another 'early adopter' of assessments was the British Environment Agency, which first used the technique during the procurement of a 10-year contract for the delivery of capital improvement projects in the Thames Estuary. The Environment Agency felt it would obtain valuable insight – and increase the chances of finding the right partner – by integrating behavioural assessment into its tender process. In total, 25 per cent of its award criteria were assigned to performance in the behavioural assessments (Construction News, 2016). Its programme director Peter Quarmby explained in an article of the New Civil Engineer: "*Because of the long-term nature of the contract, we recognised that its success was likely to be heavily dependent on finding a partner whose organisational culture aligned with ours and whose key personnel could support the achievement of our objectives*." (New Civil Engineer, 2015)

Mitchell (2016) mentioned the use of some form of team assessments in very large British infrastructure projects by clients like Highways England. Behavioural assessments accounted for nearly a third of the selection criteria used in the procurement of its collaborative delivery framework for a 5 billion pounds upgrade of major road infrastructure. Jensen (2015) described how Network Rail used behavioural assessment in the procurement of its Wessex Capacity program. In the procurement of their Wessex Capacity Alliance, 25% of total weighting focused on collaboration including a group assessment (12%), interviews (10%) and a written submission (3%) (Financial Times, 2020).

Moreover, waterboards (Welsh Water and Thames Water), a national airport and railway agency used such assessments over the last decade (Dewberry et al, 2018). According to a report by the Ministry of Treasury, the current assessments in the British infrastructure sector generally means that every bidding party is attending so-called ‘assessment centres’ and “*collectively working through externally observed exercises*” (HM Treasury, 2014). UK Power Networks (UKPN) recently used ‘*an assessment of collaborative’* as part of the overall procurement process. A two-day behavioural workshop was held at a university (JCP, 2019

It can be concluded that the development of team assessments in the United Kingdom is in an advanced stage and relatively diverse. A majority of the national public clients in the infrastructure sector have worked with those (behavioural) assessments for the evaluation of bidding parties.

4.2.3 Team assessments in Finland

Experiences from project alliancing in Australia reached Finland in the period 2005-2010. Around that time, the Finnish Transport Agency (FTA), together with Technical Research Centre of Finland (VTT), conducted studies into the potential of different alliance models (Lahdenperä, 2009). By choosing three project alliance pilots, FTA introduced the alliance model, including its team assessments/workshops, in the Finnish infrastructure sector (Lahdenperä; 2012). The first project alliance in Finland - the Tampere/Lielahti–Kokemäki railway renovation - applied the weight of 20% while others pick 25% (Netlipse, 2020). However, the format of the selection workshop stays the same. It can be concluded that the implementation of ‘collaboration criteria’ is relatively homogeneous in Finland.

The procurement procedure that’s applied in Finland shows many similarities to the Victorian model from Australia. The assessments have a relatively broad objective since the client can also ask elaborating questions about the qualitative section of the tender’s plan: ”*Discussions may focus on changes in the content of requirements and/or design as well as on tentative innovation possibilities on which a common view is sought*” (Lahdenperä, 2009). In this model, the individuals from the client and contractor organisations that form the future joint team will focus on some project-related issues in a workshop setting. In the evaluation of the assessment, attention is paid to teamwork, collaborative behaviour, leadership and outcome.

The procurement procedure, including the workshop as described above, can be seen as a template that is used for all Finnish alliances. During all their selection workshops contractors have to organise, guide and lead teamwork by involving the client's experts in the collaboration.

4.3 Current views on team assessments

Even though very few studies focused on team assessments in the construction sector, some construction-related articles and magazines discussed the potential of these assessments. Comments and doubts are collected as they are giving an impression of the view of both institutions, researchers and market parties, and used for making improvements and recommendations (Section 10.3).

Moreover, this paragraph makes sure this research complies with the current practice and the latest knowledge on team assessments. Public documents and websites in all three countries were studied.

A recent view of the ICE (Institution of Civil Engineers) highlights the importance of conducting these assessments properly within the context of procurement regulations. *“The assessments can be seen as beneficial, but they require substantial planning and preparation to develop the transparency needed for public procurement, which can be time-consuming and expensive in terms of client and bidder resources. Behavioural assessments should be kept simple until the value of more complex models and techniques is properly established, …”* - Procurement Advisory Group (PAG) of Civil Engineers (ICE, 2018).

The statement of the ICE is considered valuable since the institute is respected and the statement is published in the recent year. Some academic researchers also discussed the assessment instrument within the procurement context. Kamminga (2012) mentioned there is still a risk of subjectivity when assessing collaboration skills and Dewberry et al (2018) raised concerns to the validity of behavioural assessment processes used in construction procurement. Their study also suggests that the performance of individuals can be more effectively and more cheaply predicted by alterative selection methods like structured interviews (Dewberry et al, 2018). However, it must be noted that Dewberry, Hayes and Sarhan (2018) did not use any directly collected empirical data from construction projects.

Others views that have been found focus on the potential and positive experiences from conducting team assessments*. "Experience has shown that effective assessment centres provide an accurate representation of the cultural capability of the partners and the behaviours of the potential leaders. This approach also places an early emphasis on behaviour, culture and the importance of leadership …*". (Infrastructure and Projects Authority, 2014). And a manager from a Dutch contracting party reflected that *"The team assessment works like a mirror and magnifier in one*" - E. van Burik in Cobouw (2016).

These relatively positive experiences are supported by a citation of the British HM Treasury. They concluded that all experiences have shown that effective assessment centres "*provide an accurate representation of the cultural capability of the partners and the behaviours of the potential team*" (HM Tresuary, 2014). Yet these citations on collaborative criteria in governmental reports are very rare. In the Netherlands, just a few reports from market consultations were found. Some of them conclude predominantly positive about the potential of team assessments while others discuss the value or specific application of the instrument. *"...A number of people prefer to let the client team judge the intended contractor team itself. This instead of a review by independent assessors. Participants also suggested making a mix between the opinion of an independent assessor and the client team...."* (Market consultation Versterking IJsseldijk Zwolle-Olst, 2019). This citation from a Dutch market consultation last year shows that some market parties question/doubt the value and role of the independent assessor in team assessments. These and previous views will be included in the interviews. The operational sections of this study (Chapter 7 and 10) will discuss the different formats and measures that could be made in order to enlarge objectivity and meet all procurement regulations.

4.3.1 Results of the review

The implementation of collaborative criteria in the procurement phase of infrastructure projects is seen repeatedly across Finland, the United Kingdom and the Netherlands. This review reconfirmed the scope of this study by focusing on these three leading countries. Moreover, the identified views in practitioners journals and studies show that there are mixed experiences and doubts about the assessment as instrument for partner selection purposes. This underlines the need of studies like these.

This review also exposed the variety of the current implementation of collaborative criteria. Moreover, it follows from this quick review that the interpretation of the selectable cases should be made with the different context of each country in mind. The origin of the identified team assessments lay in both the British Ministry of Defence and Olympics Delivery Authority as well as the (Australian) use of the alliance model. The interview protocol will not be adjusted to these differences beforehand (to prevent any bias), but the following questions in an interview will be prepared on the results of this desk study.

Part III

The main focus of this explorative study lies on collecting empirical data from projects that have applied collaborative criteria, via the team assessment in particular. This fifth part describes the design of the case study and the selected cases in more detail. Observations from Part I are used as input for the case selection and interview-protocol. Part IV will present the findings of the case study.

5. Collecting Empirical Data

The main part of this study consists of collecting and interpreting empirical data. With a dozen of interviews, this study strives to gather real experiences instead of describing the potential from a theoretical perspective. This chapter describes the decisions regarding the way these observations were collected.

First, (5.1) the case study approach as set out in Section 2.2 is elaborated in more detail. This is followed by (5.2) the selection criteria that were used, (5.3) the cases that were selected and (5.4) the way this study gathered its empirical data. Ten different cases form the basis of empirical data gathering for this research. In addition, five assessors were interviewed to include their view as an assessment provider.

5.1 Case study approach

This explorative research has the goal to identify the motives, instruments and experiences of collaborative criteria applied in practice. Following the considerations of Yin (2003), the objective and characteristics of this study resulted in conducting a case study. This research method could bring forward why certain decisions were taken, what the effects were and how the corresponding instruments were implemented (Yin, 2003). Moreover, this study is a typical exploratory one since those are most of the time conducted in the early phase of exploring a certain domain. The desk study and quick scan in this report already showed that this field of study is relatively new.

The case study will be carried out in two stages. The first phase consists of studying each case independently. This includes analysing the interviews with representatives from the client-side. In some cases there is also some project-related documents studied, e.g. Value for money report or the tender evaluation. This will lead to different brief reports of each single cases (Appendix C). The first stage will be followed by a second stage of finding explanations, differences and similarities for all observations. This will be mainly done in Part IV and V of this report. A thematic analysis is conducted previously.

An in-depth comparison (cross-case analysis) would require that cases must be mostly similar (Seawright, 2010). This is simply not possible with the diverse and limited list of suitable projects. However, there is still a common ground of all ten cases that are being studied, which will make it possible to compare the observations from different case studies and put these into a broader perspective.

5.2 Case selection criteria

All cases have to meet the case study selection criteria. The four selection criteria below will lead to a particular set of projects, which will be the set to choose from.

5.2.1 Selection criteria

- The construction project is 'public': the client is a public authority;
- The project is an infrastructure project (broad definition: rail, road or water, bridges, tunnels etc.);
- The project is (being) realised in the Netherlands, the United Kingdom or Finland;
- The client applied (among other criteria) a team assessment as selection or evaluation criterion.

The first requirement that the cases must be projects tendered by a public client, such as the national department for transport, railway authority or waterboards. This follows from the fact that the (legal) boundaries and project characteristics of public projects differ from private infrastructure projects.

Selecting collaboration aspects is also applied in the building sector (real estate and renovation); however, the contractual models and corresponding characteristics of the client-contractor relationship (can) differ from the infrastructure domain. The research goal of giving insights and advice to clients of infrastructure projects results in the requirement that each case should be an infrastructure project.

The cases should also be projects in Europa and tendered according to the European procurement law so the same 'basis of regulation' is met. This is also reflected in some benchmarks of European public clients who scope their exploration to just European member states (Trafikverket, 2016). Besides, the requirement is also practically motivated: the understanding and communication of projects in Europe is more accessible for the researcher and graduation committee. The final and most important selection criterion requires the use of a team assessment in the tender phase. This can be the only or one of the collaborative criteria.

5.3 Selected cases

The goal of this research - identifying the different formats, motives and experiences of collaborative criteria - benefits from a diverse case study selection within the set of projects following from the previous paragraph. This diversity is defined as *a diversity of implementation (formats used for the team assessment): different weights and different assessors*. In Table 1, an overview of the final cases is provided.

In order to select ten proper cases, a long list of projects was constructed first. This list was based on the review (Chapter 4) in which projects were collected that have published any document online that mentioned their use of a team assessment. Several reports of market consultations in the Netherlands were found that discussed the value of a team assessment but did not apply this instrument (yet). It must be noted that the involvement of (any colleague of) AT Osborne [the graduation company] was no requirement at all. A majority of the interviewees is found via online sources or the researcher's network.

The shortlist of cases covers the following diversity: the Netherlands (6), and the United Kingdom (2) and Finland (2); a wide range of weighting (10% till 50%); and assessors from seven different firms.

Case	Client	Type of project	Collaborative criteria
CASE N1	Client A	Dyke improvement project	Team assessment weighting 50%; Pre-selection via vision document
CASE N2	Client B	Shore improvement project	Team assessment as pre-selection criterion
CASE N3	Client C	Dyke improvement project	Team assessment weighting 35%
CASE N4	Client B	Bridge renovation	Team assessment and a vision document as selection criterion
CASE N5	Client D	Dyke improvement project	Team assessment weighting 40%
CASE N6	Client E	Bridge renovation	Team assessment weighting 25%; Vision document weighting 30%
CASE U1	Client F	Railway project	Team assessment (part of BAC) weighting 10%-30%
CASE U2	Client G	Tunnel project	Team assessment (part of BAC) weighting 20%-25%,
CASE F1	Client H	Tunnel project	Team assessment weighting 25%
CASE F2	Client I	City tramway project	Team assessment weighting 20%

Table 1 | Selected cases with its main characteristics

5.3.1 Case characteristics

A summary of the characteristics of the cases is given below. The full report can be found in Appendix C.

Case N1 is a dyke improvement project (project alliance) in the Netherlands with a contract value of around 200 million euro. The client (Client A) used a selection round (from five to three parties) based on a vision document. The award criterion team-assessment weighted 50%. The project is now in the delivery phase.

Case N2 is a shore improvement project of a large river in the Netherlands with a contract value of more than 100 million euro. The client (Client B) used a team-assessment to reduce the number of bidders from five to three. The contract, a PDC contract form, has been awarded.

Case N3 is a dyke improvement project in the Netherlands with a contract value of around 100 million euro. The client (Client C) is a regional water board and used the evaluation criterion 'collaboration', a team-assessment, which weighted 35%. The project is now in the delivery phase.

Case N4 was a bridge renovation project in the Netherlands with a contract value of less than 100 million euro. The client (Client B) used a team-assessment to reduce the number of bidders from five to three. Client B also asked a vision document from all bidders to select the five parties for the team-assessment.

Case N5 is a dyke improvement project (a project alliance) in the Netherlands with a contract value of more than 500 million euro. The client (Client D) is a regional water board and conducted a team-assessment, which weighted 40%. The project is now in the designing phase.

Case N6 is: a bridge(s) renovation project in the Netherlands with a project value over 100 million euro. The client (Client E) used a team-assessment as evaluation criterion, weighting 25%. Moreover, a written document about the contractor's vision on collaboration and organisation counted for 30%.

Case U1 is a very large railway project in the United Kingdom with a contract value of over one billion pounds. The mega project includes several contracts: the seven main civil works parts and the designing work. The client (Client F) used a behavioural assessment centre, weighting between 10% and 30%.

Case U2 is a large tunnel project for a waterboard in the United Kingdom with a contract value of over one billion pounds. The very large project includes several subcontracts like the main civil works and enabling works. The client (Client G) used a behavioural assessment centre, weighting in the range 20% - 25%.

Case F1 is a tunnel project in Finland with a contract value of around 200 million euro. The client (Client H) applied the project alliance model and organised several workshops in the tender phase. The selection workshops weighted 25%. The project is already delivered.

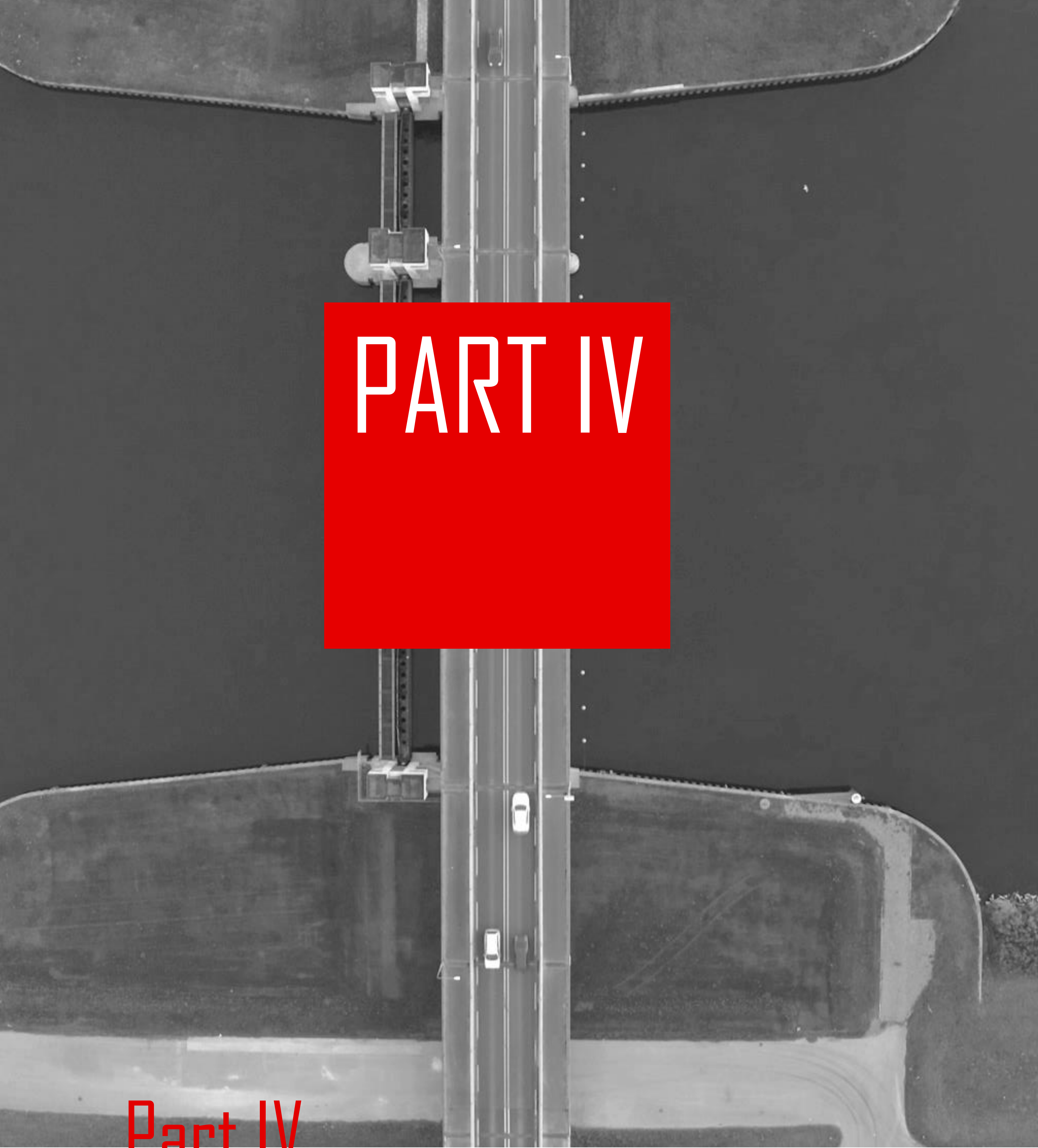
Case F2 is a tramway project in Finland with a contract value of more than 200 million euro. The client (Client I), a regional city, applied the alliance model and used selection workshops in the tender phase, which weighted 20% in the second tender phase. The project is now delivering the first part.

5.4 Data gathering

Based on the interview results (and in some case a project document review of the procurement evaluation report) a case analysis is performed. The interview data for the ten different cases is gathered via the use of eleven semi-structured interviews. As described in Section 2.3, this approach is chosen because all interviewees are only interviewed once. When the researcher is restricted by a list of predefined questions, it's not possible to dive into new observations and relevant aspects. In a semi-structured interview, the conversation follows some cornerstone questions, but the following questions can develop throughout the interview (Cohen and Crabtree, 2006). Yin (2002) identified several limitations of (semi-structured) interviews, used to gather data for cases. By developing two clear interview protocols and discussing the observations with five assessors, the weaknesses are answered as best as possible. The limitations that are the result of using semi-structured interviews are also discussed in Section 11.3. The interview protocols can be found in Appendix B for the client's representatives (B1) and assessors (B2).

The interview focussed on three pillars: the motives to use this 'collaborative criteria', the formats of the assessment used (including its extent and preparation), and the experiences afterwards. For this study, the exact function of an interviewee within the client organisation was not predefined. What is considered to be the most relevant is that the individual had a leading role in the client's organisation regarding the team assessment. Most of the time the respondents were contract managers, project managers or procurement directors. The assessors were selected on the basis of their references within the construction industry.

The (online) interviews lasted between 60 minutes to 90 minutes. The interviews have led to different brief reports of each single cases and can be read in Appendix C. These brief reports highlight the relevant observations from the interview report on a case level and give an impression of the project and criteria that have been applied. The following chapters will use certain parts from the original interview reports. These whole interview reports are not published public because these include sensitive information. The graduation committee had excess to this documentation.



Part IV

In this fourth part, the observations from the case study are presented. The different motives, instruments and experiences of assessing bidding parties will be discussed chronologically in Chapter 6, 7 and 8. At the end of each chapter, the corresponding sub-question is answered. Chapter 6 follows the broader scope of collaborative criteria, while Chapter 7 and 8 zoom in on the concrete forms and experiences of team assessments.

6. Motives for applying collaborative criteria

Over the previous years, several researchers (i.a. Portilla, 2020; Dewberry et al, 2018) and magazines (i.a. Construction Manager, 2019; Construction News, 2016) discussed the team assessment instrument. However, those publications did not focus on the higher driving motives that have led to the implementation of those criteria. The increasing use and publications about this domain make it very relevant to answer the first sub-question: w*hat are the client's motives for applying collaborative criteria in the tender process?* The motivations are particularly important in the context of a procurement procedure. Section 3.2 described the obligation as set by the procurement law to argue the need and of each tender criterion.

The first paragraph (6.1) will give an overview of the identified motives across all ten cases. These motives will be further elaborated per country since the context of the (6.2) Dutch, (6.3) British and (6.4) Finnish infrastructure sector is different (refereeing to the observations in Section 4.2). This is followed by (6.5) a reflection and discussion on the motives from a broader perspective, with both literature and the opinion of interviewees. All these observations lead to (6.6) a sub-conclusion and answer of the first sub-question.

6.1 Overview of the identified motives

This chapter focusses on the reasons to use evaluation criteria that address the future client-contractor collaboration and contractor's collaborative behaviour. Therefore, interviewees were asked about their considerations of developing the procurement strategy and tender criteria. Reasons for conducting a team assessment compared to a written document or interview are not considered relevant in this part of the study. The observed research gap (Section 1.2) touches upon the whole domain of collaborative criteria. This study aims to collect and interpret the reasons why clients are moving to this domain more and more – and not just one instrument in particular. This focus is visualised in Figure 9.

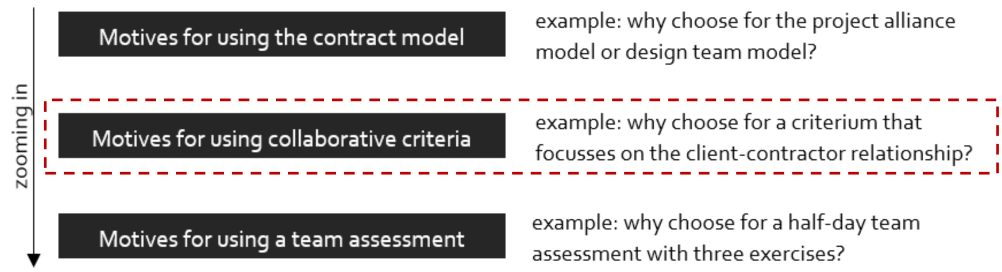


Figure 9 | The focus of the first sub-question within a broader context

The results were find via a thematic analysis of the interview data (Braun & Clark, 2006). Here, the empirical information from the interviews was documented, coded, interpreted on a case-level, and put together into similar themes. The five identified motives in Table 2 are the result of this analysis.

Motive	Description of the identified motive
Motive I	Applying collaborative criteria because the project and (relational) contract bring intensive collaboration (a joint project team) in the design or realisation phase; related to relational contracting.
*Motive I**	Applying collaborative criteria because of the complex project challenges and high uncertainty, both demanding intensive coordination and collaboration; closely related to Motive 1.
Motive II	Applying collaborative criteria because other (more traditional) criteria focusing on price and the content of the design were less possible due to the early contract involvement.
Motive III	Applying collaborative criteria to lay the foundation of the future client-contractor collaboration in the tender phase and emphasize on the importance of collaboration/soft skills;
Motive IV	Applying collaborative criteria because the partner has to share the same values and vision

Table 2 | Overview of the five identified motives

It must be noted that the motives do not necessarily exclude each other, and some of them are strongly related to one another. Because interviewees could not choose from a predefined list of motives, Table 2 is not fully comprehensive. An interviewee could agree on the presence of a non-identified motive when this would have been explicitly asked. In other words, this study gives the first impression of motives. Table 3 shows which cases can be linked to one of the motives of Table 2. The case summaries (Appendix C) elaborate on the exact formulation and presence of these motives per case.

Motive	N1	N2	N3	N4	N5	N6	U1	U2	F1	F2
Motive I	X	X	X		X	X	X	X	X	X
Motive I*	X	X	X		X	X	X	X	X	X
Motive II	X	X	X	X		X	X			
Motive III	X			X	X		X		X	
Motive IV				X			X	X		

Table 3 | Overview of the identified motives and corresponding cases

Most answers by interviewees were related to the uncertainty, complexity and dynamics that the client will probably face during the design or delivery phase (Motive I*). This motive is considered to be suited to the higher level of motives related to the contract model (see Figure 9) . When interviewees were asked more about the link between this and the collaborative criteria, the more concrete answers were always related to a more intensive collaboration (Motive I). For this reason, motive I* is seen as an indirect motive to choose for collaborative criteria and grouped under Motive I.

6.2 Dutch motives for using collaborative criteria

The six Dutch cases of this study are quite different in size: case N4 has a contract value of fewer than 50 million euros, while case N5 exceeds the 500 million euro. A majority of the cases is conducted in the domain of water infrastructure (case N1, N2, N3 and N5). It must be noted that both case N2 and N4 are projects by client B. Apart from case N1, in each case one interview was used to collect its data. Moreover, evaluation reports of case N2 and N4 were studied. Assessor 1, 2 and 3 were involved in one or more Dutch cases.

Four Dutch cases have used a collaborative criterion as an evaluation criterion. Case N2 and N4 used them to select (three) parties out of more bidding parties. Clients call this also a pre-selection or 'funnel function'.

Project	Collaborative criteria	Why collaborative criteria?
Case N1, Client A	Pre-selection: 'collaboration vision' 50% : team assessment	*Need for collaboration (project alliance), focus on new joint alliance organisation, knowledge of the market was needed*
Case N2, Client B	Pre-selection: team assessment	*Need for collaboration, people determine project success, not the contract, managing uncertainties together*
Case N3, Client C	35%: team assessment	*Need for collaboration (in joint team), managing risks together, working with market, giving (people) room for development*
Case N4, Client B	Pre-selection: team assessment and written vision	*The project offered the opportunity to put the market vision into practice with more focus on collaboration in the tender and project*
Case N5, Client D	40%: team assessment	*Need for collaboration (project alliance), managing risks and uncertainties togehter, innovation*
Case N6, Client E	25%: team-assessment and 30%: plan for collaboration	*Need for collaboration (design team), focus on equitable joint team, other criteria were possible to a lesser extent*

Table 4 | Dutch cases and their collaborative criteria

Asking the Dutch client's representatives about their motives to use these soft collaborative criteria brought the interview in almost any case towards the chosen project delivery method and relevant project characteristics. The six Dutch cases were not selected because of their early contractor involvement (ECI), but all six projects have applied some (light) form of ECI: case N1 (Project Alliance), case N2 (PDC contract), case N3 ('Dyke team'), case N4 (Joint team), case N5 (Project Alliance) and case N6 (Design Team). The right row in Table 4 gives an impression of the words and characterises that are observed to be a motive to implement the collaborative criterion/criteria. A more detailed impression is given in Appendix C.

6.2.1 Motives related to intensive collaboration (in a joint team) during the design or realisation phase

The Flood Protection Program (HWBP) and changing requirements brought new risks and challenges for the Dutch Waterboards. Client A (Case N1), client C (Case N3) and client D (Case N5) were all confronted with these changing context for their dyke improvement projects. Their motivation to use collaborative criteria is related to the early contractor involvement model.

Client C would choose for a combination of a project alliance and design team, their so-called 'dyke team', because 'the calculation method' gave much freedom and the context of rules was changing (Interviewee 4). "*And then you conclude that you are looking for someone who really does the project together with you; you need a team of good people*" (Interviewee 4). Both clients C and D were looking for two products in the tender process: a proper financial plan and a good team. "*And how do you select that team? On team competencies via soft award criteria"* (Interviewee 4). For the dyke improvement of case N1, client A used the alliance model via a competitive procedure with negotiation. The new Water law [Waterwet] of 2017, the changing context and unclear project scope were several reasons to do so (Interviewee 1). During the first phase of developing the procurement strategy, client A figured out they needed the market in a collaborative way, not in a traditional competitive way (Interviewee 2). "*It was the uncertainty of the pre-exploration and the market conditions that made us decide to tackle the challenges together ... And that is where these soft criteria like a team-assessment and vision on collaboration came in*" (Interviewee 2).

The client's representatives of case N1, N3 and N5 all linked their use of collaborative criteria to the choice for a relational contract. "*We wanted to work with a contractor who does not think 'in the contract', but who thinks in terms of cooperation. The only way to assess this is by using a team assessment*" (Interviewee 6).

Besides those Waterboards, the clients of case N2 and N6 also argued their use of collaborative criteria in terms of the project's risks and uncertainty. Client B and E both faced many risks concerning the actual state of the asset and the way those should be handled. Comparable projects by client B had many contract changes concerning these uncertainties. That made the client's team of case N2 raise the question '*how can we manage these risks and use our resources and taxpayers' money as well as possible?*'. Their answer included the PDC (Plan, Design, Construct) delivery method while Client E would choose for the design team model. Both approaches would result in more collaboration between the client and the contractor. "*We need to seek collaboration on certain aspects actively. Collaboration simply has a dominant factor in this project. With that in mind, we looked to the tender process*" (Interviewee 3).

6.2.2 Motives related to the fact that price/content criteria were possible to a lesser extent

The second motive to apply collaborative focusses on the inability to use more traditional award criteria. In a certain way, this motive could be linked to the first motive since selecting a partner in an early stage of a project (ECI) simply means that price and content assessments are not possible (to a full extent).

In contrast to many other projects of client E, there was no design on the table for case N6. Since the bridges of case N6 are monuments, the only thing that was clear for the project team is that those bridges should look the same in the end. The fact that there was no design yet made that the client could not ask for a price. "*What then remains is to award the contract on soft criteria, on collaboration: a tender process that focuses heavily on collaborative behaviour*" (Interviewee 7).

In case N1, the choice for an early contractor involvement delivery model was related to the advantage that this form of early and relatively intensive collaboration with market parties is that you go through the entire process together (Interviewee 1). The contractor is not surprised later on or confronted with a lack of information. This well-considered choice for ECI resulted in "*being less able to award on price ... and in this way attaching greater value to collaboration in the award process*" (Interviewee 1).

6.2.3 Motives related to starting and focussing on the collaborative relationship

For client B, collaborative criteria are also a way of improving collaboration throughout the procurement procedure. Both case N2 and N4 are labelled as one of the new projects by client B that embrace the recommendations of the 'Marktvisie'. This market vision resulted in a new umbrella of pilot projects by client B. The goal of this philosophy was to conduct an infrastructure project in a collaborative way with market parties, without the pre-defined more traditional boundaries and rules (Projectteam DOEN, 2017).

For case N4, this was done via a new procurement procedure that focussed more on the project and collaboration ahead. It differs per project under this umbrella how the procurement procedure is precisely shaped. The essence is similar: focussing on an improved and more sustainable collaborative relationship; 'maximum customer value and fair money for decent work' (Rijkswaterstaat, 2018). This identified motive (Motive III) could be linked to new project delivery methods like partnering but it must be seen in a broader context. The philosophy of case N4 is applicable for traditional projects as well (Interviewee 5). In case N4, Motive I and I* are not identified; it is not a very complex or dynamic project. This supports the belief that Motive III is related to but does not depends on the presence of Motive I and II.

It must be noted that this philosophy was already developed before case N4 started, but case N4 became the first project to bring the vision into practice. According to interviewee 5, the core of that new thinking was '*look at the tender phase without the current rules*'. That resulted in a team assessment and a written proposal about the cultural fit of the contractor to this new philosophy (Interviewee 5). The written proposal can be related to Motive IV as it aims to select potential partners that have the same vision and mindset as the client's project team/organisation.

Motive III is also identified in case N1 where the client was actively seeking for a different way of working together with the market. Interviewee 1 experienced the effects of mistrust and a negative attitude during previous projects. *"It is up to the client to 'do something about it', upfront in the start of the project in the tender process"* (Interviewee 1). Client A determined that this all starts with emphasizing and focussing on collaboration. A contract form should embrace this, but so should the criteria (Interviewee 1 and 2).

6.3 British motives for using collaborative criteria

Compared to the eight Dutch and Finnish cases , the two British cases of this study are quite different in size. Both U1 and U2 are a billion pounds project that consists of several contracts for the main works and, among others, demolishing and enabling work. Both Interviewee 8 and 9 worked on both case U1 and U2; assessor 4 was involved in both assessment centres. Table 5 shows a summary of the relevant case information.

Project & Client	**Collaborative criteria**	**Why collaborative criteria?**
Case U1, Client F	10%-30%*: behavioural assessment centre	*Need for collaboration, the project has an enormous public profile, very complex project challenges, and in the main works civil contracts much trust is handed over*
Case U2, Client G	20% -25%*: behavioural assessment centre	*Need for collaboration in the design phase, interface with key stakeholders, cultural differences may affect project outcome, relatively high uncertainty in the main works civil contracts*

Table 5 | British cases and their collaborative criteria

6.3.1 Motives related to intensive collaboration (in a joint team) during the design or realisation phase

Case U1 is one of the most unique infrastructure projects of this time. It is a complex and large project where a lot is at stake. Collaboration was seen crucial from the beginning (Interviewee 8). Case U1 could be seen as one of the most visible projects towards the British society of the last decades. In this way, the project has to deal with both the more external dynamics as complex challenges (Motive I) (Interviewee 9).

In both case U1 and Case 2, the civils work has been divided into several packages and tender processes. Eventually, client F (case U1) ended with four different contractors for all seven main works parts. All those four contractors participated in the assessment centre. Their assessment scores were included in the tender process of all part(s) they were bidding (Interviewee 8). The observations in case U1 and U2 can be clearly linked to Motive I. The level of complexity and uncertainty will result in an intensive collaboration.

6.3.2 Motives related to the fact that price/content criteria were possible to a lesser extent

All seven main works contracts of case U1 used an early contractor involvement (NEC) model with a high degree of collaboration for success. This ECI approach "*made it also hard to have a more traditional substantial weighting on price*", which is related to Motive II (Interviewee 8).

6.3.3 Motives related to starting and focussing on the collaborative relationship

Previous major projects in the United Kingdom (and other countries) were not failing because of budget problems, or there was no technical capacity to deliver the project. It is basically that there failing to agree on strategic discussions (Interviewee 9 referring to NAO, Major Projects Review). Interviewee 9 also refers to the observation that "*70% of strategic relationships fail/have problems due to lack of attention to cultural differences and management of the relationship*" (Cameron and Quinn, 1999).

In case U2, interviewee 8 was very quickly convinced to apply collaborative criteria. In a previous project, the contractor's competencies and expertise were scored by written submissions like the manager's cv. "*But these tell you nothing about their [future] behaviour*" (Interviewee 8). Some people got the best experience on paper, but they are 'very disturbing' and have to be replaced from a project. Besides, the fact that proper behaviour of ten years ago is not a guarantee for a future project (Interviewee 8).

These experiences within the construction sector made Client G apply collaborative criteria for case U2, especially those screening people face to face (team assessments and interviews). "*If you go purely on written submissions and historical documentation, you don't actually capture what you are getting*" (Interviewee 8.) That specific point convinced the contract manager to apply behavioural assessments with particular importance: "*if you do not meet this profile you are not going to work on my job*"(Interviewee 8).

6.3.4 Motives related to cultural alignment and shared values

For the main civil parts of case U2, it was, according to interviewee 9, very logical to apply team assessments. The need and level of collaboration would be very high. As part of case U2, client G also had some smaller contracts like the enabling works. These contractors would demolish several buildings before the main delivery phase. Although this is regular work - "*you do not interfere with them*" - client G also introduced collaborative criteria for this tenders (Interviewee 9). These enabling works had a high profile towards the public. The contractor would be the first to interact with the public physically – "*with big vehicles and noise*". Besides, a school had to be closed for a few days. *"It was key for us that the party shared the community values that are important for us*" (Interviewee 9). This reasoning made it logical to conduct assessments (20%-25%) for this part as well. This motive is new compared to the identified motives in other cases.

In the view of Interviewee 8, partner selection for such construction projects as U1 and U2 is almost like a marriage exercise. "*If you don't know the persons you will be working with, this will bring all kind of extra uncertainties"* (Interviewee 8). A contractor can really damage the reputation of a client organisation towards the public when it does not share your values. In this case, 'it is quite naïve' to focus on collaboration after award when the client organisation has way less leverage. This fact brought the conclusion for client F and G that aligning the client and contractors before the award is preferable. So concluding, for both case U1 and U2 it's not only a motive related to the contract model or level of collaboration that made these two clients apply collaborative criteria. It is also about sharing the same values (Motive IV).

6.4 Finnish motives for using collaborative criteria

Case F1 and F2 are both project alliances within the Finnish infrastructure domain. The procurement procedure of both projects is quite similar in terms of the evaluation criteria and different tender stages. It must be noted that client I was also involved in case F1 but not as the main client. As a result, client I used the experiences of case F1 for case F2. Interviewee 10 worked on case F1, which is already delivered, and Interviewee 11 works on case F2. The motives of both cases are shown in Table 6.

Project & Client	Collaborative criteria	Why collaborative criteria?
Case F1, Client H	25%: team assessment workshop (2d stage)	*Need for collaboration, establishing trust and focussing on best for project and the joint/integrated project organisation*
Case F2, Client I	20%: team assessment workshop (2d stage)	*Need for collaboration, complex and unique project for this client organisation, joint/integrated project organisation*

Table 6 | Finnish cases and their collaborative criteria

6.4.1 Motives related to intensive collaboration (in a joint team) during the design or realisation phase

For the delivery of case F1, client H used a relatively new procurement approach for that time: the alliance model via a competitive procedure with negotiation. Case F1 would be the second project alliance of client H. The project alliance in Finland is based on the idea that all parties form a joint, integrated project organisation (Lahdenperä, 2009). This Australian 'Victorian' alliance model is elaborated in Section 6.5. A consultant company helped to create the tender process and award criteria for the Finnish alliance model, not just for this project in particular (Interviewee 10).

The joint and integrated project organisation, during both design and delivery phase, is considered to be one of the main reasons for client H and I to apply an award criterion focussed on collaboration and leadership (Interviewee 10). Award criterion A4, '*the leadership ability of the alliance leaders and the project team, and alliance capability of the tenderer*', counted for 25 per cent in the last procurement stage of case F1 (Value for money report case F1, 2014). This criterion is also seen in case F2, but weighting 20 per cent (Value for money report case F2, 2016). For this criterion, the collaborative behaviour and alliance competencies of the tender's team were assessed via so-called (team) workshops (Interviewee 10).

As stated previously, client I was the second client in the alliance project of case F1. In this way, client I already had some positive experiences from using the alliance model in a high profile project. This was a strong incentive for client I using the same model for case F2 (Interviewee 11). Both market and client I agreed that the implementation of an entirely new tram system is such a risky and uncertain project that the alliance model was suitable for the implementation of the project's infrastructure.

Complexity and multiple interfaces towards other infrastructure projects, stakeholders and the public spaces were key drivers behind the decision for the alliance model and the award criteria, including the workshop (Interviewee 11). This relates to both Motive I and Motive I*.

6.4.2 Motives related to starting and focusing on the collaborative relationship

In case F1, the third motive of starting and focusing on the collaborative relationship was identified as well. Because of the partnership and joint team ahead, client H felt that it had to build the relationship already during the tender process. "*Trust is everything. We had to build trust*" (Interviewee 10). By focusing on collaboration in the tender process they were able to build trust from the first moment on.

Moreover, interviewee 10 reflects that the collaborative criteria are bringing this relationship development, which has to take place somewhere, upfront in the project. Both parties can have these discussions and relational development during the procurement or realisation phase. "*In this last alternative, the joint team has to start the project and do all team-development during normal work*" (Interviewee 10).

In line with other cases, Motive III is partly related to the other motives (as this team development is needed more in projects with intensive collaboration) but it can also be seen as a new motive. In case F1, Motive III is seen as an ambition that would also apply to more traditional projects of client H. This supports the interpretation that Motive III is a separate driver of applying collaborative criteria.

6.5 Discussing the identified motives

This chapter empirically found that collaborative criteria are closely related to the use of specific project delivery methods (Motive I and II). The desk study of this research (Section 1.2; Section 4.1) already suggested that soft parameters like collaborative criteria follow the development of partnering, alliancing and early contractor involvement in the construction sector (Kamminga, 2012; Eriksson, 2010). In the very first project alliances in the infrastructure domain, potential partners were selected on their potential fit with the project owner (Jefferies, Brewer, Rowlinson, Cheung and Satchell, 2006).

To understand the specific elements of those delivery methods that comply with the first and second motive, the different relevant delivery methods of several cases are discussed.

6.5.1 Motive I/II and early contractor involvement

The umbrella of early contractor involvement (ECI) includes serval delivery methods that strengthen the focus on relationship-based procurement. The approach of ECI started around the zero's and is adopted in countries like the United Kingdom, New Zealand, Australia and South-Africa (Koncarevic, 2013). The general belief of ECI is to involve the contractor at the early (design) stage of the project (Rahmani, 2016). This can bring benefits for the clients in terms of using the knowledge of the market more, being able to stimulate innovation and having more efficiently (value for money) (Mandell et al., 2013). Because these contractors are involved in an early stage, traditional tender price evaluation is less possible and common (Van Wijck, 2018). This fact matches Motive II of not being able to apply price or content focussed criteria.

The Dutch infrastructure sector mostly implements early contractor involvement via so-called 'Bouwteams' (Design teams). This model is seen in case N6 and intends to develop the design together in order to use the contractor's expertise about the different possibilities and construction process. It must be noted that besides the fact that the Design team-model stimulates a collaborative relationship and all parties are considered equal in the agreement, there is still observed a hierarchical relation between the client and the other team members in practice (Van Riggelen, 2019; Nielen, 2010). Although, it must be concluded that join design teams bring intensive collaboration and explain the presence of Motive I.

6.5.2 Motive I and project alliances

Alliance contracts and integrated project delivery have also some extent of early contractor (or designer) involvement to different degrees (Lahdenperä, 2012). As stated in Section 4.2, the project alliance model gained increased intention over the last decade. The model started in the construction and offshore industry in Australia, while public clients across Europa first saw difficulties in implementing the Australian model within the boundaries of the EU procurement directives (as described in Chapter 2).

Changing European laws and the possibility of the competitive dialogue method has facilitated the use of project alliances in Europe (Walker and Lloyd-Walker 2015). Case F1 would be one of the first alliance projects in Europe similar to the forms that are be used in Australia. Project alliances in the Netherlands (like case N1 and N5) use other procedures but are based on the same principles (Laan et al. 2011; Wondimu et al. 2016). All project alliances focus on stimulating an atmosphere of openness and trust, because parties share their risks and benefits (Love et al., 2010; Sakal, 2005; Walker et al., 2015). A project alliance always brings a joint team and alliance board since the responsibilities and risks are collectively carried. Their decisions are made unanimously (Chao-Duivis et al., 2007). In these projects, new organisations are formed and the (former) employer of team members have become less relevant (Elemans, 2016). This more intensive way of collaborating in the joint alliance project teams is considered one of the reasons (Motive I) that the alliance cases in this study have used collaborative criteria (F1, F2, N1, N5).

6.5.3 Motive III and literature on team development

The third motive of applying collaborative criteria addresses the future collaborative relationship and can be questioned from literature on bidding behaviour. The current goal of both parties during team assessments in the selection phase is not uniform (Lo et al, 2007; Kadefors, 2004). A tenderer benefits from doing something clever and therefore becoming first. After all, this is a competition phase.

Conducting a team assessment after the award of the contract is another possibility of building on the collaborative relationship and serving motive III (Assessor 2 and 3). Such assessment could be more simple and effective since there is only one objective: improving the interaction between the parties (instead of scoring the bidding parties within the context of procurement law). The competitive element of the tender phase is not present there, which could result in a more pleasant environment for the participants (Assessor 2). This alternative is further discussed in Section 9.1.

6.5.4 The identified motives within the broader context

As shown by the observations and discussion, it must be concluded that literature supports the observations from the several early contractor involvement and project alliance cases in this study. These two relational contracts enhance but also demand an intensive form of collaboration between the partners involved (Eriksson, 2010). This is part of the basic set-up of these approaches. Collaborative criteria follow the development of alliancing and early contractor involvement, and that seems no coincidence.

6.6 Concluding: the identified motives

This chapter focused on the client's reasons to apply collaborative criteria. Ten different cases were studied on the process and considerations behind the procurement strategy. This led to the project goals and characteristics that influenced the client's decisions towards the use of collaborative criteria.

SQ1 | What are the client's motives to apply collaborative criteria in the tender process?

This chapter empirically identified three different motives in the majority of the cases. These motives are considered to be the most critical drivers to use collaborative criteria in the tender process. The three motives are (de)find via a thematic analysis and do not exclude each other:

i. Applying collaborative criteria because the project and (relational) contract bring intensive collaboration (a joint team) in the design or realisation phase - mentioned in 9 out of 10 cases;
ii. Applying collaborative criteria because other criteria focusing on price and content of the design were less possible due to the early contract involvement - mentioned in 6 out of 10 cases;
iii. Applying collaborative criteria to lay the foundation of the future collaboration and emphasize on the importance of collaboration/soft skills - mentioned in 5 out of 10 cases;

Because interviewees could not choose from a predefined list of motives, Table 2 (paragraph 6.1) is not fully comprehensive. The exploratory nature of this study resulted in this first impression. Further research could validate and deepen the different views on collaborative criteria by a Q-methodology.

7. Instruments to score collaborative criteria

This seventh chapter elaborates on the different formats that are used to score collaborative criteria. In this way, the following six paragraphs will serve the second sub-question*: what instruments and formats are used to assess the contractor's collaborative behaviour?* One of these instruments, the team assessment, is discussed in more detail since this is the main instrument to assess collaborative behaviour. Other associated and supportive instruments are described to identify similarities and differences.

The first paragraph (7.1) will provide the wide pallet of different instruments that is identified in across the ten cases. The subsequent paragraphs discuss the different formats of the team assessment instrument by elaborating on (7.2) the scoring, (7.3) the content, (7.4) the participants of team assessments, and 7.5) the assessors and observes being involved. The conclusion (7.6) will answer the third sub-question on two different levels: the wide pallet instruments and the variety of team assessment formats.

7.1 The identified Instruments

As a result of the selection criteria of the case study, all ten cases used some form of a team assessment where the collaborative behaviour of the tenderer was assessed. However, the scope of the second sub-question is to identify all kind of instruments that are used to (support the) assessment of the tenderer's collaborative behaviour of cultural fit. While the previous chapter was all about the reasons to apply collaborative criteria, this paragraph focusses on the instruments that are used to score such criterion.

Understanding the differences between terms like 'instrument' and 'assessment' is essential for interpreting the observations of this chapter. On the left side of Figure 10, the 'collaborative criterion' is displayed and can be seen as the 'umbrella term'. The assessment (centre), which is used to provide scoring for the collaborative criterion, consists of one and sometimes more instruments. These instruments can be an interview, team assessment or written submission, and make use of several exercises. Some cases in the the United Kingdom used several instruments to score one award criterion. In that case, both instruments I and instrument II are part of the so-called assessment (centre).

As displayed in the right of Figure 10, the term 'sub-criteria' refers to the list of competencies that is used to score participants. These competencies are discussed in Section 7.2.

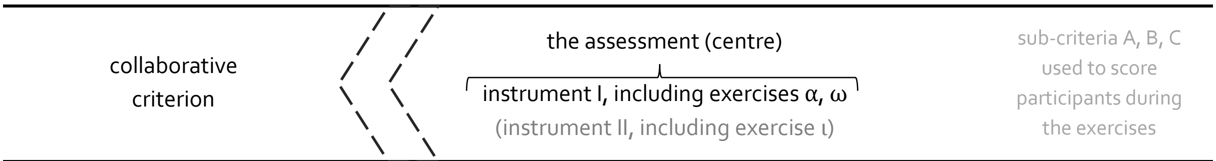


Figure 10 | Terms and concepts explained

Portilla (2020) conducted a literature review on team assessment and relational partnership and identified three categories of instruments that can explain the capability of a tenderer team to collaborate and perform: 1. *team dynamics instruments (*about the behavioural relationships in a team*),* 2. *organisational and cultural alignment instruments* (about the alignment between organisations*),* and 3. *individual matching instruments* (the composition of the team) (Portilla, 2020). These three categories are used to categorise the identified instruments across the ten cases. It must be noted that the second category is adjusted to instruments that focus on alignment of values and visions.

7.1.1 Team dynamics instruments

The focus of this study and most extensive instrument for the evaluation of the contractor's (collaborative) behaviour is a team assessment where team members from the tenderer and client organisation are present (Assessor 1 and 4; Interviewee 9). These assessments aim to assess the behaviour of a team by using project-related exercises to more general (non-business related) ones.

The citation below gives an impression of how an average team assessment works out in practice.

> Interviewee 6 (Case N5) "*We carried out an all-day assessment with each tenderer. We have chosen for a non-project related game in the morning and three cases in the afternoon. In these case, we simulated a future project team meeting about certain events that could be expected, for example the minister intervenes or there is a delay or a major contract change. The question is 'can you now deal with the dynamics offered by this project?'. During the cases you could clearly see whether the atmosphere of that meeting was according to the alliance idea we had in mind.*"

The assessment itself, the scoring, is done by independent experts (assessors) who analyse the behaviour based on a set of competencies required in a team to collaborate and perform. The components and content used for these team assessments vary from project-related workshops and scenario's to several very basic general exercises. This variety is shown in Table 7. A definition of these contents follows from the interview reports and is presented in the table description.

Instrument	N1	N2	N3	N4	N5	N6	U1	U2	F1	F2
Team assessment: tenderer+client	X	X	X	X	X		X	X	X	X
- Business related exercises	X	X	X	X	X		X	X		
- General exercises	X	X		X	X		X	X		
- Workshop exercises									X	X
Team assessment: tenderer						X				
- Business related exercises						X				
- General exercises						X				

<u>Business related exercises (cases)</u>: developed simulations of how the bidders' organisation demonstrates when problems occur in the project setting. There are two variants: cases dealing with challenges of the project itself and cases focussing on a similar project. A case should be written in such a way that it actually tests behaviour and not knowledge (Assessor 1 and 4).
<u>General exercises</u>: exercises and games that have nothing to do with the business/project but are very related to behaviour. Examples: people had to design and built a chair made out of blown-up balloons or to write down a story (Assessor 1, 3, 4).
<u>Workshop exercises (cases)</u>: sessions about real project issues like the risk portfolio. These sessions are also used to observe the tenderer's ability to collaborate and understanding of the required behaviours, but also facilitate the possibility to address questions on real project aspects which may be relevant for the offer in the procurement phase (Interviewee 10; Assessor 5).

Table 7 | Team dynamic instruments

7.1.2 Cultural & vision alignment instruments

The second category of instruments focusses on the alignment of organisations and their visions. This can be done by a written submission about collaboration and norms & values. Table 8 shows all instruments that have been identified in one of the cases. A project management plan is not considered to be a cultural and vision alignment instrument since its focussing on more practical and procedural aspects.

Instrument	N1	N2	N3	N4	N5	N6	U1	U2	F1	F2
Written answers to certain cases							X			
Tenderer's vision on collaboration	X			(X)		X				
Culture alignment survey							X	X		
Panel/team interview							X	X		
Individual interview							X			

<u>Written answers to questions about certain collaboration challenges/cases:</u> several prescribed cases which are answered on paper
<u>Vision document</u>: a document by bidding organisations about their view on the project challenges and client-contractor relationship ahead (including organisation, values, ambition, how to deal with conflicts) (Interviewee 7; Assessor 1)
<u>Culture alignment survey</u>: a questionnaire showing a graph with degree of alignment. A standard industry tool is used to find out the ideal culture for the client. In some areas you want alignment, while in other areas you don't want that since it brings value (Assessor 4)
<u>Interview</u>: a team or individual interviews with board members, delivery people or key team members about certain collaboration challenges, the corporate culture and the way safety or other relevant criteria are handled (Assessor 4)

Table 8 | Cultural & vision alignment instruments

The necessity for assessing organisational cultural congruence emerges because a new project organisation is a product of joining two (o more) independent organisations. The fit of their cultures and values has been reported to be an element that explains their performance (Cadden et al, 2013).

7.1.3 Individual matching instruments

The last category of instruments aims at giving insights into the composition of individuals in the team and how this composition influences its performance. The team dynamics (as assessed in team assessments) is influenced, among other factors, by individuals. Nevertheless, the team assessment instrument does not include the analysis on an individual level. These individual matching instruments do. Table 9 shows the identified individual matching instruments across the ten cases.

Instrument	N1	N2	N3	N4	N5	N6	U1	U2	F1	F2
Personality traits				X			X	X		
Behavioural biographies							X			

Personality traits: is a insights discovery, big fice, mbti etc, emotional intelligence: EQ-i-test
Behavioural biography: is a structured way of collecting previous experiences of individuals so it becomes comparable. It decreases control of the interviewee like delivering your cv (Assessor 4).

Table 9 | Individual matching instruments

Portilla (2020) concluded that this last category of instruments should be mainly applied to increase the prediction value and/or explain and underpin the behaviour observed during the team assessment. This conclusion supports the main focus of this study related to the team dynamics instrument.

7.1.4 Weight and stage of the instrument in the tender process

Each instrument of Section 7.1.1 and 7.1.2 could be used as evaluation criterion or as pre-selection. The purpose of selection criteria is to limit the number of candidates to a predetermined number (usually 5, sometimes 3) (Pianoo, 2020). Two of the Dutch cases have used the team assessment as a selection criterion.

Case N1 and N5, both project alliances, allocated a high weight of respectively 50 and 40 per cent to the team assessment. Client A (Case N1) discussed different percentages. "*It should be able to make a difference on this point and that conclusion resulted in a 50% weighting of the team-assessment*" (Interviewee 1). Within client C, a sensitivity analysis was done for the criteria (a matrix with all possible results). "*Do you prefer someone with a bad plan and a brilliant team assessment or vice versa?*". A percentage of 35 per cent came out of that discussion as a compromise (Interviewee 4).

This kind of sensitivity analysis is also done in the United Kingdom. Interviewee 8 notes that it's not just the weight of a criteria; it's the sensitivity of those 20 per cent to that other 80 per cent "*Because people often mistake, there are many ways of getting marks"* (Interviewee 8). For all subparts of case U2, the ultimate weight of different award criteria and the degree of importance of collaboration was a decision on board level. It must be noted that a collaborative criterion with 10 per cent still has a value. Even if it will not make a difference in the decision making, it still acts as an instrument to address the importance of collaboration (Interviewee 9). This signal function is also addressed in section 8.5.

A majority of the interviewees agreed on the position of a team assessment in the tender phase. When the tender phase lasts about six months with various dialogue discussions, irritations could arise easily (Interviewee 6). In other words, positive or negative experiences have arisen and can influence the assessment. Doing a team assessment at the end of the tender phase is less pure compared to an assessment at the very beginning (Assessor 1; Interviewee 6). In most cases, client teams have not reviewed any of the written submissions for the final award criteria when entering the assessment environment.

Some clients in the United Kingdom, not any involved in this study, applied a behavioural assessment within the competitive dialogue meetings (Interviewee 8; Assessor 4). According to several interviewees, this is no actual team assessment completely focussed on collaborative behaviour because this is quite impossible in terms of a level playing field due to the changing content (Assessor 1; Interviewee 4 and 9).

7.2 The scoring of team assessments

After choosing for a team assessment, the question arises 'what kind of behaviour do you want to see in the assessment?' In other words, 'what are looking for and scoring as an assessor?'. It must be noted that this study is not conducted within the social science domain. However, the context of the sub-criteria and scoring applied is relevant for understanding the variety of team assessment formats.

7.2.1 The client's perspective on the team assessment

Before the variety of specific competencies is discussed, clients were asked about their higher perspective on the team assessment scoring: '*what kind of behaviour are you looking for?*' The client's team of case N1 was looking for people that were able to delve into each other's interests, to put aside own views, and to share responsibility (Interviewee 2). Three other views are summed up below.

> Interviewee 1: "*How do you act if things are really going to get unpleasant?', this calls for a certain sense of putting pressure on a contractor and looking at how the team then performs/ behaves*".

> Interviewee 3: "*Sometimes you have to look to the common project goal from conflicting interests. You want to have a team assessment and case simulation where this emerges and grinds*".

> Interviewee 6: "*What do you do when something doesn't work out?*"

These views are the starting point in the discussions with the assessor about designing the assessment.

7.2.2 Scoring of the bidding party

The set of competencies (sub-criteria) that are used to score the tenderers vary depending on the kind of project and the consulting firm doing the assessment. For instance, the assessors involved in case N1, N2 and N4 used 1) The ability to set clear goals, 2) The ability to take mutual responsibility. 3) Open communication, 4) Mutual respect, 5) Flexibility in cooperation, and 6) The ability to take the initiative. This set of competencies is based on 'het wiel van Vroemen' (Vroemen, 1995) and displayed in the left case-column of Table 10. To give an impression of the differences between the sub-criteria that have been used across the different cases and countries, Table 10 shows three cases and their sub-criteria.

Sub-criteria/competencies	Case N4	Case U1	Case F1
The ability to set clear goals	X		
The ability to take mutual responsibility	X	X	
Open communication, reflection and feedback	X		X
Mutual respect and empathy	X		
Flexibility in cooperation	X		
The ability to take initiative and leadership	X	X	
Organisational skills			X
Decision-making and problem-solving skills (solution management)		X	X
The ability to build and reinforce mutual trust		X	X
Commitment/ability to operate in accordance with principles/ targets of the alliance			X
Relationship building and stakeholder management		X	

Table 10 | Sub-criteria/competencies (Based on interviewee 5, 9 and 10)

To score each competence, most assessors use an ordinal scale. One option is to use a five-point ordinal scale and give a score depending on the performance of the tenderers in terms of each competence (mark 1: poor, mark 2: insufficient, mark 3: sufficient, mark 4: good, and mark 5: excellent) (Portilla, 2020).

Some assessment companies even use tick-boxes (while observing certain behaviour), which gives them a number (Assessor 4). This is believed to be a poor mechanism because you should have a qualitative approach to train the assessors and allow them to make the judgment in a professional way. The four stages of assessing (Observation/Recording/Classification/Evaluation) should be done separately (Assessor 1 and 4). Focussing on the sub-criteria used, for assessor 5 the literature on project alliances was leading because their workshops aim to test the tenderer within the context of an alliance (see fourth row in Table 10).

According to assessor 5 (Finland), the higher the understanding of the essential structural features of project alliances (based on Ross, 2003), the higher the project alliance competence is. Therefore the assessor wants to see behaviour that represents how these five sub-criteria are understood. An example:

"best-for-project" is seen in the behaviour of the tenderer of how the team can prepare the decisions together with the client before they go to the project alliance board (Assessor 5). This is quite different compared to the sub-criteria used for a Dutch project alliance. It must be concluded that the sub-criteria (and focus of the assessment) varies greatly per case, country and assessor.

7.3 The content of team assessments

Most cases have applied both business (project) related cases as well as (a lower number of) general exercises. Assessors and project and contract managers across the Netherlands and United Kingdom do agree on the value of this mix (including non-business related exercises).

Non-business related cases could be more demonstrating the behaviour of the tenderers, or it can at least shake people a little bit according to interviewee 9 and assessor 4. *"In my view it should be a proportional balance between business related cases and general exercises. If you consistently talking about very serious matters at some point you will lose your audience"* (Interviewee 9). Assessor 1, 3 and 4 agree that one of the problems in business/project related cases is writing them up in such a way that they test behaviour and not knowledge. So project-related designing can be more hard for the assessors involved (Assessor 4).

As set out before, the workshop as seen in Finnish cases is significantly different from the Dutch and British approach regarding the team assessment content. Client H and I did not simulate any random problem during the workshop rounds, because they focus rather on the real product/offer during this process of workshops (Interviewee 10). Assessors and client choose several project elements like risk evaluation, quality of groundwater and then ask the question to the tenderers: '*are there some kind of risks we have noticed yet? Or have we seen some risks too high or too low?*'. Within several sub-groups the bidding parties have to plan and organise the workshop. An example of assessor 5 is a small workshop (not related to any of the cases of this study) where the general design was developed in such way that costs decreased for more than a million euros (Assessor 5). This Finnish approach is significantly different from the more scenario and gaming setting in which the British and Dutch assessors are moving.

However, it must be noted that there are elements of the Finnish workshops which are more similar to the other ones. In an example of assessor 5 the tenderer had three hours to execute surprise tasks they did not know before. The output of the task was to make a proposal for the executive board of the future alliance board. This case also shows some similarities to the business-related cases (with surprising elements) in the United Kingdom and the Netherlands. Nevertheless, the origin stays different. The Finnish approach is more content focussed. It's more about the real value the team can already generate together from the starting point (Assessor 5). This is preferred above analysing the tenderers teams "potential" by different tools (Assessor 5; Interviewee 10). "*From the client perspective, the team that can generate great value from the starting point is usually better than the team that could be really good in the future"* (Assessor 5). This different belief supports the motive to study more formats than just the more similar Dutch cases.

7.4.1 The assessment environment

All assessors expressed that it is crucial to create an ambience where the tenderer's team can show authentic behaviour (Assessor 1, 2, 3, 4 and 5). However, when you dive into the description of the setting of those team assessments, there are quite some significant and interesting differences in the Dutch and British cases. Because the Finnish workshops include no games with an enlarged conflict of interest, the setting and ambience of these workshop are automatically less tense.

In the Netherlands, two views related to the assessment environment were observed. First, the view of interviewee 7 (case N6) and their assessor (assessor 3). While developing the assessment for Case N6, two variants came to the table 1) 'turn it into a military assault course' and 2) 'turn it into a more friendly case where you can also test those skills'. The client preferred this last one since it would be more the starting point of the desired future environment (Interviewee 7). Assessor 3 indicated that from the perspective of behavioural psychologists it doesn't matter which one was chosen (Assessor 3). Second, you have Dutch (and British) assessors who make more use of interventions and a specific tense to bring the authentic behaviour up-front. According to this perspective, a case with a conflict of interests is needed to score the competencies and show how tenderer's would behave when thing get challenging (Interviewee 1).

'Pleasant environment.'	'Intervention and tension'

One of the personal favourites of interviewee 8 is a decapitation where the assessor is removing the project manager from the tenderer's team like he has an accident. "*That's really interesting because you get a different dynamics*" (Interviewee 8). Besides, some British assessment providers also makes use of actors to come and play specific roles, for example, a resident on-site complaining about the noise (Interviewee 8).

7.4.2 Duration of the assessment

Case U1 and U2 used different forms of comprehensiveness of their team assessment. The team of Client F designed several variants of the behavioural assessment for case U1 (Interviewee 9). The case included multiple behavioural assessments, not only for the main works civil but also for design tasks. These different variants were made because some contracts needed an assessment, but these should be in balance with the value of the tendered part and the tender costs (Interviewee 9). Client F decided to assess the designers, but with a light variant of the behavioural assessment. According to interviewee 9, this variant can be seen as the minimum of behavioural assessments and is half day per tenderer.

The Dutch case N6 and some assessments of case U1 used half a day per tenderer. All other cases used one full-day program (minimum of five hours). Improving the scoring, getting used to the setting and ensuring all participants show authentic behaviour are arguments to do so (Assessor 1; Interviewee 1 and 6)

half day	full day	two or more days
N6, (U1)	N1, N2, N3, N4, N5, U1, U2	F1, F2

7.4 The participants of team assessments

The content and extensiveness of the assessment also affect the (amount of) participants being involved, and vice versa. This chapter identifies what kind of considerations are relevant for the people you'll invite for and assess during a team assessment. Since British assessment centres also include interviews with key people, this paragraph discusses these participants (next to the team assessment) as well.

7.5.1 The Dutch approach

In Dutch team assessments (with the tenderer and client), the group that's being invited is in almost any case the core team consisting of several managers from both sides. This group of key figures will, after being awarded the contract, work together and "*make or break the project*" (Assessor 1). Rijkswaterstaat (the Dutch Department of Waterways and Public Works) had specified their five roles as 'IPM' [Integral Project Management], which includes a project manager, a manager project control, an engagement manager, a technical manager and a contract manager (Rijkswaterstaat, 2019b). This group of five people is seen on the right and is reflected in the dark grey triangle of Figure 12. It is the top level of the day to day project organisation.

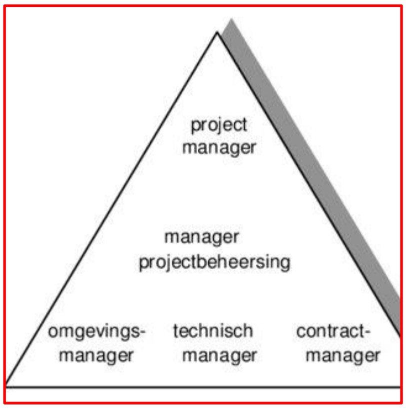


Figure 11 | IPM model

When the client's side is represented by those five managers, the contractor is asked to represent the same five disciplines. This could be five people, but also four or six (Interviewee 3). The tender guideline of case N5 stated *"A team covering the same disciplines is expected from the tenderers... Furthermore, when putting together a team, it can be useful to know the profile of the core team members of the client. This profile is included in Appendix H of this Tender Guidelines"* (Tender guideline case N5).

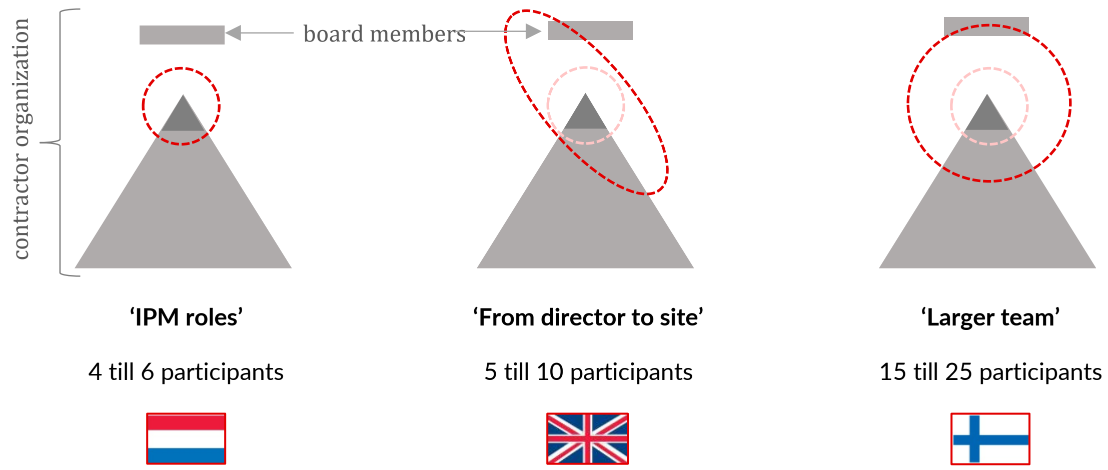


Figure 12| Participants from tenderer organisation in the assessment (centre)

7.5.2 The British approach

In the United Kingdom, the selection of participants seems to be more depending on the specific project. "*Who you involve is driven by the requirements of the contract and where the critical interfaces will be. Every judgement of who is invited for a behavioural assessment should be in the context of the critical factors of delivery*" (Assessor 4). This considerations leads to the key figures most of the time, but not automatically.

There are two more observations that differ from the selection of those participants for Dutch assessments. First, some projects split their participants. The assessment with the designing people gets a score and the assessment with the delivery people. For the scoring on the collaborative criterion, the two scores are summed up. According to assessor 4 this depends on what the client is willing to do. "*You normally look to all key people being involved during the whole project*" (Assessor 4).

As illustrated in Figure 12, the British cases also included key directors in their assessment centre. Case U1 did this via CEO interviews (in this study identified as 'cultural alignment instrument') and case U2 included some key directors in the assessment itself (Interviewee 8 and 9). Besides, these clients also considered involving delivery people to assess the cultural fit between organisations. It must be noted that this is not the about the team assessment participants but also the other instruments identified in Section 7.1.

The participants of case U1 and U2 included managers from the project team. In some other British projects, the participants come from all over the organisation "*... a (finance) director, but also specialists or even delivery people – and different project domains, so engineering, collaboration, project management*" (Assessor 4). This follows the belief that assessor's focus on what is critical for delivery. It all depends on the risk assessment upfront before it is decided whom to invite for those assessments. *"We ask the question: who are the people that affect these risk most? That drives us to determine the people involved*" (Assessor 4).

7.5.3 Involving board members

The client's team of case U1 was very interested in the head of the company of the bidding parties because they believed that the board reflects the culture of a company. When the board level person shares the values of the client, it's likely he will pass these values to his project team(s) (Interviewee 9). " *... So it's very critical in my opinion to meet CEO's or key directors*" (Interviewee 9). This is done in addition to the team assessments with the project and design managers that are invited anyway. These interviews with key directors for case U1 were based on the lessons of case U2 (Interviewee 8).

According to interviewee 9, team assessments and interviews are particularly relevant for joint ventures. If there are three entities that are forming a consortium there are three CEO's involved. "*It could occur that they have never talked to each other before this tender. The behavioural assessment centre pushes them down to sit down and decide how to present themselves*" (Interviewee 9). In this way, the assessment encourages not only interaction between the members of the project team but also on the board level. "*When these three [CEO] guys met just before the interview, we will most of the time see that*" (Interviewee 9).

Assessor 4 underpins these observations by interviewee 9 and 10 and is convinced of the fact there are some good reasons to think this is appropriate, for example when corporate policies may affect delivery. "*We have flown them from Japan and the states for interviews*" (Assessor 4). Assessor 4 adds to this that you have to be sure to have appropriate and specific questions for these people. It's considered a waste of time to get these board members in the assessment room or ask them about the corporate approach of a business unit of their organisation when it has nothing to do with the project (Assessor 4; Interviewee 9).

7.5.4 The Finnish approach

In Finland, the number of people being involved in the workshops (case F1 and F2) is relatively high. This is due to the fact that the purpose of these workshops is not only assessing the contractor's collaborative behaviour and alliance competencies but also work on real issues of these large projects (Interviewee 11). In other workshops (for other Finnish projects) just a few people have been invited (Assessor 5).

During the workshops for case F1, the client team was present with nearby ten to fifteen people (Interviewee 10). At this time, this was all new for everyone within the client organisation. From the client-side the project managers, members of the alliance board, experts from the tunnels and roads and some other technical experts for lightning and tunnel technical systems were present. Similar roles were present from the tenderer's organisations (Interviewee 10). The consortiums were with nearly twenty people present and all of them should work on the project in the future as well. During the tendering for F1, there were more people present than nowadays. "*Now we realize that some people are not needed for these workshops. It was part of the learning process*" (Interviewee 10).

The more recent workshops for case F2 included around 25/30 people from the client and about 25/30 from the market parties (Interviewee 11). The reason for this, as interviewee 11 explains, is that the tramway project alliance is so large that it cannot be implemented by a five-person client project staff.

In order to ensure the commitment of each department of the client's organisation to the project, each of these units has its own representative in the 'Alliance Steering Group' and were all invited. Interviewee 11 underlines that it is completely understandable that this kind of commitment is not possible or desirable for such projects done by larger clients. Although, the tramway project is for client I in particular a very high profile project which has a major impact on the development of the whole city (Interviewee 11).

It can be concluded that the Finnish approach of today still differs from the Dutch IPM model. The workshops for project alliances include people on several management levels. The members of such 'Alliance Steering Group' are not assessed in the Dutch cases N2 or N5 (also project alliances). Figure 13 below shows the number of participants from the client and tenderer's side in each case.

	N1	N2	N3	N4	N5	N6	U1	U2	F1	F2
Client's team	5	5	5	5	6	-	7/10	6/10	~15	~25
Tenderers team	4/5	5/7	5	4/5	6	5	7/10	6/10	~20	~25

Figure 13 | Selected cases with its number of participants in the team assessment

7.5.5 Assessing the client's team

The assessors can hand over some points of attention to the client's team for their future behaviour, but most of the time there is no separate report written down (Assessor 1 and 2). This is also seen in the United Kingdom. However, some clients, among client F, explicitly chose to evaluate the client team. To do this in the same way and with the same robustness as the tenderer's team, a second assessors team was deployed (Assessor 4). "*I don't suppose it matters if you do that independently or within the team-assessment with the tenderer, but I prefer the latest one – and you people have to know that in advance*" (Interviewee 8).

In the end, the client's team got a report with their competence level and risks (a combination of individual reflection and for the team as a whole). When the client and contractor started the job after awarding, both reports were shared as part of the joint team development (Assessor 4). This observation in case U1 is one of the best examples that the British use of team assessments is not just about partner selection, because that would just require assessing the tenderer. It's even more about the collaborative relationship.

7.5 The assessor of team assessments

In all cases, external assessors (and observers) were involved in preparing, assessing and scoring the capacity of the bidders to collaborate within the proposed delivery method. The client picks an assessment provider ('assessment firm' or 'occupational psychologist company') in the beginning of developing the procurement strategy. Assessors are already helpful in an early stage of the procurement phase since the collaborative criteria need to be defined with sub-criteria in the procurement/selection guideline.

7.3.1 The number assessors being involved

The number of assessors involved in the assessment itself depends on the number of people that are being assessed. In British behavioural assessments, the number assessors is significantly more compared to Dutch and Finnish cases. In case U1, the client's team needed a very robust process for all their billion pounds main works contracts so they simply expended the number assessors (Interviewee 8). The client involved seven people that were already working within the client organisation together with seven behavioural psychologists. Those group of fourteen people act as assessors (Interviewee 8). Most British assessments include four until six assessors, which are all from occupational psychologist companies (Assessor 4).

Some firms even let their assessors submit a cyclomatic profile at the beginning of their career to test what their personality profile is. "*We don't change their personality but find out mitigations that have to be made in the team composition of assessors*" (Assessor 4). It must be noted that none of the firms in the United Kingdom is very massive, so most of the time they call in associates to increase the number of assessors in large projects (Interviewee 9). The centres of Case U1 were that big that no firm would pretend they could do all the assessments on their own (Interviewee 8).

The assessment provider can also be a university. Case F1 chose to involve some experts from the nearby university. Interviewee 10 argued that these people already had some relevant material on the shelve, and the reputation of most universities is excellent. "*We told all consortiums that these specialists were coming, and this was very easy for them to accept*" (Interviewee 10). Multiple interviewees indicate that they have considered a few companies or independent assessors. Especially their experience within the construction sector is considered to be valuable. When client G was one of the first clients in the United Kingdom to apply assessments, they had to explain the procurement context and the risk of being challenged. Nowadays, assessors are familiar with the context of this industry, which is also important for developing the case together. "*The psychologists in this country are getting better and better*" (Interviewee 8).

Based on previous assessments and studies, assessor 4 guarantees a prediction of their results. "*If the client follows all we asked them to do – and if they go to all the training – our prediction of the future collaborative behaviour of that team its within plus or minus of our conclusion from the assessment*" (Assessor 4). No other assessor involved in this study can or has made such a statement.

7.3.2 Guidelines for assessors and assessments

When discussing level playing field, two out of the five assessors involved in this study mention the relevance of specific guidelines for conducting proper assessments. In the United Kingdom, where the first assessment for a public client's partner selection already took place around the zeros, there are several guidance's in place: the BS11000 Collaborative Relationships Standard and ISO 44001 and 44002 (Interviewee 9). ISO 44001 (2017) specifies requirements for the sufficient identification, development and management of collaborative business relationships and ISO 44002 (2019) gives guidelines on the implementation of ISO 44001 (ISO, 2020). So the ISO 44001 is about applying behavioural assessments, the partner selection part itself, and ISO 44002 tells you about doing it (Assessor 4).

Recent years, the British Psychologistic Society also published a guidance manual on how to conduct these assessment. The document provides details on how many times you should test a given sub-criteria before you can make a judgement (Assessor 4). The Netherlands has such organisation as well, the Netherlands Institute of Psychologists – also having a 'Professional code for psychologists' [NIP beroepscode]. The reviews done by assessor 1 are confidential reports and are carried out under this code (Assessor 1). Assessors who do not refer to such guidelines in the interviewees still use their own procedure. For example, the assessor of case N3 delivered the final reviews of the team assessment to the notary. The client's team was not familiar with these reports or even the scoring. The assessment committee provided the scoring on the other criteria to the same notary. In this way, it was an independent party who summed up and revealed the total results (Interviewee 4). According to interviewee 9, these guidelines are essential since a poorly conducted team assessments reflects on the whole domain. It creates unnecessary bad experiences which makes it harder to convince those clients later on of this value (Interviewee 8 and 9).

7.3.3 Involvement of independent observers

Both cases in the United Kingdom and Finland (U1, U2, F1, F2) worked with independent observers during the team-assessments and workshops. This phenomenon is not seen in any of the Dutch cases.

In all four cases, the observer - or several observers within a compliance team - has the role in giving feedback to the client's team in a separate meeting after the assessment (Interviewee 9 and 10). "*This can be very simple, for example, you have a bathroom break on that day so you have to repeat this every single assessment*" (Interviewee 9). Client F had independent observers to make sure it was all done as being declared. They happened to be from a procurement discipline, but they were only looking to the process itself and writing a report at the end of every day like 'yes, the process was observed'.

The observers in Finland have no connection to any tenderer or the client. The observer of case F1 had, for example, a long career and worked for another organisation. Client H told all consortiums that if anything is bothering you during the assessment, you can talk to this person (Interviewee 10). It was also the way that they ensured that they [the client's team] were behaving correctly and there were no benefits for one party in particular (Interviewee 10). During the first workshop, that client F organised even a few lawyers were invited to observe and ensure a level playing field. After each workshop, the client team had a feedback session about 'did everything went ok or was something bothering you and uncomfortable?' (Interviewee 10). This example supports the conclusion that clients and the assessors, observers and lawyers that are involved take the responsibility and principle of a level playing field very seriously.

7.3.4 Ensure a level playing field

The previous paragraph already touched upon the importance of creating a level playing field. Involving observers and making sure you have the right (number of) assessors can be seen as the first step of creating equal opportunities and environments for every tenderer.

The second step that is considered to be important by both assessors and the client's team member is the right preparation prior to the first real team assessment. "*... That preparation is really valuable"* (Interviewee 5)*; "*W*e had to train our self as well. With the help of a consultancy company, we created a guideline 'how to behave as a client during workshops*?" (Interviewee 10). The interviewees of each case see their preparation as a crucial step to ensure a level playing field (Section 8.2).

Interviewee 9 underlined the importance of focusing on this level playing field with two concrete examples. During the team assessments for case U2, the lunch could not be provided at 12 o clock, as planned, but only at 1 o'clock. It was the first assessment day, and according to the principles of a level playing field (creating the same setting and the same conditions), interviewee 9 changed the timetables of all coming behavioural assessments. In this way, client F is really focussed on creating the absolute same environment (Interviewee 9). What's also seen in the Netherlands and Finland is that those one-day assessments are spread over a couple of weeks. "*You need to understand that the assessors can't do each day for five working days on"* (Interviewee 9). This same argument is valid for the client team. Within case U1, it was even decided that the Monday mornings and Fridays were avoided for this reason.

7.6 Concluding: the instruments and formats used

In this chapter, the different instruments used for scoring collaborative criteria in the tender process were identified. In this way, the chapter answered the following sub-question:

SQ 2 | what instruments and formats are used to assess the contractor's collaborative behaviour?

The first part of this sub-question is answered by giving an overview of all identified instruments, which can be found in Table 11. It must be noted that the left column of team dynamics instrument is the dominant set of instruments used since it was one of the selection criteria. Next to a team assessment (in nine out of ten cases with the tenderer & client), several cases also made use of cultural alignment instruments; like written documents and (panel) interviews, and individual matching instruments like an EQ-i test.

Team dynamics instruments	Cultural and vision alignment instruments	Individual profile matching instruments
Team assessment (tenderer & client) a) Business related exercises b) General exercises c) Workshops	Written answers to questions about certain collaboration challenges/cases	Personality traits: Insights Discovery, Big Fice, MBTI etc Emotional Intelligence: EQ-i-test
	Tenderer's vision on collaboration	Behavioural biographies
Team assessment (only tenderer) a) Business related exercises b) General exercises	Culture alignment survey	
	Panel interview	
	Individual interview	

Table 11 | All instruments relating to collaborative criteria as identified across the ten cases

The second part of the sub-questions focussed on the variety of team assessment formats. All identified team assessment instruments involved independent assessors to score the bidding parties on a predefined list of (soft) competencies. Moreover, guidelines and a robust procedure were used to ensure a level playing field and conduct the assessment according to the procurement regulations. Besides, all British and Finnish clients involved observers to 'protect' the process and ensure a level playing field by observing the client's team behaviour during all different assessment exercises. In most of the cases, the assessors have a more extensive role than just observing behaviour during the assessment. Assessors were already involved during the stage of designing the market strategy and its tender criteria.

Across the ten cases, the number of assessment participants varied from 4 until 25 individuals per side. Besides, it was found that the variety of formats relates to the different definitions of the tender sub-criteria: team assessments that focus on the cultural fit between client and tenderer bring different competencies on the scoring card compared to focussing on the tenderer's alliance capability. Most cases have applied both business (project) related cases as well as (a lower number of) non-busines related exercises.

To conclude, the team assessment has some corresponding cornerstones in the Netherlands, the United Kingdom and Finland but the specific application (participants, content, duration, scoring) differs per country, project and last but not least, the assessment provider (assessors) involved.

8. Experiences of conducting team assessments

A valuable part of this study consists of describing first-hand experiences of conducting team assessments. Previous studies about team assessments, among others Portilla (2020) and Dewberry et al (2018), did not primarily focus on those experiences or just mentioned a few projects. This chapter answers the third sub-question by using the experiences of sixteen interviewees across three countries: *what are the client's experiences in conducting team assessments*? The experiences that are observed multiple times (two or more) are discussed in one of the five paragraphs of this chapter.

The client's (and assessor's) experiences are discussed chronology: first (8.1) the experiences regarding the communication with the market, subsequently (8.2) the briefing sessions beforehand, (8.3) the experiences from the team assessment and soring itself, (8.4) the client-contractor collaboration subsequently and lastly (8.5) a reflection and discussion on the value of the instrument. The observations as set out in this chapter have two purposes: 1) answering the third sub-question, and in this way showing the sentiment around team assessments across client organisations, and 2) generating input for Section 9.1 and 10.3.

It must be noticed that the experiences of all interviewees, the client's representatives, are based on their specific use of the instrument. Interviewees can only reflect on their own implementation and have limited experience and knowledge of other formats. The experiences as presented in this chapter find their origin in one of the cases. When interviewees refer to an observation from other previous assessments, this is mentioned explicitly, but these projects are anonymised as well.

8.1 Communication between the client and bidders

The first set of experiences focusses on all kind of written and face to face communication about the team assessment process with the bidding parties. This includes sharing the reasons to apply the criteria in the beginning and managing expectations of the concrete procedure of the assessment.

8.1.1 The client's procurement strategy

A vast majority of the Dutch interviewees addressed the importance of sharing information with the market about the client's motives to apply team assessments (Interviewee 3, 4, 5 and 7). At the same time, clients indicate that they are struggling in determining what to share and to what extent. "*In several projects, I observe that handing over more information is not always satisfying or serving the purpose*" (Assessor 2). Assessor 2 noticed that clients are sometimes tending to share too much or too little information, both making it hard for bidding parties to determine 'what is it really all about?' – referring to the project challenges and client's higher goals (Interviewee 5). Next to tender guidelines, sharing information with the market is done via market consultations, and dialogue and tender kick-off meetings (Van Elburg, 2008).

According to interviewee 5, the tender guideline is essential to let market parties understand why the client takes specific steps in the tender and what they want to achieve with these procedures: "*A party that slips through the net and wins the tender but does not understand what you aim to achieve is completely not in the client's interest*". (Interviewee 5). In case N4, the client used almost two-thirds of the tender guideline for explaining the tender goals, criteria and client's vision (Interviewee 5). Two interviewees argued that sharing these considerations might be even more important in the domain of collaborative criteria and goal of enhancing collaboration throughout the procurement phase (Interviewee 3, 4).

8.1.2. The assessment procedure

Next to the rationale for the team assessments, clients also need to address the course and content so expectations are appropriately managed. In case N4, the client organised two kick-off meetings during the tender phase: one of them before the team assessment when there were three bidding parties left. One of the sessions during this kick-off meeting was a workshop by the assessors about the subsequent assessment. "*In this way we informed the market well about what they could expect with regard to the team* *assessment*" (Interviewee 4). Case N2 also succeed in this goal, as confirmed by the participating tenderers in an external evaluation of the tender phase: "*The participating project teams of the tenderers have had a clear and consistent view and sufficient prior information to the assessment*" (Non- disclosed, 2019).

The need to put effort in the expectation management upfront is reflected in an observation by interviewee 7 (case N6). One of the tenderers that had lost thought that they had not presented well in the team assessment in terms of the content of the case(s). *"And that was not true at all; they had scored well. The substantive result of the case did not matter at all*" (Interviewee 7). It is crucial that the market is informed about this difference since participating in a team assessment is still relatively new for construction parties. Managing expectations can prevent wrong interpretations and even disputes.

This observation is supported by the British Nuclear Decommissioning Authority that has conducted several assessments in 2007 and 2008. "*Early in the process, explain to bidders the rationale for the assessment, the issues that they will address and how they might prepare*" (Doyle and Jones, 2009). Tender documents and a face-to-face meeting should address the purpose, structure, review and processing of the assessment in detail. This is considered to be crucial for new procurement exercises like team assessments.

8.2 Briefing sessions and practising

Looking back on the steps taken before the team assessment itself, interviewees consider their preparation as a crucial step to ensure a level playing field and in this way, exploit the value of assessments.

> Interviewee 5 (Dutch): "*As the client team we played the assessment ourselves once - so we had prior knowledge before the first real assessment. That preparation is really valuable."*
>
> Interviewee 8 (British): "*There was a one-day briefing session with those client people involved in the assessments: the do's and don'ts and the importance of a level playing field."*
>
> Interviewee 10 (Finnish): " … *We had to train our self as well [the client team]. With the help of a consultancy company, we created a guideline 'how to behave as client during these workshops*?"

All five assessors involved in this study underline these client's experiences. A dry run is considered crucial to conduct proper team assessments (Assessor 1, 2 and 4). "*Despite the specific implementation, we organise a whole day of practising with the client's team. This is necessary to get the right start in the first real team assessment*" (Assessor 1). However, several Dutch clients did not see this importance on forehand.

Client's representatives are allowed to react responsively during the assessments but need to build up a certain tension in their reaction. This should go the same each assessment (Interviewee 4; Assessor 2). Because Finnish clients invite significant more members to their workshop, this preparation is seen as even more crucial. In case F1, quite a lot of experts from the client-side were taking part*. "It should be clear for everyone how to behave: don't take the leading role*" (Interviewee 10). Briefing all members took several hours before the workshop and was considered more crucial when this was new for every participant.

Across nine cases of this study, the client's team participated in the team assessment. Those teams had to repeat the assessment for three, four or sometimes five times. The briefing sessions and doing a dry run are initiated by the assessor involved since assessments are new to most clients. However, interviewees within each case agree on the importance of at least (half) a day practising after participating (Interviewee 1, 3, 4, 5, 6, 8, 9 and 10). A briefing session and dry run of the team assessment is crucial to ensure a level playing field and gives the participating members confidence (Assessor 2). In case N6, the client's team did not participate in the real team assessment. Nevertheless, the team has gone through the assessment to show its importance. This also served as a practice round for the assessor (Interviewee 7; Assessor 3).

As explained in Section 7.5 about one of the team assessments in case U1, a member of the client's team decided to be the tough [client] project manager. Briefing sessions in advance and (even) more practising can prevent such unexpected behaviour (Interviewee 8).

8.3 Course of the team assessment

All interviewees within the ten cases were asked about their personal experience from the team assessments that were conducted. This experience could reflect a feeling of the whole client team, but in the end, this is the perception of the client's reprehensive that has been interviewed.

The first and more general finding that can be observed in a majority of the cases is that the end-result of all team-assessments was according to expectation. The preparation and instrument itself resulted in the right setting: differences between tenderer's teams became visible.

> Interviewee 5: "*After the second day we were looking at each other with big eyes. It was so different: the behaviour of the candidates ... Same case, same team from our side, only the team from the contractor was different. There was such a different atmosphere and process. After all the team assessments, we could really identify what was distinguishing between the parties*."

> Interviewee 4: "*I was very satisfied with the team assessments afterwards. The good preparation had paid off ... You could see that a candidate had one good manager, but the rest of the team let him swim. And the tendered that was ultimately selected actually moved along as a real team.*"

> Interviewee 6: "*The assessment has lived up to expectations. It is distinctive. An example: a market party that participated in the team assessment was just put together and had not seen each other beforehand. This came out immediately during the team assessment*."

> Interviewee 10: *"I was happy after the first workshop for this project. We could see some differences between the bidding teams .. We had the same feeling which team was doing well*."

While some interviewees doubt in advance if the assessments would bring different results, the experiences show that each tenderer brings a new atmosphere and course of the team assessment (Interviewee 4, 6, 10). This view is supported by the experiences of assessors (Assessors 1, 2 and 4).

8.3.1 Sharing and substantiating the scoring

After the assessment and award of the contract, proper communication is again considered necessary since the scoring has to be shared. Even when assessors have taken notes of each individual involved, the final review is always conducted on a team-level. A tenderer's scoring is communicated as one cumulative grade (Interviewee 4). All five assessors that have been interviewed emphasised that they have put much effort in these scoring reports. Partly this is related to the procurement context and clients do not want to be challenged, especially when the assessment turned out to be decisive (Interviewee 3).

Regulations do not require to debrief all of the contractors, but it has been greatly apricated and logical to do so since tenderers have spent much effort in the bidding process (Assessor 4). Interviewees unanimously share the view that this is greatly apricated by the market (Assessor 1, 2 and 4; Interviewee 5 and 10).

It differs per country and case until what extent the reports have been shared. In case F1, each tenderer could read the scoring and report of the other parties, while most Dutch cases chose to just hand over the tenderer's own review. In some projects, the client cannot see the report without the approval of the tenderer (Assessor 1). It's essential that the client and assessor make a well-considered decision at front.

For case N3, interviewee 4 and assessor 2 set up a scoring card for the team assessments. Potential outcomes (scorings) could be a two, a four, a six, an eight or a ten. The tenderer that was given a six in the end felt disappointed but also a bit personally insulted (Interviewee 4). The client had interpreted the number '6' more like the conclusion 'we could have done this project well with you' while the contractor's feeling with the number '6' was more 'with the grade seven you can come home, and a six feels like a blame' (Interviewee 4). The team assessment and its review feels clearly more personal than the review of a more traditional written submission by the whole tender team.

The comprehensive review of the tenderer's scoring is greatly apricated but also serves the increasing transparency of the whole process (Interviewee 1, 3, 5 and 10). Moreover, expectation management and a well-considered scoring review are needed (Interviewee 4; Assessor 4). It must be noted that no client or assessor involved in this study has experienced tenderers that challenged the scoring or assessment approach. "*There has been plenty of pushbacks in other areas, but never on behavioural components*" (Interviewee 8). It seems that the robust scorings and vigilant attitude serve its purpose.

8.4 Subsequent collaboration

Besides the experiences focussing on the team assessment environment, interviewees were also asked about their perspective on the effect of the assessment on the following collaborative relationship with the contractor(s). Since it is considered to be impossible to study the causal relationship between the different assessments and collaborative relationship post-contract, the observations of this paragraph are one of the best approaches to see the effect of team assessments.

8.4.1 Starting a collaborative relationship

Team assessments are conducted to score the tenderer, but these moments of interaction between (one of) the tenderer(s) and client's team are also seen as (one of) the starting point(s) of the future collaborative relationship. Several interviewees expressed their experience that the competitive dialogue meetings, which took place after the assessment, have benefited from the intensive interplay during the assessment (Interviewee 1 and 5; Assessor 1). The first dialogue meeting starts at a level of interaction where usually the second dialogue meeting would have been (Interviewee 2). This works out during the whole tender phase. *"It was more difficult than I expected to say goodbye to two tenderers in the end. That was the result of the intensive procedure with the competitive dialogue and team-assessment*" (Interviewee 1).

This observation of interviewee 1 does not only address the assessment but also other contact moments, which makes it harder to reflect on the team assessment independently from the rest of the tender phase. Interviewee 10 addressed that it was much easier to act and work as one team from the very beginning due to the workshops and competitive dialogue meetings. The basis of trust has been there from the first day, which was partly due to those team-assessments (Interviewee 2, 10). " *Collaboration requires maintenance, and it requires trust. That means doing evaluations with each other, celebrating successes, sharing a lot. The team assessment was the starting point for building our collaboration*" (Interviewee 6).

This first observation regarding the benefits for the competitive dialogue meetings and later collaboration can be linked to the fact that a team assessment is an extra contact moment where people get to know each other's behaviour and characters. More contact moments between both parties simply lead to a more developed relationship (Walker and Hampson, 2008).

The second finding touches upon the deliverables of the team assessments: observations of the stronger and weaker spots of an integrated team. Several cases of this study have used these lessons in their PSU [Project start up] and following coaching sessions. "*You can continue with the development points that were already there [from the team assessment]. We see that the interaction and reflection really goes three steps deeper than it would on a normal project*" (Interviewee 2). This experience is also seen in Finland. The result of the project alliance procurement model, including these workshop, was that both teams knew each other already well when they were signing the contract (Interviewee 10 and 11).

In case U1, the client organisation set up a behavioural team that took over from interviewee 9 after awarding and would focus on the collaboration. In this way, the collaborative criterion and corresponding assessment was just the first benchmark. The client applied subsequent assessments in the delivery phase to see if a team developed and keep addressing attention to this topic (Interviewee 9). In case N6, all six competencies that were assessed via the team assessment still have a place in the project today via coaching and on-boarding programs (Interviewee 7). In case U1, the client also asked the assessors to assess the client's team during the team assessment separately and deliver a second report. When the client's team and the contractor's team started the job after award, both reports were shared as part of the joint team development (Assessor 4). It must be noted that there is difference between a scoring report, that is needed according to procurement regulations, and a learning report that focusses on points of attention for the future collaborative relationship (Assessor 2).

The two findings of this paragraph strengthen the observation that, across multiple cases, the interaction during the team assessment, in which both client and tenderer are present, positively influenced the competitive dialogue meetings and following team development. This observation is related to Motive III.

8.4.2 Returning to old patterns

Besides those positive findings, which can be categorised as positive experiences and effects of the team assessment, interviewees also addressed differences between ambition and reality. In case N6, the experiences of conducting a team assessment were prevalent positive, and the intention of the contractors side were good as well. Nevertheless, "*sometimes they [design team members from the contractors)] subconsciously fell back into the old pillars and patterns*" (Interviewee 7). This is also seen at the client-side. Client F, which had set up a behavioural team, faced difficulties in building on the outcomes and lessons of the assessment. The client organisation got overtaken by events and other challenges (Interviewee 8).

Everyone starts very collaboratively after the award of the contract - "why not?" – but when there are problems, corporate entities tend to evert the contractual financial positions (Interviewee 8). This observation touches upon the essence of this study: the collaborative behaviour is needed most on these moments of (potential) conflicts (Assessor 1). According to interviewee 8, it is essential to have a construal model that supports collaboration and helps sustain it. When a client applies a very lump sum risk allocation combined with collaborative criteria, the commercial aspects will always overrule (interviewee 8). This relation between the project characteristics and the assessment is further discussed Section 9.3.

It can be concluded that the team assessment I just the start of a longer process of emphasising on the collaborative relationship. Both parties can easily fall back into old patterns. Just like the previous observations of this chapter are related to Motive III, this critical note can be as well. Clients who bet heavily on team assessments to start and focus on the collaborative relationship in the tender process may face this behaviour of sticking to old patterns and pillars. In other words, Motive III should not be a naïve driver.

8.5 Potential value of team assessments

This last paragraph of Chapter 8 aims to zoom out even more than the previous paragraph did. To determine the value of team assessment in the tender process, client's representatives are asked about the observed positive and negative effects. In this way, experiences are discussed from a less operational point of view.

8.5.1 The team assessment as a signalling function

In most cases, the client was surprised by the strong signalling function of collaborative criteria, and the team assessment in particular. According to several interviewees, this signal is about the importance of the soft elements and human factor for project success. By conducting a team assessment in the tender phase, clients show that they strive for an improved collaborative relationship (Interviewee 5).

The clients of case N1, N3 and N4 figured out that one or more tenderers considered several compositions of the leading project team because their competencies could be crucial for being selected (Interviewee 1, 4 and 5). *"Just by using an assessment, others start thinking about cooperation. It also has a signalling function."* (Interviewee 5 and 9). The winning tenderer of case N3 had put much effort into the composition of their team: "*the consortia that was selected have told that they did not let the management decide who would join their team, but an organisational psychologist. That is quite special*" (Interviewee 4).

Even when the scoring of a team assessment is not decisive in the partner selection, it still has value. "*We know what this company is about and we know 'what is wrong'. If certain people scored very low in the assessment, we would use this in the negotiations after award on which members would keep on board*"

(Interviewee 9). This intervention is not seen in the Dutch or Finnish cases. However, it can be concluded that the motive to emphasize on the importance of soft skills works out in most cases, irrespective of the specific application and weighting of the team assessment.

8.5.2 The team assessment and client organisation

The signalling function and emphasis on soft elements is also relevant within the client organization. "*The consequence of using these criteria, including a team-assessment, is that you are on both sides looking on for people who are able to work together and collaborate in a very early stage of the project*" (Interviewee 2). Besides this observation from case N1, also in case N3 and N5 it was explicitly found that the organisation looked to the client's team in a different way and with a positive end-result (Interviewee 4 and 6).

On the other hand, this study also observed a more critical footnote about the client organisation related to team assessment across British and Dutch cases. Case U1 had put much effort into their assessment centre and tested the contractors in multiple scenarios during a team assessment. However, client F figured out which cultural and behaviours they wanted, but they did not look internally to themselves on a comparable level (Interviewee 8). In some sub-contracts, the contractors ended up with a more traditional client team that was not particularly collaborative. "*We sometimes joke about it, but some experienced guys here still see collaboration as 'you do what I want .. Just collaborate with me*" (Interviewee 8).

Assessing your own key people as a client is considered to be essential. This can be prior to or during the assessment. Especially in larger and more traditional client organisations, it is considered to be difficult to get the best people that fit this new approach. "*Within our organisation, we may also have to take a closer look at whether people fit the task. On our side, that may be done too much based on capacity and availability*" (Interviewee 3). In this line of reasoning, the attention for collaborative competencies is not that dominant within the client organisation as this same organisation has asked the market.

8.5.3 The team assessment compared to other instruments

Several interviewees argued their preference for real-life team assessments compared to the other instruments as highlighted in Section 7.1. All written instruments (tests, a vision document, surveys) do not say anything about the team you will see on the project after the award of the contract (Interviewee 4 and 5). All kind of collaborative criteria can have raised specific attention (the signalling function as described previously), but in the end, the added value and predictive value of a team assessment are more substantial (Interviewee 4, 6, 8 and 9; Assessor 4). The difference of a team assessment compared to other criteria is that the commission authority gets a feeling about what the teamwork will be like during the upcoming project. It is not about beautiful plans, it is about behaviour (Interviewee 6). Projects are all about trust, and the real-life assessments are a new starting point for that (Interviewee 10).

It can be concluded that a majority of the interviewees that have been involved stated that, in their view, the team assessment has a clear added value compared to the other instruments since this instrument is assessing authentic behaviour of the team who will be on the project later on. An assessment of (half) a day makes it impossible to pretend to be someone else. In other words, team assessments do make it possible to assess authentic behaviour and avoid window dressing (Interviewee 1, 5 and 9).

However, it is derived that the added value of a team assessment is no automatic value that's created with every assessment. It can be matched to the previous interpretation about the assessment set-up and role of the assessors involved. "*I completely support and see the benefits of behavioural assessments, but how we delivered it is crucial for the benefits*" (Interviewee 9). The value of team assessment is very closely related to the decision what you [a client] will do with the results of the assessment after the award of the contract. Assessor 4 compares it with undertaking a risk assessment. "*Unless you mitigate the risks by post-contract award activity, it's nearly a waste of time. There needs to be a total end to end approach, form procurement to project start-up to delivery. You need the full cake*" (Assessor 4).

8.5.4 The growing attention for team assessments

The increased intention for team assessments also has a potential downside. It can lead to easy copying and less customization. Some clients seem to implement the assessment in more or less the same way of another successful project (Interviewee 6). In the British construction sector, team assessments have become well-known over the last decade. Interviewee 8 sees a phase coming of 'everyone is doing this so we should do this as well'. "*But without discrimination, the industry needs to step back a little bit*" (Interviewee 8). "*The idea 'we have to focus on the collaborate relationship because everyone does it' is a wrong attitude*" (Interviewee 9). The implementation of collaborative criteria and the following relationship should depend on the project, the contract and the project dynamics.

Every client should ask themselves the why-question, but perhaps in a more distinctive way than before (Interviewee 8). In the Netherlands, this call for a custom approach and distinction is heard as well. "*I get the impression that the team assessment can become a unity struggle. And that's a bad thing.*" (Interviewee 6). Client's representatives that are involved in this study want to send the warning that a team assessment is no miracle cure and it should be prevented to become an easy copyable instrument (Interviewee 3, 6 and 9; Assessor 2). "*The team assessment should not be a blueprint solution for something that does not lend itself to that. And that's what I'm afraid of if we only make it bigger*" (Assessor 2).

The growing use of team assessments increasingly demands the need for customisation. It is still possible to conduct an unique assessment, but it is also harder and even more important now there are teams that have already done them several times (Interviewee 8). Interviewee 8 raises the attention to keep conducting these assessments unpredictably. This is mainly the responsibility of the assessors involved.

8.6 Concluding: experiences of applying team assessments

The team assessment experiences have been discussed chronology in order to answer the third sub-question. Assessor's experiences are used as an additional view to strengthen the client's observations.

SQ 3 | What are the client's experiences from conducting team assessments?

From each stage of implementing team assessments, one or even more sub-conclusions were extracted. Two higher-level conclusions address the broader value of the instrument while all others focus more in detail on the experiences from the course of the team assessment itself.

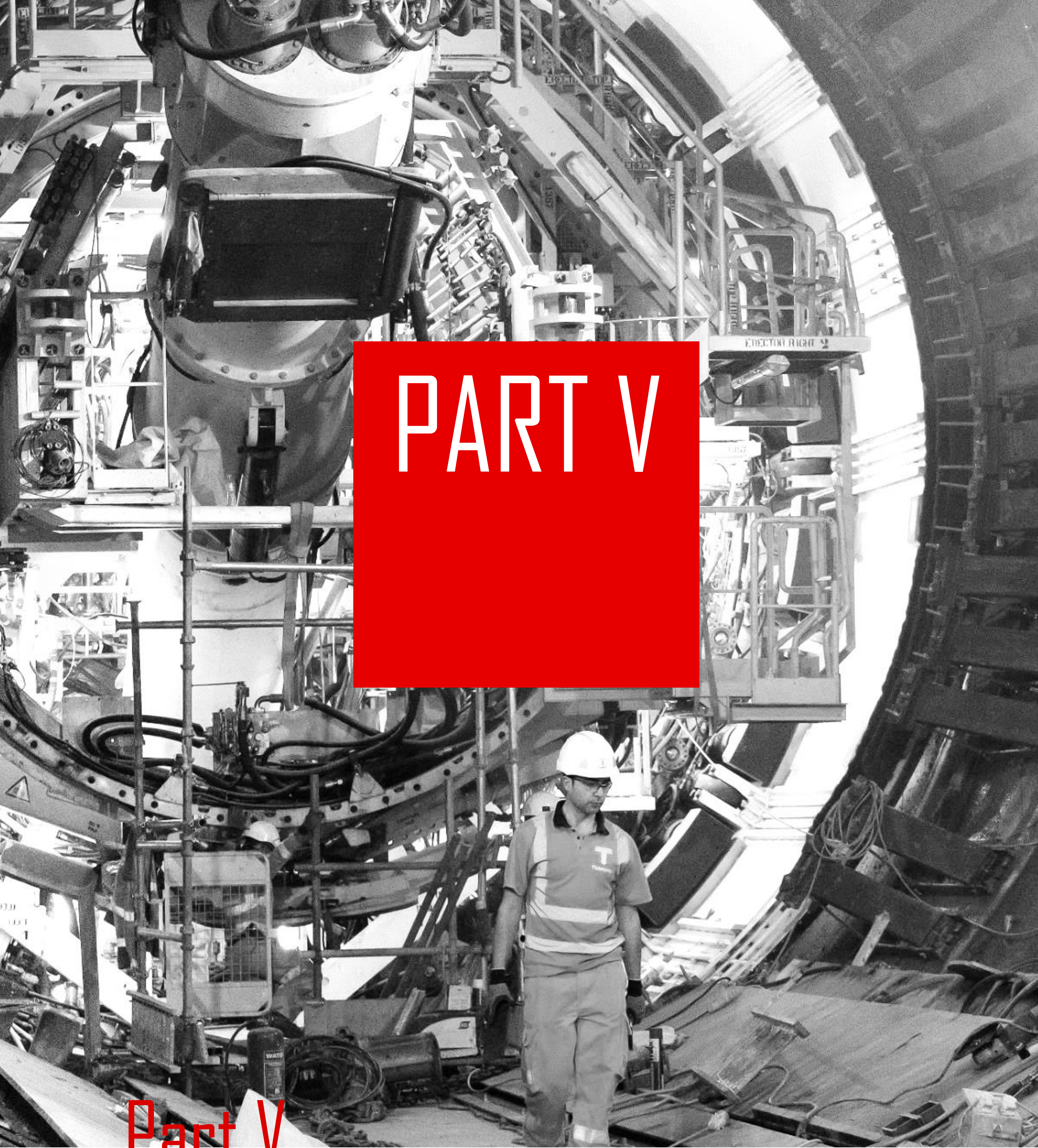
The higher-level client's experiences of team assessments are all very positive. Clients see the instrument as a valuable thing that has a relatively strong signalling function. Interviews noticed that bidding parties are looking to collaboration and their team-selection in a different way. More attention is paid to the soft elements compared to more traditional tender criteria. Moreover, the dialogue meetings and project start-up after the assessment benefit from the interaction and reflection of the team assessment.

Despite some remarks and prerequisites for further use, all ten clients are predominantly positive about the specific course of their team assessments. The most significant remark being made by both clients and assessors is the fact that the assessment is no miracle cure for current problems. Ongoing attention is needed to enhance collaboration. The assessment as part of the tender process is just one way of doing so.

A piece of relevant advice that is given by five interviewees focusses on the fact that both parties can easily fall back into old patterns. A team assessment is just the start. Interviewees address certain differences between ambition and reality. The intention of both sides are always decent; however, people still could fall back into the old pillars subconsciously. It can be concluded that the team assessments are just the start of a long process of emphasising the collaborative relationship.

This study found four more detailed experiences of the specific course of the team assessments.

1. **Expectation –** The market needs to know what to expect from the assessment and the client's team (it is played on behaviour, not on the content). Managing expectations can prevent such wrong interpretation and possible disputes after the tender phase.
2. **Preparation –** A practice day and dry run by the client team is crucial to ensure a level playing field. All interviewees agree on the importance of at least (half) a day practising as client team. It also gives the participating members certain confidence.
3. **Reflection -** Each tenderer makes a different course and outcome. Some interviewees doubted in advance if the assessments would bring different results. The experiences show that each tenderer's team brings a new atmosphere and course of the team assessment.
4. **Evaluation –** Putting much effort in the tenderers scoring and review is appreciated and needed. A comprehensive review of the tenderer's scoring is greatly apricated but also serves increasing transparency of the whole process towards the market.



Part V

In this fifth part, the observations from the previous chapters are interpreted. Chapter 9 discusses the value and use of collaborative criteria. In addition, the (improved) procedure for implementing team assessments is described in Chapter 10. This is also the answer to the fourth sub-question. Both chapters will refer to previous conclusions, but also make use of new observations from interviewees and the literature.

9. Comparison & interpretation

By interpreting and comparing the observations from the previous part, this chapter aims to zoom out from the specific cases and weigh the benefits and footnotes of collaborative criteria. This will lead to advises and considerations that can be used in answering the fourth sub-question in Chapter 10.

Hence, (9.1) the value of criteria and instruments is discussed first. In this way, the findings related to the team assessment are put into the broader context of this study. A similar approach is used for (9.2) comparing the similarities and differences between the countries that have been studied. Subsequently, it is explored what kind of (9.3) infrastructure projects are fit for collaborative criteria in the tender phase.

9.1 Reflecting on collaborative criteria

Based on the observations of this study, the value of collaborative criteria is discussed.

9.1.1 The motives of applying team assessment

Based on the observations in Chapter 8, it can be concluded that both clients and assessors consider the team assessment as a valuable instrument that has a strong signalling function. Across several cases, interviewees noticed that bidding parties are looking towards collaboration and their team-selection in a different way: more attention is paid to soft elements compared to more traditional tender criteria (Section 8.6). Moreover, the subsequent dialogue meetings and project start-up after the team assessment have benefited from the interaction and reflection of the assessment.

In all cases, a conscious choice was made to use a team assessment next to the other instruments for three reasons: 1) it assesses authentic behaviour instead of an ambitious written plan, 2) the future collaborative relationship has already been brought forward and developed in the tender process, 3) the challenging environment of simulated assessments results in knowing where one stands, the good and bad points. Despite the remarks that are being made as well, these three drivers of team assessments are considered to have added value when assessments are conducted correctly.

Figure 14 shows these three distinguishing elements of a team assessment together with the answer on the first sub-question, In this way, the higher and more specific drivers of team assessments are derived.

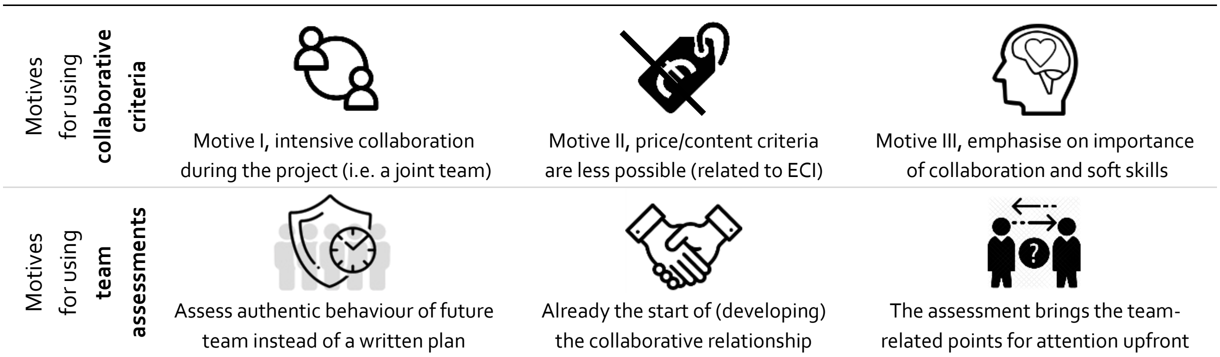


Figure 14 | The motives of using collaborative criteria and team assessments

The lower row of Figure 14 displays the value of team assessments that belongs to this specific instrument in particular, and cannot be identified across any of the other instruments of Section 7.1. Although, it must be noted that the interview is also an instrument to assesses more authentic behaviour/responses.

9.1.2 Discussing the alternatives of laying more emphasis on collaboration

Two of the cells in Figure 14 focus on applying team assessment to start and focus on collaboration throughout the tender process. The most significant remark being made by both clients and assessors in this study is the fact that the assessment is no miracle cure for (expected) problems. Ongoing attention is needed to enhance the client-contractor collaboration. The assessment as part of the tender process is just one way of doing so. "*Which instrument you use for 'collaboration' in tendering is less important in my opinion. The team assessment is a tool for that, but so is coaching later on*" (Interviewee 5). The team assessment, together with other collaborative criteria, resulted in a relationship that clients were looking for, but there is no causal link found between one instrument and the rest of the project.

Conducting a team assessment after the award of the contract is another possibility of building on the collaborative relationship (Assessor 2 and 3). Such team assessment could be more simple and effective since there is only one objective: improving the interaction between the parties (instead of scoring the bidding parties within the context of procurement law). The competitive element of the tender phase is not present there, which could result in a more pleasant environment for the participants (Assessor 2).

Moreover, the current goal of both parties during team assessments in the selection phase is not uniform. A tenderer benefits from doing something very clever and therefore becoming first (Lo et al, 2007). After all, the tender phase is a competition phase until the moment of awarding the contract.

On the other hand, by following this suggested application the partner selection itself is not based on the collaborative behaviour of tenderers. Moreover, the signalling function of collaborative criteria (see Section 8.5) is missing. It can also be harder to align values after awarding of the contract when the client organisation has less leverage (Interviewee 8). Expanding coaching and structural evaluations during the realisation phase are encouraged by assessors (Assessor 1, 2, 3 and 4), but the client's representatives mostly see this as a supplement to a partner selection with collaborative criteria (Interviewee 5, 6 and 10). To sum up, Figure 12 does not apply in its entirety when team assessments are used post-contract.

9.1.3 Resilient partnership

As stated in the previous paragraph, the most significant remark being made by both clients and assessors in this study is the fact that the assessment is no miracle cure for (expected) problems.

The recent report '*Toekomstige opgave Rijkswaterstaat*' mentioned that an insufficient identification and management of risks creates an undesirable incentive to the market that leads to a low bid price, but also increases the risk profile and puts pressure on the collaborative relationship later on (Rijkswaterstaat, 2019). The introduction of collaborative criteria and a team assessment in the tender process will not change these incentives. It is no guarantee that there will no longer be any problems or conflicts between contract partners in the realisation phase of a project; simply too much will change during the execution. However, a proper collaborative relationship leads to resilient teams (Assessor 1).

"*I see too often that the conflict is not allowed to be there. And cooperation or collaboration is not about that*" (Assessor 2). According to several assessors we have to deal with the conflict well, instead of leaning to get the conflict out. At this moment, the construction sector is still too often thinking of an extra rule when something goes wrong. That does not bring resilient teams, or resilient partnership (Ruijter, 2019).

Instead, we should be able to fall back on a common standard instead of supplementing and tightening up those rules. This view supports the use of other cultural alignment instruments (Section 7.1) that bring a shared view instead of more rules. "*We need to take more time to think about this and understand each other"* (Assessor 2). A team assessment can enhance this process, but it can also be wrongly seen as a new instrument that prevents problems. This observation also relates to the discussion about the participants of the assessment. Dutch interviewers and assessors are not in favour of involving the tenderer's directors via interviews or assessments. "*It's much more valuable to look at how the project team looks at governance: 'how to deal with and take the right leadership role'*"(Interviewee 4). Clients and assessors can make the assessments even bigger, but above all, we might have to accept that things can go wrong (Interviewee 6; Assessor 2). Project teams need the skill to develop to deal with these situations (Assessor 2). When that's done, clients have to make sure that the culture of cooperation does not erode, for example when new colleagues arrive and others leave. Continuous maintenance is necessary (Hertogh, 2017).

9.2 Differences between the three countries

Across this study, the variances between the three countries already have been highlighted on a more detailed level (i.a. Section 6.5; Section 7.4; Section 7.6). This paragraph will not answer a new sub-question but observe the conclusions of the first three sub-questions in unity and show similarities and differences between the application in the Netherlands, United Kingdom and Finland. By doing so it is possible to derive recommendations for the Dutch construction sector.

9.2.1 Different applications of team assessments

The differences between the three countries on aspects like content and participants are highlighted first. British projects use a combination of online and face to face instruments for assessing collaborative behaviour (Construction News, 2016). So-called 'assessment centres' are created compromising all kind of instruments. In case U1 and U2, four to six instruments were applied, separately scored and made up the total score. The British approach differs from Finland and the Netherlands in their extensive way of defining the best instruments and people to involve. The content and extent of the team-assessment could even depend on a risks assessment undertaken by the client in the beginning of developing the assessment.

On the other hand, the British and Dutch team assessments do show some fundamental similarities on the type of exercises and assessment ambience. British and Dutch assessors apply a combination of general and business-related exercises that are embedded in a fully simulated environment, while Finnish workshops also focus on solving real project issues in order to improve the bid of potential alliance partners. Because the Finnish workshops are not automatically about conflicting interests, the setting and ambience of the workshop is automatically less tense. These workshops are not only assessing the tenderer's collaborative behaviour and alliance competencies but also improving the identified list of risks for example. Dutch and British interviewees underline that it is crucial to create an ambience where a tenderer is forced to show authentic behaviour. This can also mean looking for a particular tense where interests could be conflicting and interventions during the simulations are helpful.

Concerning the assessors, in the average British behavioural assessments, the number assessors is significantly more compared to Dutch and Finnish assessments. In all non-Dutch cases, observers were appointed by the client to give feedback afterwards and 'protect' the process. These people ensure that the client's team is behaving correctly, and there are no benefits for one tenderer in particular.

In the Netherlands, the client's side is in almost any case represented by five managers. A bidding party is asked to represent the same disciplines. In the United Kingdom, the selection of participants within the assessment centre depends on the specific project risks and challenges. Some British cases included interviewing key directors as part of the assessment centre, and also the Finnish cases included different roles. One Finnish workshop involved people from several levels of the future alliance organisation. These members of an alliance 'steering group' are not assessed for Dutch project alliances. This could be a relevant consideration for future projects since Dutch interviewees of project alliances addressed the importance of the interaction between the 'parent organisations' (Interviewee 2 and 6).

9.2.3 Concluding: looking to the British and Finnish sector from the Dutch perspective

On the basis of this desk and case study, it can be concluded that the British sector is leading in this domain. Their use of behavioural assessments differs from the other countries in its wide range of formats; including more instruments, assessors and observers. Meanwhile, the British and Dutch assessments show similarities in their course, type of exercises and ambience. In both countries, a combination of general and business-related exercises is used in a fully simulated environment, while Finnish workshops also focus on solving real project issues in order to improve the bid of potential alliance partners.

Based on these differences, this study suggests that both Dutch clients and Dutch assessors could learn from the comprehensive guidelines and application of British team assessments. The common British application is not about a higher weighting of these criteria; on the contrary, it is the level of detail that is different. Dutch assessments could be improved by including prior risk assessments, more instruments and observers. Moreover, the embedding of collaborative criteria in the tender phase can be enhanced by appointing a behavioural team on the client-side. Future Dutch project alliances could also look at the extensive workshops by Finnish clients of project alliances. Their workshops involve several levels of the future alliance organisation, which is not seen in the Netherlands yet.

9.3 Infrastructure projects that are fit for collaborative criteria

The overlapping characteristics of ten cases are not automatically defining the domain of the infrastructure projects that are fit for collaborative criteria. Projects like case N4 can be seen as the active vanguard, while the potential of these relatively new criteria can be relevant for a larger set of projects.

In order to provide input for Chapter 10, the potential of collaborative criteria is discussed from a broader perspective than just the cases of this study. All interviewees (including the assessors) were asked about their view on the projects that should consider using collaborative criteria, Across the sixteen interviews, two dominant project related views were identified - some interviewees mentioning both. A third view does not focus on the project characteristics, but the attitude and culture of the client organisation involved.

9.3.1 Collaborative criteria have a potential for all kind of projects

The first view follows the fundamental belief that any construction project may, to a greater or lesser extent, benefit from these collaborative criteria. This essence is reflected in the citations below.

> Interviewee 5: "*Project work is people work. You can run a great project with a bad contract but not with a bad team. I'm really convinced that for every project you can do a little more here."*

> Interviewee 6: "*Numerous contract forms can benefit from this focus on collaboration. The contract is not decisive: if you select the right people, you have a better chance of a successful project."*

> Interviewee 8: "*Certainly in the infrastructure domain we will never be away from levels of high uncertainty. So in that kind of arena, you need this kind of methods*."

This first view is driven by the belief that you have to select the right partner (and team) for your project and organisation, irrespective of the project delivery method. A client could always benefit from drawing attention to collaboration in the tender process (Interviewee 7). Nowadays, more and more clients recognise that every large contractor can do almost anything on a technical level, but the soft elements determine whether the project will be a success (Eriksson, 2010). "*Choosing your partners will, therefore, become more important in the construction sector*" (Interviewee 6).

What most projects have in common is that there is a client and contractor that have to deal with changing situations. The interaction between the client and contractor, and those with other stakeholders, is crucial in these changing situations, which brings the question 'what do you do when it gets challenging' (Interviewee 6; Assessor 1). Every public client should be acquainted with these soft criteria.

9.3.2 Projects with intensive forms of collaboration

Most interviewees have nuanced the view of Section 9.3.1 by stating that different projects demand for different approaches. This second views adds a nuance to the first perspective. Collaborative criteria and its corresponding assessments should be mainly applied to projects involving an alliance partnerships or intensive and longer forms of client-contractor collaboration. This view is reflected in the citations below.

> Assessor 1: "*the more integrated collaboration… the more uncertain the scope, the more risks need to be sorted out …, which means that you need to make greater use of those collaboration skills*"

> Assessor 2: "*I see this instrument [the team assessment] suit to more long-term partnerships, client-contractor collaboration of ten to twenty years with a high level of mutual dependence.*"

> Interviewee 5: "*This makes sense in projects where the environment is so complex, and you have to collaborate intensively – when the environment and alignment are very much in motion."*

This view matches Motive I as identified in this study: collaborative criteria particularly suit projects with intensive collaboration (and a joint team) in the design or realisation phase. However, it is not believed that the increasing attention for early contractor involvement and two-phase model (Rijkswaterstaat, 2019) automatically bring more collaborative criteria. Some clients who are applying a design team are primarily driven by the technical challenges they're facing (Interviewee 5). When there is just limited complexity and dynamics in terms of the environment, stakeholders, process and regulations, it is not logical to put heavy emphasis on collaborative criteria. Following this line of reasoning, Motive I and II should be nuanced by stating that early contractor involvement is not the driver, but the intensive collaboration in specific. The cases in this study had to deal with uncertainties concerning their surroundings and requirements, so from the perspective of the collected empirical data, the statement of interviewee 5 is not refuted.

9.3.3 The client organisation

While the first two views focus on the project characteristics to determine the relevance of collaborative criteria, this third view looks at this dilemma from a different angle. It includes the (culture of the) client organisation and signature of the client's team members. The trigger of this view was interviewee 7 who noticed enthusiastic reactions within the client organisation but also stated that this will not automatically result in more team assessments over the coming years. "*It also depends on the people on the client's side on such a project. Everyone has their own signature*" (Interviewee 7). This signal is seen across more cases in the Netherlands and the United Kingdom. Since the Finnish alliance model can be seen more as a blueprint that is being improved after each alliance, these differences of signature are not detected there.

As discussed in Section 8.5.2, client F figured out which cultural and behaviours they wanted, but did not look internally to themselves on a comparable level (Interviewee 8; Assessor 4). A new director within the client organisation had bad experiences with applying team assessments. The view of such director had an impact on the embedding of collaborative criteria. You could even argue there are just two options: you're in favour of these relatively new criteria and think it's important or you think this is a waste of time. "*Both sentiments have been expressed on board level, and that will be the case in the future*" (Interviewee 8). This aspect is put forward in the factors of Section 10.2.

Implementing team assessments like those of case U1 and U2 (relatively large assessment centres for several sub-contracts) costs around half a million pounds (Interviewee 8). However, there is no study yet showing how much you will going to save by this investment. "*This is civil works mentality: we try to protect the public purpose, but this [a very critical attitude] is the nature of publicly funded clients and projects. It depends on how you're client leadership are*" (Interviewee 8). After all, the implementation of collaborative criteria depends – maybe even more than other criteria – on the leadership, preferences and culture within the client organisation. Some clients want to get a gut feeling in the tender process: 'will this be a good team?' while others may want to know more about the content and knowledge of the candidates (using BVP). "*It is the type of client, not only the type project*:" (Assessor 2).

9.3.4 Concluding: projects that are fit for collaborative criteria

This section discussed three different views on the potential of collaborative criteria within the construction sector. Although the essence of each view differs, they are not mutually exclusive. The first view can be complemented by the nuance of the second and third view.

The first view followed the basic belief that any construction project may, to a greater or lesser extent, benefit from these criteria. Most interviewees also nuance this view by saying that some projects and contracts are more suited to these soft criteria. This is also supported by studies of Kumaraswamy and Anvuur (2008) and Eriksson (2008). The second view states that projects with a dynamics environment, processes and regulations benefit more from the focus on collaboration in the tender process. A third view looked at this dilemma from a different angle and includes the (culture of the) client organisation and signature of the client's team members. This is supported by findings in Sections 8.5.2 and 8.4.2.

10 Implementing team assessments

The conclusions of both Part IV and the previous chapter are used to derive the general steps of a team assessment as part of the partner selection process. Subsequently, these steps are enriched with measures and considerations to improve team assessments in practice. This chapter answers the last and fourth sub-question: *what is the procedure of implementing team assessments, and how can this be improved?*

This answer is given by (10.1) identifying the assessment procedure, 10.2), providing factors that show if collaborative criteria add value, and (10.3) summing up actions that should be introduced in order to answer specific doubts concerning the team assessment. While the operational level of this chapter mostly serves the working field, the process and factors also add new knowledge to the existing literature.

10.1 Team assessment procedure

Since this is one of the first studies focussing on the use of team assessment for partner selection purposes, it aims to give an impression of the procedure that is being followed. The general procedure includes five steps. Steps number 1, 2 and 3 correspond to pre-assessment activities and take several months; step 4 is the team assessment itself; and step 5 is related to the scoring and completion of the assessment.

1

The process starts with the client's team considering the application of collaborative criteria and questioning which specific instrument. In most cases, this follows from the considerations regarding the project delivery model, project challenges and the ambition related to the client-contractor relationship (Section 6.6). Interviewee 1 and 6 both visited other infrastructure projects in this early phase to learn lessons from clients with similar challenges and objectives. Moreover, so-called market consultations are held to discuss the market approach and contract model. In the market consultations of case N1, N6 and F2, market parties made positive statements with regard to the team assessment and tender process. These reactions can help to convince the internal client.

2

In this stage, the client involves an team assessment provider. This firm is also called an occupational psychologist company or assessment firm. At the end of this stage, the client has to argue the use of its tender process and evaluation (sub) criteria to the market.

Most of the time, assessors are being involved when the client has decided to assess collaboration in a general way: '*could you help us shape this collaborative criteria and assessments*?" (Assessor 4). Although, clients already have an idea about the specific instruments and sub-criteria. This will only be stronger over the coming years as a result of the growing attention for this topic, and can also be a danger since easy copying and less customization will affect the value of team assessments (Section 8.5).

In some projects, clients already have published the sub-criteria (*competencies*) of the assessment in the tender guideline without having discussed these with the assessment provider. However, the experience and expertise of assessors is considered to be helpful in this stage of choosing appropriate sub-criteria and communicate the relevant considerations to the market (Section 7.1; Assessor 1, 2 and 4). Sub-criteria are a combination of the assessment provider's vision (competencies of collaborative behaviour) and the project related challenges (Section 7.2). It must be noted that the sub-criteria used by the assessors and the mix of exercises must comply with the project characteristics and foreseen project challenges in order to maximize the value and legitimacy of the team assessment (Section 3.2).

3

The precise course and content of the team assessment is designed in co-creation between the client's team and assessors (assessment provider). Epically those cases (*exercises*) that are business related need the expertise's of both parties. Such cases needs to address the challenges that are foreseen by the client's team in coming project, but the case also needs to be written down in such way that it is actually testing behaviour and stimulating authentic behaviour (Assessor 1 and 4). In six cases of this study, already-existing general exercises were used by the assessment provider (Section 7.1).

During step 2 or the beginning of step 3, the client publishes the tender guideline and corresponding notes. In this communication, it is considered to be crucial to manage expectations related to the team assessment (Section 8.1). "*Explain to bidders the rationale for the assessment centres, the issues that they will address and how they might prepare*" (Doyle and Jones, 2009). Moreover, the client's team members are presented.

The stage of preparing the team assessment is completed by practising at least (half) day. In this way, each member of the client's team is familiar with the exercises and briefed about how to behave (Section 8.2). Across the ten cases of this study, a vast majority of tenderers that have participated in an assessment took this very seriously and involved an (other) occupational psychologist company to prepare them for the assessment (Interviewee 1, 3, 5, 7, 8, 9, 10; Assessor 1, 2, 3, 4).

4

The team assessment is brought into practice during a full day per bidding party. The assessors will take the lead in explaining the day and team assessment since the client team is present as participators (Interviewee 4). During the day, the client's team and tenderer's team work on the prepared conflicting exercises while the assessors are observing all behaviour and "*writing a full notepad per day*" (Interviewee 1). In most cases, only the tenderer's team is assessed, while the client participates in the activities to provide the context and interaction that the contractor would face during realisation.

5

After the team-assessment, the team of assessors gets together to score the tenderer. The exact approach for this stage differs per team assessment provider (Section 7.2). Extensive assessments procedures usually follow a mathematical approach because more risks come in when you have an open to end discussion (Assessor 4). When the differences between the observations of the assessors are too big (statistically different), assessor have a mandatory discussion with all the evidence. Other assessors use a similar approach and include the client's team to make some minor changes (after doing the first analyse independently as an assessor) (Assessor 5).

To score each sub-criteria, assessors use a minimum of a five-point ordinal scale and give a score depending on the performance of the tenderer's team as a whole. The final score of each tenderer is the sum of the scores obtained from every sub-criteria (*competencies*). The team assessment results come with a report with detailed information about the observed bidders' behaviour and the argumentation of these results.

The assessment provider presents the results and their final report to the tenderers (or via the client). The tenderers receive their results and feedback from the client or in debriefing sessions with the assessor (Section 8.3). This helps to make the process (more) transparent and contributes to tenderer's learning process. It can be considered to bring the score to a civil-law notary where the client delivers the scoring of the other award criteria to be as independent as possible (Interviewee 4).

10.2 Determining when collaborative criteria add value

This paragraph focusses on seven factors that could help determining when construction projects benefit from the use of collaborative criteria in the tender process, and in particular team assessments. These factors are relevant for the first step (of Section 10.1).

10.1.1 Seven factors

This study used three different sources to propose factors that can help to determine if those criteria an de team assessments add value. These sources are previous conclusions of this study (Section 6.6 and Section 9.3), literature from the desk study, and in-house research of a large British public client (HS2 Ltd). Observations of HS2 Ltd were used as a starting point, and both validated and enhanced by the conclusions of this study. The result is shown in Table 12. The seven factors are all presented with its description and supporting observation. Some of them are closely related to one another.

1. Need for close collaboration (e.g. a joint project team) Projects where there is a need for close collaboration between the client and contractor (Supported by motive I; Walker et al, 2008; Kamminga, 2012)	**2. Complex project challenges in a dynamic environment** Projects where the challenges are complex and require on-going mutual problem solving (Supported by motives I*; Eriksson & Westerberg, 2000)	**3. Contractors interface with key stakeholders/public** Projects where the partner interfaces directly with the client's customers and/or key stakeholders (Identified in case U2; no literature found that supports this factor)
4. Soft competencies are critical Projects where leadership, team-working or other 'soft' competencies are critical success factors (Supported by motive III; Eriksson, 2010; 2014; Kamminga 2012)	**5. Cultural differences (values)** Projects where mismanaging organisational cultural differences could impact project delivery (Supported by motive IV and Mahalingam et al, 2015)	**6. Trust needs to be handed over** Projects where high levels of trust need to be handed over to the partner (Identified in case U2. (Supported by Laan et al , 2011; Kadefors, 2004)
7. Attitude of the client organisation Clients who are committed to look at themselves, select client's team members on collaborative attitude/behaviour, and to focus on collaboration post-contract (Supported by Section 8.5; 9.3; Davis, Edwards & Baccarini,2008)		

Table 12 | Factors that show when collaborative criteria add value

The first six factors of Table 12 find their origin in the documentation of HS2 Ltd (2017) on 'BAC' [Behavioural Assessment Centres] and show many similarities with the motives as presented in Section 6.6. HS2 Ltd is a public commissioning authority that has conducted internal research about the factors that argue the use of assessments (focusing on cultural fit and collaborative behaviour) in the tender process.

The seventh factor is based on the observations and conclusions of this study and looks to the potential value that collaborative criteria could add from a different angle. This new factor focusses on the (culture of the) client organisation and signature of the client's team members. After all, the implementation of collaborative criteria depends on the leadership and culture within the client organisation (Section 9.3.3). This study found that some clients did not look internally to themselves on a comparable level that was demanded from the bidding parties. Client organisations should have to take a closer look at whether people fit the task to make sure collaborative criteria, and team assessments in particular, can add value (8.5.2). Next to this focus on the client's team members, the new seventh factor also addresses the attitude of the client's organisation after awarding. Conducting a team assessment can be compared to undertaking a risk assessment: *"unless you mitigate the risks by post-contract award activity, it's nearly a waste of time"* (Assessor 4). This call for an end to end approach is relevant next to the other six factors because it determines if collaborative criteria add value in the overall project (also after awarding).

The six project-related factors of Table 12 should also be seen in the context of Section 9.3 that discussed the domain of projects that are fit for collaborative criteria. On the one hand, there is the first view following the belief that any construction project may, to a greater or lesser extent, benefit from these criteria. The factors of Table 12 apply to a majority of medium to large infrastructure projects. Section 10.1.2 will show this by presenting two example projects. On the other hand, there is a second view stating that projects with a dynamic environment and intensive collaboration clearly benefit more from the focus on collaboration in the tender criteria. This view matches the proposed factors of Table 12 as well.

Table 12 can be seen as a tool which combines several observations of this study. Both motives (Section 6.6) and interpretation (Section 9.3) are used to strengthen (and validate) these factors. However, the factors need to be validated (again) in future research since the set-up of this study was not designed to do so.

10.1.2 The relevance of having these factors

This study embraced the idea of designing factors that could help clients in their process of considering collaborative criteria. The (six project-related) factors show that different kind of projects could make use of collaborative criteria and facilitate the strategy discussions within the client organisation.

Several interviewees expressed their concern that 'the success and increased intention' for team assessments could lead to too much copying and too little customisation (Interviewee 6 and 9). The attitude '*everyone is doing this so we should do this as well*' (Interviewee 8) results in incorrect experiences, missing out the potential, and eventually even appeals from bidding parties. By contrast, a client has to consider and argue the value and implementation of team assessments every single project again (Interviewee 6 and 9; Section 8.5). Table 12, and in particular the added seventh factor, aims to start that discussion.

There is no formula belonging to these seven factors that tells someone to apply collaborative criteria above a specific number. It is considered to be the combination of those seven factors that give an impression of the potential extent and weight of collaborative criteria. When more factors apply to a large extent, the (set of) collaborative criterion/criteria should be given a higher weight (Interviewee 9). Figure 15 aims to provide an impression of the broad applicability of those criteria with two example projects.

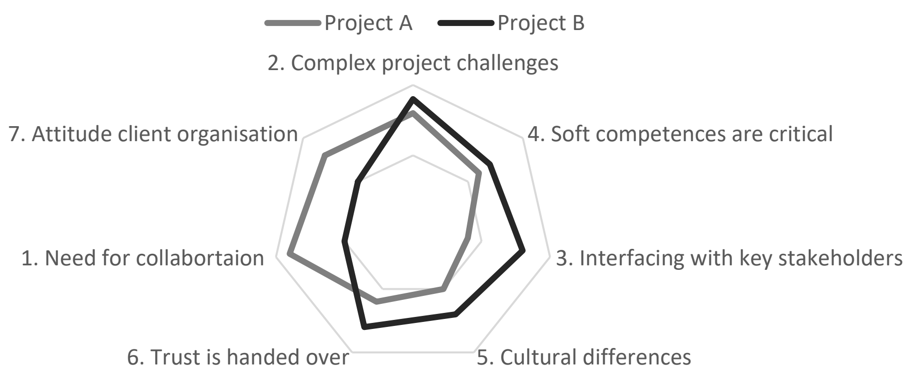


Figure 15 | Example of two projects scoring on factors that show the value of collaborative criteria

The cases of this study included a design team model on the one hand (N6) and a complex mega-project that includes several large civil works sub-parts (U2) on the other hand. Figure 15 shows an example of the application of these seven factors (as presented in Table 12) for two example projects: a design team contract like case N6 (project A) and a larger and more traditional Engineer & Build contract (project B). It can be noticed that the projects score differently in the domain of these factors.

Despite the difference project characteristics, both example projects could benefit from the use of collaborative criteria. These seven factors can help to argue their added value in the tender guideline.

10.3 Improving the team assessment procedure

This study identified several question marks concerning the implementation of team assessments (Section 4.3; Section . Those doubts are answered as best as possible with the observations of this study.

In Section 4.3, several views on team assessments were collected. One of them was the statement of the Procurement Advisory Group (PAG) of the UK based Institution of Civil Engineers (ICE, 2018) of two years ago*: "The assessments can be seen as beneficial, but they require substantial planning and preparation to develop the transparency needed for public procurement, which can be time-consuming and expensive in terms of client and bidder resources. Team assessments should be kept simple until the value of more complex models and techniques is properly established, …"* (ICE, 2018). This study did not design or validate more complex techniques. The domain of social sciences is psychological science is leading in doing so. However, this study found several measures that are seen as a relevant step of improving the team assessment instrument.

Some of these measures have already become the standard for certain assessors, while others are observations in just one of the countries that have been studied. Furthermore, it must be noted that there are plenty of area's (i.a. the specific exercises) that have no best answer but are always project related (Section 8.5). Table 13 sums up different measures and its origin in this study.

Potential weaknesses	Considered measures (improvements)
Bidders don't appear as they are	1) Introduce certain conflict of interests that raise authentic behaviour (Section 7.3; 8.3); 2) Make sure to apply a mix of exercises during the team assessment, which should at least last half a day, to ensure authentic behaviour (Section 7.3, 8.5);
Not being able to ensure a level playing field	1) Take the time (a minimum of half a day) for the preparation (including a 'dry run') before the first real team assessment day (Section 7.5; 8.2); 2) Perform the team assessment in the first stage of the partner-selection process when there is no acquaintance yet between both parties (Section 7.1); 3) Involve observers and lawyers (within a compliance team) that have a role in observing and i.a. giving feedback to the client's team about their behaviour (Section 7.5; 9.2); 4) Take the team assessment very seriously as a client team in a whole: speak out to each other on undesirable behaviour or being less committed (Interviewee 8, 4);
Being challenged afterwards	1) Manage expectations as best as possible on the front via the tender guideline and even a tender start/follow-up with al bidding parties and the assessors (Section 8.1; 8.3; 10.1); 2) Be as transparent as possible about the scoring of the team assessment: consider to share the review to bidding parties and hand them over in real-life (Section 7.5; 10.1)
Arouse suspicion about neutrality of scoring	1) Be clear about the procedure, deliver the final scoring to the notary. In that case the client is not familiar with the scoring and provides his scoring there as well (Interviewee 4) 2) Make clear which regulations and guidelines are being followed (Section 7.5)
Assessing people that will leave soon on the project	1) Stimulate bidding parties to look for team members that are intrinsically motivated to be on the project by explaining the project goals and client's commitment (Interviewee 10); 2) Apply mini-assessments or interviews with the same competencies of the team assessment when key figures change later on (Interviewee 7); 3) Consider culture alignment instruments (Section 7.1) to make it a wider review of the tenderer's organisation and asking tenderers to explain their commitment (Interviewee 10)

Table 13 | Concrete improvements that can be added to the team assessment procedure

10.4 Concluding: implementing team assessments

The observations of this study were used to define and improve the steps of the team assessment procedure.

SQ 4 | What is the procedure of implementing team assessments, and how can this be improved?

The observed steps of the assessment procedure, including the suggested improvements, are displayed in Figure 16. The Dutch perspective was the starting point of deriving these improvements.

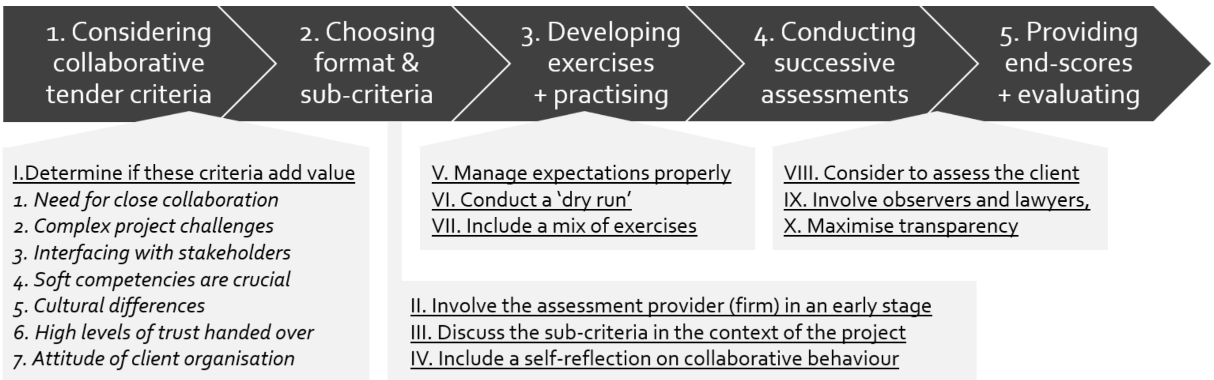


Figure 16 | The (improved) team assessment procedure

PART VI

Part VI

In this last part, the research methods and its findings are discussed. This will be followed by the final conclusion of this study. Chapter 11 and 12 must be viewed in its entirety considering the fact that the identified limitations of this research affect the conclusions and recommendations of a research. The final conclusion answers the main research question and will make use of the answers to the four sub-questions.

11. Discussion

This chapter includes (11.1) a discussion on the findings of this study, (11.2) a reflection on the goals of this study and (11.3) a description of the limitations that can exist due to constraints and decisions concerning the research method. These research limitations are discussed with the belief that writing about them towards the end of this study strengthens the conclusions and recommendations. This chapter can put them into the right academic perspective. Discussing the findings has quite the similar objective. It also aims to bring several observations of this study into perspective of other studies.

11.1 Discussing the findings

Based on the answers of the sub-questions and interpretation in Section 9.1, this chapter discusses these findings from new perspectives.

Selection of team members within the construction industry

Clients reflecting on the collaborative criteria see a particular value in the 'signal function' of these criteria to the market. This signal is think of this human factor, which is so important for project success. The clients of case N1, N3 and figured out that one or more tenderers considered several team compositions for the team-assessment and project because their competencies were that important for being selected. Even organisational psychologists were involved, which is irregular for the contractor culture.

This attention to the selection of team member is not new. It is a whole domain of studies and statements on the importance of the right team members selection, on both the owner as supplier's side. Since organisational issues have become more critical drivers in delivering projects, collaborative behaviour is one of the factors that is taken into account in the selection of the members of a project team (Slevin & Pinto 2004). As already set out by Walker in the early nineties, the team selection and the considerations regarding their formation are an important determinant for the course of the project and recognized across several domains (Walker, 1996). New research could focus on the interface of individual selection methods and the connection with team assessments (outcomes) in the construction industry.

Several researchers (Antoniadis, 2012; Eriksson, 2010) state that since soft criteria have a more dominant place within the tender phase of construction projects, there should also be new tools available. We should focus on mechanisms that go one step further than just a team-building activity or one individual test. Antoniadis mentions that the focus on collaborative behaviour and individual development should start in the early days of someone's career (within the organisation). Concluding, developing individual profiles from the start could support team formation and assessments during the tender phase (Antionadis, 2012).

The culture of the client organisation

Client F figured out which cultural and which behaviours they wanted, but they did not look internally to themselves on a comparable level (Interviewee 8). Some contractors ended up with a more traditional client team that was not particularly collaborative. Interviewee 3 also reflected on their own client team and those of comparable projects: "*Within our organisation, we may also have to take a closer look at whether people fit the task. On our side, that may be done too much based on capacity and availability*." In this way the attention for collaborative competencies is not that dominant as client B demands from the market via their selection and award criteria (Interviewee 3). The conclusion that the client has still to take a step here is also seen by interviewee 1. In other words, the culture of the client may counteract the use of collaborative criteria. Interviewee 5 notices that its client organisation is extremely "blue" with the desire to standardize. The uncertainty and freedom of this approach, the philosophy, is also difficult for some. Assessor 4 sees the positive aspects of these criteria more as a change of culture than a change of the working method. This realisation could make it even more challenging to go through this transition as a governmental organisation. This all is not about using an extra set of rules (Interviewee 5).

Interviewee 5 and 3 both needed the label of being a pilot project to experiment with new procurement procedures and collaborative criteria. This observation – it's challenging to change procedures in the current project environment – can also be found in the literature.

Bradley & Parker (2000) used a survey of around 1000 employees and found that public agencies in Queensland (Australia) have a culture "which reflects a focus on rules and regulations, with little flexibility". A few years later, Davis, Edwards & Baccarini (2008) looked into the decision-making process of procurement methods by public clients in Australia. They found that more traditional methods are still preferred, even though relationship-based contracting is known to optimize the project outcome: "*An embedded culture of uncertainty avoidance invariably*" (Davis et al, 2008). The researchers used several focus groups and discovered that the culture within the organisation had a strong effect on this behaviour (the decisions regarding the procurement strategy). The researchers used Hofstede's dimensions and developed the hypothesis that the observed culture is about '*uncertainty avoidance' (*Edwards & Baccarini, 2008). This embedded culture hinders the implementation of collaborative criteria, and above all the procurement methods and procedures that underlie this intention of using a team assessment. It is not found why the European public clients within the construction domain have a similar embedded culture. Another identified culture within the construction industry in literature is that the current management style is still mainly engineer-focussed (Pries et al., 2004). This culture makes it harder to implement changes in culture, as interviewee 5 describes the implementation of collaborative criteria. However, it must be noted that most of this research is conducted more than a decade ago. New research could focus on the client's cultural and institutional barriers (laws and traditional procurement procedures) for implementing changes in the domain of the collaborative relationship between the owner and contractors.

Also, studies in Europe write about the role of the client organisation in changing procurement procedures construction industry. Several reports by the Swedish public transport agency mentioned this role even as a "driver of change and innovation" (SOU, 2002). Their finding is supported by the belief that the owner of construction projects has a crucial role in their chose for the delivery method and governance form of the project (Eriksson and Laan, 2007). Following this line of reasoning, change can only come from clients and not (only) by market parties. This is a relevant observation: innovative and effective building techniques can be put forward to the market (for example, via early contractor involvement), but new procurement techniques and instruments that enhance collaboration need to be supported by the public client.

11.2 Discussing the research objective

Following the discussion of the findings, it's also relevant to discuss the achievement of the (higher) objective of this research based on the observations and interpretation of this study.

The research goal proposing an assessment procedure

The discussion on the previous page about the desire to standardise within public organisations is also relevant for the third research goal of this study. This study aimed to proposes an assessment procedure for the partner-selection process to give tools to clients who would like to apply collaborative criteria. The desire to make them as concrete as possible (in order to enlarge the relevance of this study) could also be seen as the same reflex that is observed within several client organisations.

Multiple interviewees express that the implementation of the collaborative relationship and collaborative criteria depends on the project, the contract and the dynamics (Interviewee 6 and 9). According to interviewee 6, you have to customise criteria and assessments for each project. The most concrete deliverables of this study (Section 10.1; 10.2; 10.3) aim to provide concrete tools for client organisations (some of them still need to be validated), without stating which specific application is best.

The foremost objective of this research: improving the relationship already during the tender phase

This study aimed to contribute to the higher goal of improving (not automatically increasing) the client-contractor relationship during the tender phase. This goal is part of the Dutch 'Marktvisie' (2016).

In line with the conclusion of De Jager (2016), it must be noted that the procurement phase of a construction project will always be in a certain competitive environment, despite the measures taken. The procedure involves several bidding parties that compete with each other in a regulated context (De Jager, 2016; Graells, 2015). This makes it (more) hard to focus primarily on improving collaboration during the tender phase. In other words, the improvement of collaboration (as related to the third identified motive in chapter 6) can only go as far in public procurement.

Nevertheless, procurement regulations and procedures are changing over the years. The introduction and obligation of EMAT-criteria but also the increasing use of the innovative partnership and dialogue rounds can be seen as a development in the direction of this domain. This makes it possible to focus more and more on the relationship during the tender phase. When introducing the project alliance in Finland, Väylä also maximized the procurement possibilities as set by the European law (Lahdenperä, 2002). De Jager (2016) already stated that these opportunities should be utilized instead of focussing on everything that cannot be achieved. This attitude is crucial for implementing collaborative criteria.

11.3 Limitations of this research

The chosen scope and research method automatically result in some limitations. Each study has its limitations, and this research is no exception. Since those limitations have an impact on the following conclusions and recommendations (Section 12.2), those are taken into account and weighted previously.

The quantitative approach

Most extensive data (observations) of this research were gathered via interviews. It must be noted that the follow-up questions during the interviews were always resulting from previous observations by the researcher. Because the interviewee and the researcher were not restricted by the predefined questions, the conversation and sub-topics developed while the interview was going on. In this way, the semi-structured interview method had its limitations. Section 5.4 already described some precautionary measures on this point – based on the weaknesses identified by Yin (2003).

Several very relevant observations of this research – some of them discussed in the previous paragraph of this chapter – were the direct result of this semi-structured interview with follow-up questions. In other words, reducing this potential issue of bias would have had a direct effect on the value of this study. Some findings could simply not be gathered with a list of predefined questions. Nevertheless, it is essential to keep the influence of the researchers perspective in mind.

Representation of the cases

The way this research collected its data is already discussed, but the selection and representation of the cases is important for the reliability in terms of the applicability of the recommendations. This study did not select the cases on the basis of their contract and project delivery method, but the cases show some similarity in terms of early contractor involvement (Section 6.5). In other words, the cases that were being studied do not represent an average (more traditional) project within the infrastructure domain. The recommendations of this research could be relevant for a construction with a more traditional contract; however, the motives and experiences identified do not automatically apply to for these projects.

Studies towards relatively new mechanisms and methods hardly use representative samples but first focus on the phenomenon in particular. This study is very focused on the projects and clients who already have implemented these partner-selection assessments since there is very limited literature available on this topic. By choosing for cases that have already implemented these criteria, it is not found why clients refrain from the use of collaborative criteria. This could be the objective and scope of a new and broader research.

The case study design

The design of this research, including the case study, was primarily focussed on gathering a broad set of experiences - from different projects, different cases. This resulted in a meagre amount of interviewees per case. Since conducting more interviews is highly time-consuming, the researcher had to find the optimal balance between the input per case and the number of cases, which will lead to a larger set of experiences. For almost every case, just one interviewee was conducted within the client organisation, and the intended representation of two interviewees per case has not been realized.

This limitation must be nuanced by different measures. Several answers by one of the interviewed respondents were supported by one of the interviewed assessors. The list in Table 14 shows the connection between the assessors involved and ten different cases [projects]. A major part of the assessor's interview focussed on general aspects of the instruments; however, some of their examples were based on a procurement exercise that's also a case in this study. In this way, the assessor interviews can also be seen as a second view on the projects that have been selected and just had one respondent.

Case	Client	Interviewee(s)	Assessor(s)
CASE N1	Client A	Two interviewees (Interviewee 1 and 2)	
CASE N2	Client B	One interviewee (Interviewee 3)	Assessor 1 was involved in this assessment
CASE N3	Client C	One interviewee (Interviewee 4)	Assessor 2 was involved in this assessment
CASE N4	Client B	One interviewee (Interviewee 5)	Assessor 1 was involved in this assessment
CASE N5	Client D	One interviewee (Interviewee 6)	
CASE N6	Client E	Both contract and project manager were present during the interview (interviewee 7)	Assessor 3 was involved in this assessment
CASE U1	Client D	Two interviewees (Interviewee 8 and 9)	Assessor 4 was involved in this assessment
CASE U2	Client F	Two interviewees (Interviewee 8 and 9)	Assessor 4 was involved in this assessment
CASE F1	Client G	One interviewee (Interviewee 10)	
CASE F2	Client H	One interview (Interviewee 11), Input was checked/validated by one other colleague	

Table 14 | The interviewees linked to the cases

Within the chosen balance of the number of respondents versus cases, the potential issues have been answered as best as possible. Nevertheless, this case study design still brings certain limitations. New research projects should also include observations from other respondents from the client's team.

Reliability of the results

Interviewing just one representative of the client's project organisation also raises questions regarding the trustworthiness and reliability of the observations. To ensure reliability in a qualitative study, the examination of trustworthiness is crucial (Golafshani, 2003). For example, interviewing a relative outsider within the client's project team – concerning the partner-selection assessment – could result in less reliability. In order to collect the most reliably observations, the researcher has spent much time in finding the best respondent (team member) per case. Except for interviewee 11 who had to include a colleague to answer all questions, every respondent confirmed that the researcher contacted the right person within the client team. For case N1, N6, U1, U2 and F2, an extra respondent was involved to get the full perspective and answer all questions correctly. In a majority of the cases it was also confirmed by one of the assessors that the researcher contacted the right representative from the client-side. Those responses strengthen the belief that a new study with the same cases will – probably - end up with the same respondents. Following this line of reasoning, the same respondents should also lead to the same observations. The consistency of the respondents was checked by a technique known as respondent validation. The initial responses and observations were shared via an interview report afterwards.

The outline as set out above assumes that new studies will select the same cases (with similar scoping and case selection criteria). For the Dutch cases, this may be expected since there are just around ten larger infrastructure projects that have used this partner-selection assessment. In Finland and the United Kingdom, the number of relevant projects is a lot bigger compared to the two selected cases per country. By contacting and involving experienced professionals in this field, the reliability of this part of the study is strengthened as best as possible. Two of these experienced professionals are assessor 1 (one of the leading Dutch assessors), assessor 4 (a highly respected assessor from the United Kingdom who is leading in this field for two decades) and interviewee 10 (a deputy director for a large public client in Finland).

The client's perspective vs the contractor's perspective

This study was conducted from the perspective of the client, this could be considered as a limitation of this study. The objective of this research is to provide an overview of the motives and experiences from the client's side, as the procurement procedure is developed by this party. All procurement criteria are chosen and implemented by a commission party, not the contractor. When it's the clients perspective that a project has benefited from the use of collaborative criteria, that experience has value on its own. This line of reasoning was already given in Section 2.4. Nevertheless, it must be noticed that a collaborative relationship is always about the interaction between both parties. Similar interviewees with tender and project managers from a contractor's perspective will lead to different results and recommendations for the application of these criteria in specific. Since the client's perspective is seen as the best starting point, new research could focus on the experiences of both bidding parties that are awarded the contract and have lost. This is also supported by several interviewees that expressed the importance of developing and improving instruments with the market, just like the relevance of interacting via market consultations.

12. Conclusion

This last chapter summarises the answers to the different sub-questions and provides a final conclusion. In this way, the main research question of this study is answered. A second aspect of summing up conclusions is making recommendations for both the working field and research field. The (12.1) final conclusion is addressed first and followed by (12.2) recommendations.

12.1 Final conclusion

Construction-related literature is not clear about the value and formats of assessing the bidders' collaborative behaviour. This explorative study aimed to fill this gap by identifying the client's motives, the formats and the experiences of conducting team assessments in the tender process. Based on these observations, a general assessment procedure is presented and improved.

Most sub-conclusions of this study are derived via one of the ten cases across the Netherlands (six cases), as well as the United Kingdom and Finland (two cases each). Other results are not related to one of the cases (projects) but directly linked to the view of one of the sixteen interviewees, including five assessors.

SQ1 | What are the client's motives to apply collaborative criteria in the tender process?

This is the first study to identify the client's motives to apply collaborative criteria for partner selection in infrastructure projects. These motivations is highly relevant since arguing them is needed according to procurement regulations. For this first sub-question, it was not considered relevant which specific instrument was used. The focus lay on identifying the drivers of applying collaborative criteria in general.

In this study, three different motives were mentioned in at least five (out or ten) or more cases:

i. Applying collaborative criteria because the relational contract brings a joint team and/or intensive collaboration in the design and/or realization phase; (mentioned in 9 out of 10 cases);
ii. Applying collaborative criteria considering the fact that criteria focusing on price and content were possible to a lesser extent, so soft criteria came in; (mentioned in 6 out of 10 cases);
iii. Applying collaborative criteria to lay the foundation of the collaborative relationship in the tender phase and emphasize on the importance of soft skills; (mentioned in 5 out of 10 cases);

It must be noted that interviewees could not choose from a predefined list, and the three motives do not exclude each other. For example, the first and second motive are closely related and could be linked to early contractor involvement and relational contracts like the design team and project alliance model. This study did not select the cases on the basis of their contract and project delivery method, but nine of them applied some form of early contractor involvement (ECI). The motives to choose for collaborative criteria, as well as the increasing use, can be linked to the development of design teams in the Netherlands, NEC3 in the United Kingdom and project alliances in Finland (Kamminga, 2012; Lahdenperä, 2012; Eriksson, 2010).

SQ2 | What instruments and formats are used to assess the contractor's collaborative behaviour?

This study identified a diverse mix of instruments that are being used to evaluate the contractor's cultural fit and capability to collaborate with the client. Next to an assessment, which was conducted nine out of ten cases with the client's team in the room, several cases also made use of cultural alignment instruments; like written documents and (panel) interviews, and individual matching instruments like an EQ-i test.

All identified assessment instruments involved in any time independent assessors to score the bidding parties on a predefined list of (soft) competencies. Moreover, guidelines and a robust procedure were used to ensure a level playing field and conduct the assessment according to the procurement regulations. The number of team assessment participants varied from 4 until 25 individuals per side and shows the variety of formats that are being used across the cases. A majority of the cases applied both business (project) related cases as well as (a lower number of) non-busines related exercises. It was found that the variety of applications relates to the client's objectives and the different definitions of the tender sub-criteria. Team assessments that focus on the cultural fit between client and tenderer bring different competencies on the scoring card compared to focussing on the tenderer's alliance capability/behaviour. This study also reveals that the team assessment instrument has some corresponding cornerstones in the Netherlands, the United Kingdom and Finland, but the specific application (participants, content, duration, scoring) differs per country, project and last but not least, the assessors that are being involved

A majority of the client's representatives and assessors that were involved argued their use of a team assessment with both the tenderer's and client's key figures since the participants from the client-side provide the context and interaction that the contractor would face during realisation.

SQ3 | What are the client's experiences from conducting team assessments?

The client's experiences from conducting team assessments forms the answer to the third sub-question. Clients (and assessors) see the team assessment as a valuable instrument in the tender process that has a strong signalling function. Within cases it was noticed that bidding parties are looking to collaboration and their team-selection in a different way. More attention is paid by bidding parties to soft elements compared to more traditional tender criteria. Moreover, the dialogue meetings and project start-up after the team assessment benefit from the interaction and reflection during the assessment.

Despite some remarks and prerequisites for further use, all clients are predominantly positive about the specific course of their team assessments. However, clients who expressed the third motive of improving collaboration via collaborative criteria, see all instruments as just one of the ways to do so. This study found seven more detailed experiences of the specific course of the team assessments:

- Expectation – Bidders need to know the rationale for the criteria and what to expect at the assessment;
- Preparation – A dry run and practising by the client's team is crucial to ensure a level playing field;
- Implementation – Usually a full assessment day and a mix of exercises is required to score properly;
- Distinction – Each tenderer's team brings a different interplay and outcome, despite the same case;
- Evaluation – Putting effort in the review of the scoring is greatly appreciated and serves transparency;
- Interaction – The following competitive dialogue and interaction benefit from the assessment interplay;
- Collaboration – Ongoing attention is needed to utilise the assessment focus on collaborative behaviour.

SQ4 | What is the process of implementing a team assessment, and how can this be improved?

The last sub-question focussed on combining the lessons learned by deriving the team assessment procedure, and propose several improvements to some of those steps. The whole procedure starts with determining the added value of collaborative criteria (the left box of Figure 17). Seven factors are seen relevant for determining this added value. Next to six project-related factors, this study found that the attitude of the client's organisation is crucial for the potential value that collaborative criteria add.

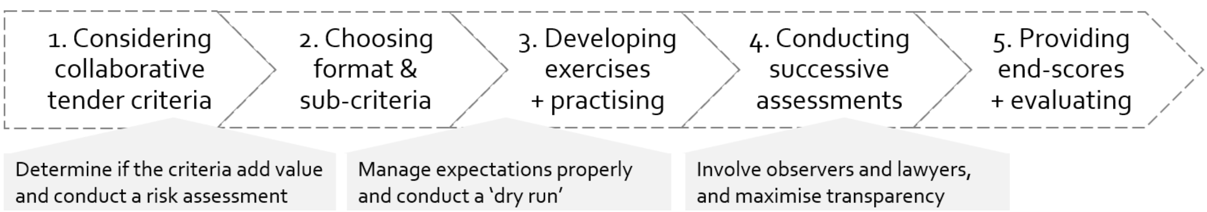


Figure 17 | Steps of implementing collaborative criteria

The second step focusses on the second sub-question of this study: determine the specific format and sub-criteria. Step 3 corresponds to pre-assessment activities and takes a few months, while step 4 is the team assessment itself. Step 5 is related to the scoring and completion of the assessment. Some improvements to those steps are shown at the bottom of Figure 17. Clients could apply this procedure to assess bidders based on their collaborative behaviour towards forming high performing client-contractor teams.

Answering the main research question

The main question aims to give a total impression of the use of team assessments today:

> **What are the client's motives and experiences of assessments that evaluate the contractor's collaborative behaviour as part of the tender process in infrastructure projects?**

This study reveals that the (increasing) application of collaborative criteria can be linked to the implementation of relational contracting and early contractor involvement (including a joint client-contractor) as well as the objective of clients to focus on and use the tender phase to start working on the collaborative relationship between the client and future contractor. Clients prefer the use of (among other instruments) a team assessment to assess tenderers on collaboration since these real-life assessments are able to assess authentic behaviour of people who will be on the project later on instead of only written bids.

Despite several remarks and prerequisites for further use, a substantial majority of the interviewees underlines the value of conducting team assessments in the tender phase. However, the assessment itself is no miracle cure. Other instruments can lead to (some of) the same goals, and ongoing attention is needed to enhance collaboration and utilise the benefits of the assessment. The increasing use of these instruments also brings some concerns. The attitude of the client organisation is crucial for a proper implementation and it could result in copying behaviour without addressing the project challenges in the assessment and sub-criteria. Such incorrect applications could harm the overall positive experience. In short, when the assessment is done correctly, it is considered to have value during both tender and award phase.

Soft skills of key people will be more and more in front in the infrastructure sector. The team assessment has proven to enable the client to make a more informed choice for a tenderer on its collaborative behaviour. It also puts 'collaboration' high on the agenda of all parties right from the start of the tender process. In this way, team assessments, also called workshops, can fulfil the potential of being one of the instruments to facilitate high performing client-contractor teams and put more emphasis on soft parameters.

12.2 Recommendations

The recommendations are divided into two parts: those for the working field and those for researchers.

Recommendations for public clients

The four most important recommendations for the working field focus on operational aspects of conducting team assessments for partner selection purposes. However, one the most remarkable observations of this study consists of a higher level warning towards public clients within the infrastructure domain.

The increasing attention (could) lead to the attitude of 'everyone is doing this so we should do this as well', while the specific application of collaborative criteria should depend on the project challenges and the client's view on collaboration. According to one of the assessors, the team assessment should not be a blueprint for something that does not lend itself to that. Besides, this study reveals the importance of the client's attitude and leadership within the client's team to utilise the added value of these instruments.

This message of reflection together with the four operation recommendations below are in particular relevant for public clients that are new with these kind of instruments. This also applies to advisory firms like AT Osborne that are supporting clients in the stage of developing a procurement strategy.

1) Discuss and determine the added value of a team assessment (or other collaborative criterial) for the project in a comprehensive way. Use the assessor's expertise and involve them in an early stage.

2) Make sure to apply a mix of exercises and take at least half a day for each assessment to ensure authentic behaviour. Perform the team assessments in the first stage of the partner-selection process when there is no acquaintance yet between both parties. Moreover, make clear which regulations and guidelines are being followed by the client's team, the observers and the assessors. The content and focus of the business related exercises/cases could be derived by conducting a risk assessment for the project upfront.

3) Take a minimum of half a day for the preparation and an intensive dry run before the first real team assessment. Designate one or more observers (within a compliance team) that have i.a. the role of giving feedback to the client's team about their behaviour (to ensure a level playing field).

4) Manage expectations upfront via proper communication (the tender guideline) and a potential tender start-up with. Following the assessment, be as transparent as possible about the reviews and even consider (together with the assessor) to share the full reports and give feedback during face to face meetings.

Recommendations for future research

The recommendations for future research are based on the interpretation and discussed limitations.

- <u>Including the tenderer's perspective and experience</u>. Interviewees with tenderer's team members could lead to new valuable recommendations since it must be noticed that a collaborative relationship is always about the interaction between both parties. Future research could focus on several in-depth cases and collect the experiences of all parties involved: the client's team, the tenderer that has been awarded the contract, and bidding parties that have lost.
- <u>Linking the success factors of projects to the sub-criteria that are being used for partner-selection.</u> The most interesting relation of this domain is the link between success factors of a project and the sub-criteria that are used to score tenderers during the assessment. This study showed that both assessors and clients have different views on what competencies are the most important for partner selection purposes. Future research could focus on the identified success factors of collaboration within the construction industry and appropriate sub-criteria for team assessments.
- <u>Improving and designing new instruments from a psychological science perspective</u>. The domain of this study lies on the interface of psychological science and procurement practises. However, this thesis was conducted from the perspective of procurement and mainly focussed on public clients. Future research could focus more on improving the assessment format or introducing tools from other domains to the construction sector. This includes questions like 'what type of tenderer and team fits the client organisation best?'. This requires more interaction between the domain of social sciences and project and procurement management over the coming years.

Literature list

Andersen, B, Hosseini, A., Lædre, O., Torp, O., Olsson, N., & Lohne, J. (2016). Selection criteria for delivery methods for infrastructure projects. Procedia-Social and Behavioral Sciences, 226, 260-268.

Antoniadis, D. (2012). Complexity and the process of selecting project team members. Journal for the Advancement of Performance Information & Value, 4(1).

Arlbjørn, J. S., & Freytag, P. V. (2012). Public procurement vs private purchasing. International Journal of Public Sector Management.

Bloemendaal, R., & Van der Geest, R. (2011). Het experiment A2 hooggelegen. The Netherlands, Veerkracht-Rijkswaterstaat.

Boukendour, S., & Hughes, W. (2014). Collaborative incentive contracts: stimulating competitive behaviour without competition. Construction Management and Economics, 32(3), 279-289.

Bosch-Sijtsema, P. M., & Postma, T. J. (2009). Cooperative innovation projects: Capabilities and governance mechanisms. Journal of Product Innovation Management, 26(1), 58-70.

Bochenek, J. (2014). The contractor selection criteria in open and restricted procedures in public sector in selected EU countries. Procedia Engineering, 85(0), 69-74.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.

Bresnen, M., & Marshall, N. (2000). Building partnerships: case studies of clientcontractor collaboration in the UK construction industry. Construction management and economics, 18(7), 819-832.

Bresnen, M., Goussevskaia, A., & Swan, J. (2005). Organizational routines, situated learning and processes of change in project-based organizations. Project Management Journal, 36(3), 27-41.

Brinkman, P, 202. (Masterscriptie) Het twee fasen proces. De aanbestedingsrechtelijke houdbaarheid van het twee-fasen-proces zoals voorgesteld door Rijkswaterstaat

Chao-Duivis, M. A. B., et al. (2012). A Practical Guide to Dutch Building Contracts. The Hague, ibr.

Chao-Duivis, M. A. B., & Wamelink, J. W. F. (2013). Juridische aspecten van ketensamenwerking: Naar een multidisciplinaire benadering (Vol. 41). Instituut voor Bouwrecht.

Cheung, S. O., Ng, T. S., Wong, S. P., & Suen, H. C. (2003). Behavioral aspects in construction partnering. International journal of project management, 21(5), 333-343.

Cohen, D. J., & Crabtree, B. F. (2008). Evaluative criteria for qualitative research in health care: controversies and recommendations. The Annals of Family Medicine, 6(4), 331-339.

Commonwealth Department of Infrastructure and Regional Development, 2006). National Alliance. Contracting Guidelines Guide to Alliance Contracting

Cobouw (2016). 'Samenwerken doe je niet alleen' published on 5 december 2016

Cobouw. (2019a). Aantal mislukte aanbestedingen stijgt_ Wie is de dupe_ - Cobouw.nl

Cobouw. (2019b). Beter borgen van samenwerking in nieuw bouwteam _ Duurzaam Gebouwd

Cobouw. (2019c). Stapelcontracten rukken op: Clusterverbod wassen neus

De Koning, J. (2018). Het nieuwe bouwteam. Retrieved September 16, 2018

Dewberry, Chris & Hayes, Alan & Sarhan, Saad. (2018). Behavioural Assessments in Construction Procurement: A Bandwagon of Institutional Waste?.

Doyle, J and Jones, R (2009) The £4.4bn assessment centre. “CPO Agenda, Spring”, 44-50.

Dusseldorp, T., Put, Van der Put, J. & Rupert, J. (2012). Naar een gezonde spanning in projectenland (pp.17-33). In: Van Muijen, J., Rupert, J. & Tours, H. (Red). Spanningen in en rondom organisaties. Kluwer: Deventer.

Drews, F (2017). Designing Meta-Organisations: An Empirical Study of Boundary Setting in Large Infrastructure Projects. A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities.

Dreschler, M. (2009). Fair competition: How to apply the ‘Economically Most Advantageous Tender’(EMAT) award mechanism in the Dutch construction industry.

Elemans, B. (2016) (Masterscriptie) Developing a Project Alliance Suitability Scan for Dutch Infrastructure Projects. Based on experiences from The Netherlands and Australia

Engebø, A., Skatvedt, Å., & Torp, O. (2019). Soft elements in collaborative project delivery methods. In Research Conference Proceedings. IGLC.

Eriksson, P. E. & Westerberg, M. (2010). Effects of Cooperative Procurement Procedures on Construction Project Performance: A Conceptual Framework. International Journal of Project Management, 29 (2), 197-208.

Eriksson, P. E. (2015). Partnering in engineering projects: Four dimensions of supply chain integration. Journal of Purchasing and Supply Management, 21(1), 38-50.

Eriksson, P. E. (2010). Partnering: what is it, when should it be used, and how should it be implemented?. Construction management and economics, 28(9), 905-917.

Eriksson, P. E. (2008). Procurement effects on coopetition in client-contractor relationships. Journal of construction Engineering and Management, 134(2), 103-111.

Eriksson, P. E., & Laan, A. (2007). Procurement effects on trust and control in client-contractor relationships. Engineering, Construction and Architectural Management.

European Commission 2017, Efficient and professional public procurement.

Financial Times (2016), Bidders for HS2 contracts tested for their best behaviour (received on 10 January 2020 via premium access)

Florence Yean Yng Ling, Shi Ying Ong, Yongjian Ke, Shouqing Wang, Patrick Zou. 2014. Drivers and barriers to adopting relational contracting practices in public projects: Comparative study of Beijing and Sydney. International Journal of Project Management 32:2, 275-285.

Flyvbjerg, B (2006), Five Misunderstandings About Case-Study Research, Aalborg University

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. The qualitative report, 8(4), 597-607.

Hertogh, M. J. C. M. and E. Westerverld (2010). "Playing with Complexity: Management and organisation of large infrastructure projects."

Hertogh, M. J. C. M. et al (2017). Kennistraject-Gaasperdammertunnel Deel 1: 2014-2016. Rijkswaterstaat, IXAS, Nederlands kenniscentrum voor ondergronds bouwen&ondergronds ruimtegebruik

HM Tresuary, 2014 Improving Infrastructure Delivery: Alliancing Best Practice in Infrastructure Delivery', HM Treasury, HM Government

HS2 Ltd (2017) BAC presentation, received on 1th of May 2020

IBR, Instituut voor Bouwrecht, (2016). the Dutch Proportionality Guide, vertaald document

Infrastructure and Projects Authority (2014). Improving Infrastructure Delivery: Project Initiation Routemap. Organisational Design and Development Module. October 2014.

ISO. (2017). 44001:2017 - Collaborative business relationship

Jefferies, M., Brewer, G., Rowlinson, S., Cheung, Y. K. F., & Satchell, A. (2006). Project alliances in the Australian construction industry: a case study of a water treatment project.

Jensen, C (2015) "Network Rail's Wessex Alliance selection process: choosing for collaboration"

Kadefors, A. (2004). Trust in project relationships—inside the black box. International Journal of project management, 22(3), 175-182.

Kamminga, P. (2012). Aanbesteding Een Goede Basis Voor Samenwerking? Een Analyse Van Aanbestedingsregelgeving Vanuit Samenwerkingsperspectief.

Kamminga, P. (2009), Welke Lessen Zijn Er Te Trekken Uit Empirische Studies Naar Succes En Falen Van Alliancing En Partnering? (Which Lessons Can be Drawn From Empirical Studies into Alliancing and Partnering Success?) (October 15, 2009). Available at SSRN: https://ssrn.com/abstract=2294207 or http://dx.doi.org/10.2139/ssrn.2294207

Kamminga, P. (2019). Samenwerking en contractuele relaties? Dat kan je maar beter eerst ontwerpen-Samenwerking by design.

Kwawu, W., & Hughes, W. (2005, September). The impact of relational contracting on the construction industry. In Procs 21st Annual ARCOM Conference (pp. 7-9).

Laan, A., Noorderhaven, N., Voordijk, H., & Dewulf, G. (2011). Building trust in construction partnering projects: An exploratory case-study. Journal of Purchasing and Supply Management, 17(2), 98-108.

Lahdenperä, P. (2009). Project alliance: the competitive single target-cost approach.

Lahdenperä, P. (2012). Making sense of the multi-party contractual arrangements of project partnering, project alliancing and integrated project delivery. Construction management and economics,

Liang, R., Zhang, J., Wu, C., Sheng, Z., & Wang, X. (2019). Joint-venture contractor selection using competitive and collaborative criteria with uncertainty. Journal of Construction Engineering and Management, 145(2), 04018123.

Lloyd-Walker, B. M., Mills, A. J., & Walker, D. H. (2014). Enabling construction innovation: the role of a no-blame culture as a collaboration behavioural driver in project alliances. Construction Management and Economics, 32(3), 229-245.

Lloyd-Walker, B., & Walker, D. (2015, April). Collaborative project procurement arrangements.

Lo, W., Lin, C. L., & Yan, M. R. (2007). Contractor's opportunistic bidding behavior and equilibrium price level in the construction market. Journal of construction engineering and management, 133(6), 409-416.

Hutchinson, A., Creating successful collaboration; select the right partner; JMJ Associates, London

Mahalingam, A., Levitt, R. E., & Scott, W. R. (2005, February). Cultural clashes in international infrastructure development projects: which cultures matter?. In Proceedings.

Mandell, S., & Nyström, J. (2013). Too much balance in unbalanced bidding. Studies in Microeconomics, 1(1), 23-35.

Marrewijk, A. V., & Veenswijk, M. B. (2016). Changing institutional practices in the Dutch construction industry. International Journal of Project Organisation and Management, 8(1), 44-62.

McKinsey (2019). Toekomstige Opgaven Rijkswaterstaat.

Mitchell, S (2016) Behavioural assessment in procurement: A legal perspective. "Construction Manager", Feb, 36-39.

New Civil Engineer (2015) Debate, Behavioural assessments. Author: Hancock, M

Pansiri, J. (2007). How company and managerial characteristics influence strategic alliance adoption in the travel sector. International Journal of Tourism Research, 9(4), 243-255.

PIANOo,a, (2019), Procurement procedure

PIANOo,b, (2020), Selectiecriteria. nl/metrokaart/selectiecriteria

Portilla, I. P. (2020) (Masterscriptie) Assessing the capacity of contractors to collaborate and perform in a client-contractor team.

Projectteam DOEN (2017), Evaluatieverslag aanbesteding Project DOEN, Rijkswaterstaat

Puckett, K (2007) Have you got what it takes? "Building Magazine", 2 Feb. 42-45.

Rahman, M. M, Kumaraswamy, M. M., Ling, F. Y.,., & Phng, S. T. (2005). Constructing relationally integrated teams. Journal of Construction Engineering and Management, 131(10), 1076-1086.

Rahman, M., & Alhassan, A. (2012). A contractor's perception on early contractor involvement. Built Environment Project and Asset Management, 2(2), 217–233.

Rahmani, F., Khalfan, M. M., & Maqsood, T. (2016). Analysing the drivers for Early Contractor Involvement (ECI) adoption by construction clients. International Journal of Procurement Management, 9(4),

Rahmani, F., Khalfan, M., & Maqsood, T. (2013). The use of early contractor involvement in different countries. AUBEA, 2013, 1-10.

Rijkwaterstaat (2016). De Marktvisie. Retrieved from http://debouwcampus.nl/co-creatie-lab/vernieuwingsagenda/item/marktvisie

Ruijter, H. A. (2019). Resilient Partnership: An interpretive approach to public-private cooperation in large infrastructure projects.

Sarhan, S., Pasquire, C., King, A., & Manu, E. (2018). Institutional waste within the UK construction procurement context: A conceptual framework. Engineering Project Organization Journal, 8, 36-64.

Sødal, A. H. (2014) Early contractor involvement: advantages and disadvantages for the design team (Master's thesis, Norway, NTNU).

Slevin, D. P., Cleland, D. I., & Pinto, J. K. (2004). Innovations: Project management research 2004. Project Management Institute.

SOU, 2002. Fra˚n Byggsekt till Byggsektor. Stockholm, Byggkostnadsdelegationen, Suprapto, M., Bakker, H. L., & Mooi, H. G. (2015). Relational factors in owner–contractor collaboration: The mediating role of teamworking. International Journal of Project Management, 33(6)

Suprapto, M., Bakker, H. L., Mooi, H. G., & Hertogh, M. J. (2016). How do contract types and incentives matter to project performance?. International Journal of Project Management, 34(6), 1071-1087.

Scheepbouwer, E., & Humphries, A. B. (2011). Transition in adopting project delivery method with early contractor involvement. Transportation research record, 2228(1), 44-50.

Straatbeeld (2019). 'Selecteren op samenwerking kan, mag en lukt'

Stam, R (2016) Relationschip contracting Arangements (Thesis)

Ten Hoeve, G. (2018). Improving collaboration between client and contractor in integrated contracts in the Dutch construction sector.

Turton, T (2015) "Collaborative Delivery Framework – Presentation to CIHT – Better Procurement and Delivery Practice in the Highway Sector

TwynstraGudde (2016). Met het oog op goede samenwerking, Aanbesteden en contracteren

Välilä, T. (2019). An overview of economic theory and evidence of public-private partnerships in the procurement of (transport) infrastructure. Utilities Policy, 62, 100995

Van de Velde, D., Beck, A., Van Elburg, J., & Terschüren, K. H. (2008). Contracting in urban public transport. Amsterdam: European Commission, 123.

Van Riggelen, R. (2019) Bouwteam: For more collaboration in the construction industry

Van Wijck, D. (2018) Early Contractor Involvement in the Netherlands: The potential of ECI in public construction projects

Van Huuksloot, C. (2014) (Ballast Nedam) Possibilities for early contractor involvement in infrastructure projects in the Netherlands

Van der Geest, E., & Bloemendaal, R. (2011). Het Experiment A2 Hooggelegen - een geslaagde nieuwe weg. Den Haag: Veerkracht.

Van Os, A., Van Berkel, F., De Gilder, D., Van Dyck, C., & Groenewegen, P. (2015). Project risk as identity threat: Explaining the development and consequences of risk discourse in an infrastructure project. International journal of project management, 33(4), 877-888.

van Valkenburg, M., Lenferink, S., Nijsten, R., & Arts, J. (2008). Early contractor involvement:A new strategy for "Buying the best" in infrastructure development in the Netherlands.

Verschuren, P., & Doorewaard, J. (2015). Het ontwerpen van een onderzoek [Research design]. Amsterdam: Boom Lemma uitgevers.

Walker, D., & Hampson, K. (Eds.). (2008). Procurement strategies: A relationship-based approach. John Wiley & Sons.

Warsen, R., Nederhand, J., Klijn, E. H., Grotenbreg, S., & Koppenjan, J. (2018). What makes public-private partnerships work? Survey research into the outcomes and the quality of cooperation in PPPs. Public Management Review, 20(8), 1165-1185.

Wilkinson, S., & Shestakova, Y. (2006). Collaborative procurement on the rise.

Wondimu, P. A., Hailemichael, E., Hosseini, A., Lohne, J., Torp, O., & Lædre, O. (2016). Success factors for early contractor involvement (ECI) in public infrastructure projects. Energy Procedia. (NTNU).

Wondimu, P., Hailemichael, E., Svalestuen, F., Hosseini, A., Lohne, J., & Ldre, O. (2017). Implementation of early contractor involvement (ECI) in Norwegian bridge projects procurement.

Wondimu, Abebe (2019) Early Contractor Involvement (ECI) Approaches for Public Project Owners

Yin, R. K. (2011). Applications of case study research.

Appendixes

Appendix A. Case overview

Appendix A1 | Case overview

Case	Client	Type of project	Case report	Interview reports
CASE N1	Client A	Dyke improvement project	Appendix C1	Appendix D1, D2
CASE N2	Client B	Shore improvement project	Appendix C2	Appendix D3
CASE N3	Client C	Dyke improvement project	Appendix C3	Appendix D4
CASE N4	Client B	Bridge renovation	Appendix C4	Appendix D5
CASE N5	Client D	Dyke improvement project	Appendix C5	Appendix D6
CASE N6	Client E	Bridge renovation	Appendix C6	Appendix D7
CASE U1	Client F	Railway project	Appendix C7	Appendix D8, D9
CASE U2	Client G	Tunnel project	Appendix C8	Appendix D9, D8
CASE F1	Client H (I)	Tunnel project	Appendix C9	Appendix D10
CASE F2	Client I	City tramway project	Appendix C10	Appendix D11

Table 15 | Selected cases with the case and interview reports

Appendix B. Interview protocols

Appendix B1 | Interview protocol clients

This appendix includes the list of interview questions for the interviews with the client's representatives. The interview is set up as semi-structured, this means the interviewee has the freedom to elaborate on personal experiences on the assessment and further personal knowledge.

BASIC INFORMATION

- Collecting the basic information. This is needed for all cases. Some aspect below could be answered via email as well, for example by sending the procurement guide in advance.

Basic information (interviewee)

- Full name;
- Client/organisation;
- Function/role (when started and did this change during the project?);
- Years in the industry/working for this organisation;

Basic information (project)

- Project name and location;
- Project partners (consortium/alliance partners);
- Year initiative/procurement strategy started;
- Year procurement phase (market information) started;
- Year (expected) delivery (so the duration of collaboration is known);
- Intensity of collaboration [joint project management team? etc];

Basic information (procurement procedure)

- Project delivery model/contract/collaboration form;
- Procurement procedure (competitive dialogue etc);
- Used award criteria, it's definition and weight;
- Amount of bidding parties (per stage);
- Involved observers/assessors;

PART I

1. Which of the procurement award/selection criteria focused on 'collaboration' aspects?

2. Could you describe the motives to use this 'collaborative criteria'?

3. Which characteristics or goals of this project support/underline the use of 'collaborative criteria'?

4. Could you describe the process of coming to the procurement strategy of this project?

3a. Is there a guideline or previous project that lead to the use of 'collaborative criteria'?

5. Could you describe why the criterion has the current weighting ?

6. Could you describe the benefits/impact that were expected in advance, when applying this 'collaborative criterion'? (Client's team, client's organisation)

PART II

- The following questions focus on the team *assessment that was used to score the criterion.*

1. What aspects of client-contractor collaboration did you find important and were put forward in the team assessment? (Sub criteria, competencies, focus points of assessment)

2. What was the content of the assessment (business related cases/games/workshop etc) and why?

3. Could you describe the assessment in terms of the duration and course of the day?

4. Could you explain the decision regarding the assessment participants: who are going to assess?

5. What was the role of the assessors/observers during the assessment procedure?

6.. Could you describe the assessment-procedure (besides the assessment-days)?

PART III

- The following questions focus on the (personal) experiences/view of the interviewee

1. What were your experiences from the assessment-day? Was it according expectations?

2 In what way is the assessment linked to the collaborative relationship post-contract?

3. Could you reflect on the expected benefits/impact of this criterion and the client-contractor collaboration/match during the rest of the tender process and after award?

4. What are the underestimated or undervalued benefits and footnotes of applying these criteria?

- The following questions focus on the use of 'collaborative criteria' in general.

5. How do you see the potential and use of team assessments in the coming years?

6. How would you describe the domain of infrastructure projects that could benefit from using 'collaborative criteria' in the procurement phase? What do these projects have in common?

ENDING

Appendix B2 | Interview protocol assessors

This appendix includes the list of interview questions for the interviews with the five assessors. The interview is set up as semi-structured, this means the interviewee has the freedom to elaborate on personal experiences on the assessment and further personal knowledge.

BASIC INFORMATION

- Collecting the basic information. This is needed for all cases. Some aspect below could be answered via email as well, for example by sending the procurement guide in advance.

Basic information (interviewee)

- Full name;
- Firm/organisation;
- Function/role;
- Years in the industry;

Basic information (firm)

- Type of company;
- Number of assessors;
- Experience in these assessments;

PART I

1. What kind of award/selection criteria have you seen that focus on 'collaboration' aspects?

2. Could you describe the process of advising the client in developing the tender guideline?

3. Could you describe your role as assessor in this process?

4. What kind of instrument are there available to assess collaborative behaviour?

5. What is 'collaboration' or 'collaborative behaviour' in the context of your assessments?

[adjust interview to the experience/expertise of the assessor]

PART II

- The following questions focus on the team *assessment itself.*

1. How do you find out what competencies (sub-criteria) are relevant and justified to use?

2. What is the content of your assessment (business related cases/games/workshop etc) and why?

3. Could you describe the average assessment in terms of the duration and course of the day?

4. Could you explain the decision regarding the assessment participants: who are going to assess?

5. Would you consider inviting other professions for team-assessments and why?

6. What was your role during the assessment itself? Are there any observers involved?

7. Could you describe the way/method you assess and score the tenderers ?

8. Could you describe the standards and guidelines that are applicable?

PART III

1. How do you reflect on the value of team assessments in the broader context?

2 In what way is the assessment linked to the collaborative relationship post-contract?

3. Could you reflect on the expected benefits/impact of this criterion related to the client-contractor collaboration during the rest of the tender process and after award?

4. What are the underestimated or undervalued benefits and footnotes of applying these criteria?

- The following questions focus on the use of 'collaborative criteria' in general.

5. How do you see the potential and use of team assessments in the coming years?

6. How would you describe the domain of infrastructure projects that could benefit from using 'collaborative criteria' in the procurement phase? What do these projects have in common?

ENDING

Appendix C. Case findings

Restricted for online version

Appendix D. Interview reports

Restricted for online version